Match length realization and cooperation in indefinitely repeated games*

FRIEDERIKE MENGEL †
University of Essex
and Lund University

LUDOVICA ORLANDI‡
Nottingham Trent University

SIMON WEIDENHOLZER§
University of Essex

December 3, 2021

Abstract

Experimental studies of infinitely repeated games typically consist of several indefinitely repeated games ("matches") played in sequence with different partners each time, whereby match length, i.e. the number of stages of each game is randomly determined. Using a large meta data set on indefinitely repeated prisoner's dilemma games (Dal Bó & Fréchette, 2018) we demonstrate that the realized length of early matches has a substantial impact on cooperation rates in subsequent matches. We estimate simple learning models displaying the "power law of practice" and show that participants do learn from match length realization. We then study three cases from the literature where realized match length has a strong impact on treatment comparisons, both in terms of the size and the direction of the treatment effect. These results have important implications for our understanding of how people learn in infinitely repeated games as well as for experimental design.

Keywords: Experiments, Indefinitely Repeated Games, Cooperation, Social Dilemmas .

JEL Classification Numbers: C70, C90.

*We thank Maria Bigoni, Tilman Boergers, Anna Dreber, Matt Embrey, Guillaume Frechette, Dan Friedman, Drew Fudenberg, Ed Hopkins, Tatiana Kornienko, Volodymyr Lugovskyy, Ryan Oprea, Ronald Peeters, Karl Schlag, Giancarlo Spagnolo, Robert Sugden, Emanuel Vespa, Leeat Yariv, Sevgi Yuksel, an anonymous associate editor and two anonymous Reviewers for helpful comments and/or for providing data access. We also thank audiences at the LEG 2019 workshop at Bar Ilan, University of California Davis, the University of East Anglia, the University of Edinburgh, the University of Heidelberg, Oxford University, the SAET 2019 conference, the VIBES seminar series and at the Vienna University of Economics and Business Adminstration for their helpful comments. Financial support by the British Academy (SG162637) is gratefully acknowledged.

†Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ *and* Department of Economics, Lund University, Tycho Brahes vag 1, Lund (SE). *e-mail*: fr.mengel@gmail.com.

‡Department of Economics, Nottingham Trent University, Nottingham NG1 4FQ. *e-mail*: ludovica.orlandi@ntu.ac.uk.

§Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ. *e-mail*: sweide@essex.ac.uk.

1 Introduction

Infinitely repeated games are of enormous importance in many areas of Economics, but also in Politics, Sociology, Biology and many other subjects. The theory of infinitely repeated games has delivered important insights into how repeated interaction changes incentives and how it can enable a wider array of outcomes including, for example, cooperation in social dilemmas. However, it does not always offer sharp predictions. For instance, in the prisoner dilemma both cooperation and defection are equilibrium actions provided that players are sufficiently patient. The multiplicity of outcomes gives an important role to empirical research on infinitely repeated games to narrow down what we can expect empirically in these games. This research has provided key insights on e.g. the determinants of cooperation in social dilemmas, the role of monitoring, or the differences between discrete and continuous time.[1]

Empirical studies of infinitely repeated interactions rely on the equivalent setting of "indefinitely repeated games" using a random continuation probability, as originally proposed by Roth and Murnighan (1978). After every round of play, there is a fixed known probability δ that the game continues for an additional round and a probability $1-\delta$ with which it ends. The length T of the repeated game is hence a random variable with expected value $\mathbb{E}[T] = \frac{1}{1-\delta}$ and standard deviation $\sqrt{\frac{\delta}{(1-\delta)^2}}$. Often researchers observe a subject in more than one indefinitely repeated game. Each of these games is then referred to as a "match" and the researcher consequently observes not just one realization of the random variable T, but a sequence of such "match length realizations". The researcher can make accurate inference from these observations on the underlying infinitely repeated game if either the realized sequences of match length realizations "correctly" represent the infinitely repeated game or if match length realizations are irrelevant for behaviour.[2] Achieving the former can be difficult as there are considerable practical difficulties involved in getting a large enough sample of different sequences of match length realizations.[3] Given that the number of match length realizations is often going to be small in practice, it is crucially important to understand whether match length realization influences behaviour.

In this paper we first demonstrate that the sequence of match length realizations has a substantial, robust and highly statistically significant effect on behaviour. Using a large data set from lab experiments studying the infinitely repeated prisoner dilemma compiled by Dal Bó and Fréchette (2018) we show that when participants initially experience relatively long matches subsequent cooperation rates are substantially higher. Specifically, when most matches in the first part of an experiment are "long" (above theoretical median length), then cooperation rates are 44% higher in subsequent matches. Intuitively, participants who experience longer matches become more optimistic about the relative benefits of conditionally cooperative strategies and cooperate more. Moreover, by comparing the impact of long matches in the first third to the

[1]See Dal Bó and Fréchette (2018) for a review of this extensive literature.

[2]In standard theory only expected match length should matter for behavior. Hence, according to standard theory match length realizations should indeed be irrelevant for behaviour.

[3]We discuss some of these difficulties in detail in Section 4.

impact of long matches in the middle third of an experiment on cooperation in the final third, we demonstrate that the effect of early matches is as least as important as the effect of recent ones. This observation is consistent with the "power law of practice" which describes the phenomenon of initially steep- and then flattening out- learning curves. We develop and estimate a simple reinforcement learning model with counterfactuals (Erev & Roth, 1998) that allows us to distinguish learning from match length realizations from learning given the expected match length. We find that indeed learning from match length realization is important.

Our results show that the environment in which early interactions take place matters for subsequent interactions as people learn from match length realizations. While match length realizations and similar stochastic factors are most likely also important factors in the field, they are difficult to isolate from other confounds. The advantage of experimental data is that it allows to separate the effects found in the current paper from alternative explanations. Settings where match length realization is likely to matter include e.g. "cultural differences" between people coming from different work environments (characterized by more or less turnover) or from different social backgrounds. The effects identified in this paper likely apply in many other settings where people learn from stochastic realizations of payoff relevant variables.[4] They also have implications for evidence based policy making. If a policy (e.g. designed to increase cooperative behavior) is evaluated over a certain fixed period, it is possible that the results of the evaluation are affected by early match length realizations even if they are exogenous to the policy evaluated.

Our findings provide valuable insights into how people learn in indefinitely repeated games. As such they can inform the development of new theories of learning in games. They also have important methodological implications for the design of empirical studies. The length of each match is typically drawn at the session level in experiments, meaning that all subjects in a given session experience the same sequences of match length draws. In fact, all papers in the Dal Bó and Fréchette (2018) meta study use this or a very similar design. The number of different sequences of match length realizations for a given treatment ranges between 1 and 10 across the different papers contained in the meta-study. Given our results discussed above we would expect that - with such small numbers of match length realizations - treatment comparisons can be affected. We provide three case studies of papers from the existing literature, which were not part of the Dal Bó and Fréchette (2018) meta study and go beyond the prisoner's dilemma: a continuous time prisoner's dilemma (Bigoni, Casari, Skrzypacz, & Spagnolo, 2015), a public good game (Lugovskyy, Puzzello, Sorensen, Walker, & Williams, 2017) and oligopoly games (Embrey, Mengel, & Peeters, 2019). We show that - for each of them - treatment effects differ depending on match length realization. We also run our own experiments and show that in some cases the conclusions drawn from the research might have been different for different match length realizations.

Our paper contributes to a substantial and active literature on indefinitely repeated games, much of it summarized by Dal Bó and Fréchette (2018). Several researchers have documented

[4] For instance, in common value auctions (as e.g. experimentally studied in Kagel and Levin (1986)) early realizations of common values and of public information could potentially influence subjects' subsequent bidding strategies through learning. Likewise, different realizations of states of the world in experimental financial markets (as in settings similar to Plott and Sunder (1982)) may alter subjects' beliefs on the probability with which states occur and impact trading strategies.

a positive effect of the length of the immediately preceding match on cooperation (see e.g. Camera and Casari 2009, Dal Bó and Fréchette 2011; 2018, Fréchette and Yuksel 2017, Bernard, Fanning, and Yuksel 2018).[5] In the context of infinitely repeated trust games, Engle-Warnick and Slonim (2006) find some evidence that there is more trust and trustworthiness in sessions that initially featured long matches as compared to sessions starting out with short ones. As they observe, this gap could have been due to individual subject or session effects since there was already more trust and trustworthiness in the beginning of the initially long sessions. We add to this literature by providing the first comprehensive analysis of the long lasting effects of (the entire sequence of) match length realization on cooperation in infinitely repeated social dilemmas. To the best of our knowledge our paper is also the first to demonstrate that people learn from match length realization using structural estimation of learning models.[6] We also advance the existing literature by discussing in detail the potential implications of these findings for measuring cooperation levels and for the design of empirical studies.

The paper is organized as follows. In Section 2 we demonstrate the main empirical finding of an effect of match length realization on cooperation using the Dal Bó and Fréchette (2018) meta study . Section 3 contains our discussion of the case-studies. We discuss implications of our results in Section 4 and Section 5 concludes. Additional theory, tables, figures and information on our own experiments can be found in an Appendix.

2 Match Length Realization and Cooperation

2.1 The prisoner's dilemma

We consider agents who play a 2×2 indefinitely repeated prisoner's dilemma like the one illustrated in the left panel of Figure 1. Payoffs satisfy $\mathtt{T} > \mathtt{R} > \mathtt{P} > \mathtt{S}$ and $\mathtt{T} + \mathtt{S} < 2\mathtt{R}$ such that mutual defection is the only Nash equilibrium of the stage game but mutual cooperation maximizes joint payoffs. Following Dal Bó and Fréchette (2018) we can normalize payoffs so that we only have two parameters, see middle panel of figure 1. The continuation probability δ indicates the probability with which the game continues for one more round. The number of stages in the indefinitely repeated game is hence a random variable T. It is common in modern experiments to play several such indefinitely repeated games. Usually participants are rematched at the end of one repeated game and play a new game with a new partner. Each such repeated game is often referred to as a "match". Typical experiments differ in the number M of such matches implemented, the expected length of a match (given by $\mathbb{E}[T] = \frac{1}{1-\delta}$) as well as the realized match length. We index the round of play within a match by t and the match by m. T^m is the realized match length of match m, i.e. the number of stages in match m.

A substantial experimental literature has studied how payoff parameters affect cooperation in the prisoner's dilemma.[7] One particularly successful approach, proposed by Blonski

[5] A similar positive effect is document for the behaviour of the previous opponent, in the sense that subjects are more likely to cooperate when they have been previously matched with somebody starting out with cooperation.

[6] A number of papers estimate learning models in repeated games (see e.g. Erev and Roth (2001), Hanaki, Sethi, Erev, and Peterhansl (2005), Dal Bó and Fréchette (2011), Ioannou and Romero (2014) or Embrey, Fréchette, and Yuksel (2018)), but they do not estimate parameters capturing the extent to which people learn from match length realization.

[7] See e.g. Embrey, Fréchette, and Yuksel (2017) and Mengel (2018) for contributions analyzing finitely repeated

	C	D
C	R	S
D	T	P

PD game.

	C	D
C	$\frac{\mathtt{R}-\mathtt{P}}{\mathtt{R}-\mathtt{P}} = 1$	$\frac{\mathtt{S}-\mathtt{P}}{\mathtt{R}-\mathtt{P}} = -\ell$
D	$\frac{\mathtt{T}-\mathtt{P}}{\mathtt{R}-\mathtt{P}} = 1+g$	$\frac{\mathtt{P}-\mathtt{P}}{\mathtt{R}-\mathtt{P}} = 0$

Normalized game.

	GT	AD
GT	$\mathbb{E}[T]$	$-\ell$
AD	$1+g$	0

GT vs. AD game.

Figure 1: Left: Prisoner's dilemma (PD) game with payoff parameters $\mathtt{T} > \mathtt{R} > \mathtt{P} > \mathtt{S}$ and $\mathtt{T} + \mathtt{S} < 2\mathtt{R}$. Middle: Normalized game where joint defection payoff $\mathtt{P}$ is subtracted from each cell and all payoffs are divided by $\mathtt{R} - \mathtt{P}$ (difference between mutual cooperation and defection payoffs). Right: Payoffs in the game induced by Grim-Trigger and Always Defect. GT played against GT yields a payoff of 1 in all $\mathbb{E}[T]$ stages. AD (GT) played against GT (AD) yields once a payoff of $1 + g$ $(-l)$ and zero in the remaining $\mathbb{E}[T] - 1$ stages.

and Spagnolo (2015), analyzes a setting where agents can only choose among the repeated game strategies "Grim-Trigger" (GT) and "Always Defect" (AD) and payoffs are given by the expected sum of payoffs of the induced indefinitely repeated game shown in the right panel of Figure 1 (see also Blonski, Ockenfels, and Spagnolo (2011) and Dal Bó and Fréchette (2011)). Provided "Grim-Trigger" can sustain cooperation in a subgame perfect Nash equilibrium ($\mathbb{E}[T] \geq 1 + g$), the resulting game constitutes a coordination game. The size of the basin of attraction of AD, denoted by `SizeBAD`, is defined as the threshold probability of choosing GT that has to be exceeded to make GT a best response.[8] Formally,

$$\mathtt{SizeBAD} = \begin{cases} 1 & \text{if } \mathbb{E}[T] < 1 + g \\ \frac{\ell}{\mathbb{E}[T]+\ell-g-1} & \text{otherwise} \end{cases} \tag{1}$$

Note that `SizeBAD` is decreasing in $\mathbb{E}[T]$ (respectively δ), conveying the intuitive idea that cooperation is easier to sustain under longer expected match durations. Dal Bó and Fréchette (2018) show that `SizeBAD` indeed predicts cooperation rates very well in a meta-study of indefinitely repeated prisoner's dilemma experiments. They also show that the length of the immediately preceding match has an effect on cooperation rates in the subsequent match. They suggest that this is either due to a minority of participants who may not understand how match lengths are determined or due to how participants update their overall evaluation of the value of cooperation through experience. They write "there is an interesting - as yet unexplored - question regarding the way that humans learn in infinitely repeated games. Is the impact of the realized length constant throughout or is the impact more important early on?"

As we will see below not only does the length of the immediately preceding match matter, but the entire sequence of match length realizations is important. Further, addressing the question posed by Dal Bó and Fréchette (2018), the impact of realized match length is not constant throughout. Early matches matter at least as much as later matches and sometimes more. We will now demonstrate these patterns empirically (Section 2.2) and then estimate simple learning models to understand how people learn in indefinitely repeated games (Section 2.3).

PD games and Dal Bó and Fréchette (2018) for a survey of the literature on the indefinitely repeated version.

[8]This corresponds to the probability of *AD* in the mixed strategy equilibrium of the game in the right panel of Figure 1.

2.2 The Effect of Match Length Realization

To study empirically whether there exists a persistent effect of match length in early matches we use the data collected by Dal Bó and Fréchette (2018). They collected data from 141 different sessions of indefinitely repeated prisoner's dilemma experiments with 2415 participants (see Table 3 in Dal Bó and Fréchette (2018)). Some of these sessions are one-shot games ($\delta = 0$), though. In our analysis we will rely on 103 sessions with $\delta > 0$. In all papers contained in their data set the sequences of match length realizations is drawn at session level, i.e. all subjects in a given session faced the same sequence of match length realizations. Figure 2 shows the distribution of the difference between theoretical median match length and realized match length in the meta-study.[9] The left panel aggregates games with different discount factors. It can be seen that match lengths are, as expected, concentrated around the median with a good amount of variation on both sides. The right panel shows separate graphs for the three most common discount factors $\delta = 0.5, 0.75$ and $\delta = 0.9$. The figure shows that for the longer games with $\delta = 0.75$ and $\delta = 0.9$ most matches are somewhat shorter than what we would expect. However, in all cases, there is a good amount of variation.

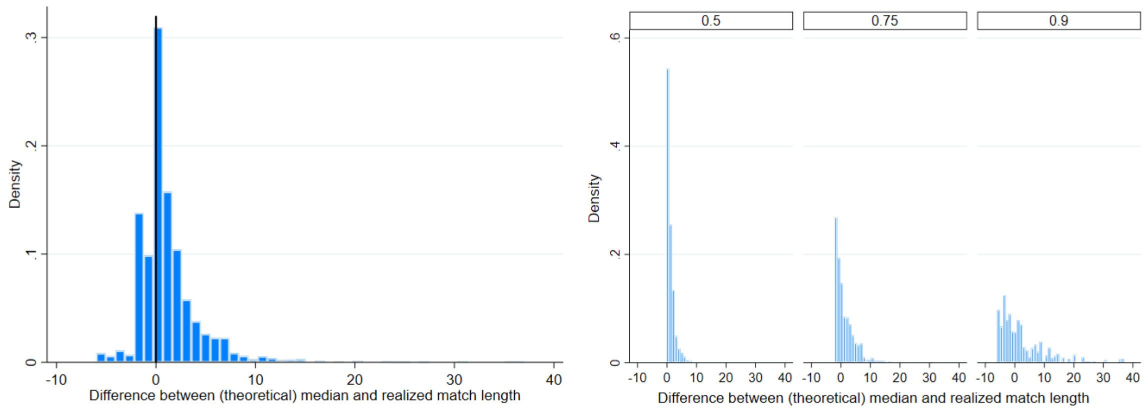

Figure 2: Distribution of the difference between theoretical median match length and realized match length overall (left panel) and separately for $\delta = 0.5, 0.75$ and $\delta = 0.9$.

We use this variation to study how match length realization in early matches affects subsequent cooperation. We define early matches as the 1st third of matches in a session and create a dummy variable Δ^{1st}_{above} indicating whether more than $\frac{2}{3}$ out of these early matches were (weakly) longer than the theoretical median length.[10] The dummy takes the value 1 in 44% of sessions. Analogously, we can also define dummies Δ^{2nd}_{above} and Δ^{3rd}_{above} which take the value 1 in 42% and 47% of sessions, respectively.

Table 1 shows the results of regressing cooperation in the final third of matches on Δ^{1st}_{above} as

[9] Appendix Figure E.1 shows kernel density estimates.

[10] The reasoning behind these choices is the following. We split matches in three groups (early, middle and late) rather than e.g. two is that it allows us to compare the effect of early (1st third) and middle (2nd third) matches on cooperation in late (3rd third) matches. This allows to address the question whether early experience or recent experience is more important for cooperation. Appendix Table D.1 shows results for alternative splits. The reason we use a dummy is that (i) theoretical medians differ with δ, which means that we cannot just use match length directly, and that (ii) it makes regression results more easily interpretable. Appendix Table D.5 shows results when we use the share of matches above median instead. Last, the reason that we use $\frac{2}{3}$ as a cutoff for the share of long matches is that it produces relatively balanced groups, though some other cutoffs would have produced that too. Appendix Tables D.6-D.7 show the results with alternative cutoffs.

	Effect of Match Length Realization on subsequent cooperation					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ^{1st}_{above}	0.142***	0.101**	0.226***	0.125**	0.101**	0.207***
	(0.053)	(0.039)	(0.056)	(0.058)	(0.041)	(0.062)
Δ^{2nd}_{above}				0.085	0.069*	0.032
				(0.060)	(0.039)	(0.068)
`SizeBAD`		-0.765***	-0.539***		-0.809***	-0.659***
		(0.069)	(0.094)		(0.068)	(0.149)
`SizeBAD` $\times$ Δ^{1st}_{above}			-0.296***			-0.241**
			(0.094)			(0.104)
`SizeBAD` $\times$ Δ^{2nd}_{above}						0.048
						(0.144)
Constant	0.321***	0.974***	0.747***	0.294***	0.994***	0.844***
	(0.028)	(0.074)	(0.098)	(0.032)	(0.074)	(0.153)
δ f.e.	NO	YES	YES	NO	YES	YES
Test $\Delta^{1st}_{above} = \Delta^{2nd}_{above}$	-	-	-	0.6903	0.6063	0.0989
Observations	34,319	34,319	34,319	18,536	18,536	18,536
R-squared	0.021	0.219	0.223	0.034	0.251	0.255

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1: Columns (1)-(3): Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by dummy Δ^{1st}_{above} indicating whether more than $\frac{2}{3}$ of matches in the 1st third of the experiment were longer than the theoretical median match length. Columns (4)-(6): Initial (first stage) cooperation rate in the 3rd third of matches explained by dummies Δ^{1st}_{above} and Δ^{2nd}_{above}. Standard errors clustered at session level. Observations stem from 103 sessions spread across 11 papers.

well as `SizeBAD` and an interaction. Early match length has a substantial impact on cooperation in later matches. If at least $\frac{2}{3}$ of these early matches are "long", then cooperation rates are higher for the remainder of the experiment as shown by the positive coefficient on the dummy Δ^{1st}_{above} in column (1). The effect size is substantial, with cooperation rates being 44% higher when initial matches were long as compared to when they were short. As expected, the table also shows a negative impact of `SizeBAD` on cooperation rates (columns (2)-(3)). Interestingly, there is also an interaction effect between `SizeBAD` and Δ^{1st}_{above}. If early matches are long than the detrimental effect of `SizeBAD` is more pronounced. This is intuitive as longer early matches could allow participants to better learn the incentives coming from the game parameters. Conversely, the interaction term also shows that the positive effect of early match length realization is stronger the more favorable the climate is for cooperation. In fact according to Table 1, the effect is positive if and only if `sizeBAD` is smaller than 0.77.

Robustness Appendix D.1 contains tables showing that these results are qualitatively robust to the inclusion of paper fixed effects (Table D.4), to considering different thresholds (Tables D.6-D.8) or to using the share of matches above median instead of a dummy variable (Table D.5). A placebo test shown in Appendix Table D.3 where we regress cooperation in the 1st third of matches on Δ^{3rd}_{above} shows that the results in Table 1 are fundamental and not e.g. driven by correlations of match lengths within sessions or observed or unobserved heterogeneity across papers or treatments, e.g. caused by different ways researchers implement match length draws.[11]

[11] We would not expect realized match length of final matches, which have not yet been played, to affect cooperation in the beginning of the experiment. Hence we would expect zero coefficients on Δ^{3rd}_{above} and the corresponding

Early vs Recent Matches Next we ask what is more important for cooperation in the final third of the experiment, early experience, i.e. match length in the 1st third, or recent experience, i.e. match length in the 2nd third of matches? Columns (4)-(6) in Table 1 show the results of regressing cooperation in the final third of matches on both dummies Δ^{1st}_{above} and Δ^{2nd}_{above}. The table shows very clearly that early experience in the 1st third of matches is very important. In all specifications the coefficient on Δ^{1st}_{above} is at least as large as that for Δ^{2nd}_{above} and exhibits a higher level of statistical significance. The interaction effect with `SizeBAD` is also more important for these matches. Early matches seem at least as important as recent matches and potentially, more important.[12]

Experience Does the effect vanish with more experience, i.e. if enough matches are played in the experiment? To answer this question we rerun specification (1) of Table 1 restricting the sample to sessions with (i) at most 12 matches in total, (ii) 12-24 matches, (iii) 24-36 matches etc.[13] Appendix Figure E.2 shows that a positive effect size can be found even in sessions that feature at least 72 matches in a session. The figure also shows a possible downward trend in coefficient sizes as more matches are played, but if at all the trend is slow and suggests that at the very least 80 matches would have to be played in a session for the coefficient to vanish. This can quickly become infeasible especially if the discount factor δ is high. Note also that there is a compositional effect in this analysis as sessions with more matches tend to have smaller δ in the meta-study (t-test, $p < 0.0001$). As with the geometric distribution we would expect more extreme outliers in match length realizations when δ is higher, the compositional effect should artificially exacerbate the effect of experience, i.e. make it seem that with more matches there is less of an effect of match length realization. That we see very little in terms of a downward trend despite this suggests that adding more matches will not easily eliminate the impact of early match length realizations.

To sum up, the results in this section have shown that there can be substantial and non-trivial effects of realized early match length on cooperation rates in the rest of the experiment. Hence, which match length realizations are drawn can potentially affect research results. This is particularly likely if few draws are made (e.g. only one draw per session or treatment). In Section 3 we will study three case studies highlighting this point.

We have also seen that early matches matter at least as much as recent matches. This is in line with a substantial body of evidence on both human and animal learning which shows that learning curves tend to be steeper initially and then flatter. This observation is known as "power law of practice" and according to Erev and Roth (1998) dates back to at least Blackburn (1936). In the next subsection we will structurally estimate learning models that do have this property and show that match length realization matters for learning.

interaction term. We do indeed find that these coefficients are close to zero and statistically not significant.

[12]Appendix Table D.2 compares the importance of early and recent matches for more different splits. Specifically the table compares the impact of match length realization in the first X-th, second X-th, third X-th...of matches on cooperation in the last X-th of matches, where X ranges from 2,...,10. For all X=2,...,9 the coefficient on the first X-th of matches is larger than that of the (X-1)th Xth of matches.

[13]We choose multiples of 12 to cut the sample as (i) they are close to the 25th, 50th and 75th percentile of match numbers in the overall sample (25th percentile is 11, 50th is 23 and 75th is 34) and (ii) 12 divides by 2, 3 and 4 without remainder allowing us to split the total number of matches in halfs, thirds and quarters as in Appendix Figure E.2.

2.3 Learning

We study simple learning models which display the "power law of practice". These are straightforward adaptations of previously studied models to an environment where payoffs depend on stochastic realizations of match length. To this end, we consider a set of agents which are recurrently matched to play a series of indefinitely repeated PD games. Following much of the literature (Dal Bó & Fréchette, 2018; Embrey et al., 2017), we restrict attention to the GT and AD strategies.[14] The payoffs of the game induced by these strategies are given in the right panel of figure 1, where the expected match length $\mathbb{E}[T]$ is now replaced by the actual match length realization T^m.

Model and Identification

In our learning model the choices of agents are determined by propensities which are updated after each match. Propensities can be interpreted as beliefs (as in fictitious play, see e.g. Mookherjee and Sopher 1997) but can also incorporate a much wider set of feelings, such as e.g. familiarity or habituation (as in reinforcement learning, see e.g. Erev and Roth 1998, Boergers and Sarin 1997). Each agent i is endowed with an initial propensity for each strategy, denoted by $\psi_s^{i,0}$ for strategies $s \in \{GT, AD\}$, which may capture pre-game experience, initial inclinations or beliefs. In this section we focus on *reinforcement learning with counterfactuals* (see e.g. Vriend 1997 and Rustichini 1999 and the special cases in Erev and Roth 1998 and Camerer and Ho 1999) where propensities $\psi_s^{i,m}$ for all strategies evolve according to

$$\psi_s^{i,m+1} = \psi_s^{i,m} + \pi(s, s^{-i,m}, T^m).$$

$s^{-i,m}$ denotes the strategy of i's opponent in match m and $\pi(s, s^{-i,m}, T^m)$ gives the payoff earned with strategy s in this case. It remains to specify a choice rule. We assume that i's probability to choose grim trigger in match m is given by the logit choice rule

$$p^{i,m} = \frac{e^{\lambda \psi_{GT}^{i,m}}}{e^{\lambda \psi_{GT}^{i,m}} + e^{\lambda \psi_{AD}^{i,m}}}, \tag{2}$$

where λ is a measure of noise, sometimes also thought of as a measure of rationality of the economic agent. If $\lambda = 0$, then $p^{i,m} = 0.5$, i.e. the agent chooses randomly with uniform probability between GT and AD. By contrast, if $\lambda \to \infty$, then the agent chooses the strategy with the higher propensity with probability 1.[15]

To estimate this model we first rewrite the updating rule recursively as

$$\psi_s^{i,m+1} = \psi_s^{i,0} + \pi^{i,1}(s, s^{-i,1}, T^1) + \ldots + \pi^{i,m}(s, s^{-i,m}, T^m). \tag{3}$$

[14] A theoretical justification for why it is sensible to restrict to these strategies is provided in Blonski and Spagnolo (2015) and Blonski et al. (2011).

[15] In environments were payoffs are stable, in the sense that they do not feature an exogenous stochastic element, reinforcement learning with counterfactuals is closely related to (smooth) fictitious play (see e.g. Fudenberg and Kreps 1993 and Fudenberg and Levine 1998) where agents play a (smooth) best response to the belief that future play will follow the past empirical distribution (see e.g. Cheung and Friedman 1997 and Camerer and Ho 1999). The equivalence holds because looking back to previous earnings of strategies is equivalent to forming beliefs based on past behaviour and then computing expected payoffs based on these beliefs. In Appendix A we show the results of simulations where we include also reinforcement learning without counterfactuals and fictitious play and where we also use the linear choice rule instead of the logit choice rule. We find that all models perform very similarly in the simulations.

If we then denote by $\Delta\Pi^m = \sum_{h=1}^{m} \pi^{i,h}(GT, s^{-i,h}, T^h) - \pi^{i,h}(AD, s^{-i,h}, T^h)$ the cumulative payoff difference between grim trigger and always defect across all past matches, we can re-write (2) as

$$p^{i,m} = \frac{e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \Delta\Pi^m]}}{1 + e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \Delta\Pi^m]}}. \tag{4}$$

Denote by $y_{GT}^{i,m} = 1$ the outcome where player i chooses grim trigger in match m and by $y_{GT}^{i,m} = 0$ where they do not. Given the choice rule above, the likelihood function associated with this binary outcome then takes the following form

$$L(\psi_{GT}^{i,0}, \psi_{AD}^{i,0}, \lambda) = \prod_{i=1}^{N} \prod_{m=1}^{M} \left(\frac{e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \Delta\Pi^m]}}{1 + e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \Delta\Pi^m]}} \right)^{y_{GT}^{i,m}} \left(\frac{1}{1 + e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \Delta\Pi^m]}} \right)^{1 - y_{GT}^{i,m}} \tag{5}$$

Estimating this model using logistic regression we have that the coefficient for the constant term of the regression is an estimate of $\lambda(\psi_{GT}^{i,0} - \psi_{AD}^{i,0})$ and the coefficient on $\Delta\Pi^m$ is an estimate of λ. Comparing these coefficients hence allows us to identify λ and the difference in initial propensities $(\psi_{GT}^{i,0} - \psi_{AD}^{i,0})$.

Extended Model The standard model does not distinguish between learning about the behaviour of opponents - given expected match length - and learning about match length realizations. To allow us to do so, we extend the standard model and consider the following updating rule.

$$\psi_s^{i,m+1} = \psi_s^{i,m} + \alpha \pi^{i,m}(s, s^{-i,m}, T^m) + (1 - \alpha)\pi^{i,m}(s, s^{-i,m}, \mathbb{E}[T]). \tag{6}$$

This nests the standard model when $\alpha = 1$. If $\alpha = 0$, then learning about match length realization (T^m) does not matter and all learning relies on expected match length. The higher α the relatively more important is learning about match length realizations. In analogy to above we define $\Delta\Pi_{\mathbb{E}[T]}^m = \sum_{h=1}^{m} \pi^{i,h}(GT, s^{-i,h}, \mathbb{E}[T]) - \pi^{i,h}(AD, s^{-i,h}, \mathbb{E}[T])$ as the cumulated payoff difference between grim trigger and always defect across all past matches conditional on the opponent's choice and assuming that match length is fixed at the expected match length $\mathbb{E}[T]$. The likelihood function for the extended logit choice model can then be written as follows

$$L(\psi_{GT}^{i,0}, \psi_{AD}^{i,0}, \lambda, \alpha) = \prod_{i=1}^{N} \prod_{m=1}^{M} \left(p^{i,m}\right)^{y_{GT}^{i,m}} \left(1 - p^{i,m}\right)^{1 - y_{GT}^{i,m}}$$

with

$$p^{i,m} = \frac{e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \alpha\Delta\Pi^m + (1-\alpha)\Delta\Pi_{\mathbb{E}[T]}^m]}}{1 + e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \alpha\Delta\Pi^m + (1-\alpha)\Delta\Pi_{\mathbb{E}[T]}^m]}} \tag{7}$$

The log-likelihood is given by

$$LL(\psi_{GT}^{i,0}, \psi_{AD}^{i,0}, \lambda, \alpha) = \sum_{i=1}^{N} \sum_{m=1}^{M} y_{GT}^{i,m} \ln\left(p^{i,m}\right) + (1 - y_{GT}^{i,m}) \ln\left(1 - p^{i,m}\right)$$

Estimating this model using logistic regression we have that the coefficient for the constant term of the regression is an estimate of $\lambda(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$. The coefficient on the difference $\Delta\Pi^m$ is an estimate of $\lambda\alpha$ and the coefficient on $\Delta\Pi^m_{\mathbb{E}[T]}$ of $\lambda(1-\alpha)$. Comparing these coefficients allows us to identify λ, α and the difference in initial propensities $(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$. The parameter we are most interested in here is α, which will tell us to which extent participants learn about match length realization as opposed to using expected match length to make their decisions. Note that our updating rule imposed the constraint that $\alpha \in [0,1]$. As the unconstrained estimates return values of $\alpha \in [0,1]$ we report unconstrained estimates below.

Our parameters of interest α, λ and $(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$ are all estimated at the population level. Wilcox (2006) has shown, though, that in pooled estimations of learning models that contain lagged variables (such as the ones above), the presence of heterogeneity e.g. in the parameter λ can cause biases in the estimated parameters (see also Cabrales and Garcia-Fontes (2000)). Specifically, Wilcox (2006) is concerned with the estimation of a parameter that measures the extent to which learners take into account counterfactuals when updating their propensities. A downward bias in this parameter leads to overestimation of the role reinforcement learning without counterfactuals plays as opposed to models (like belief learning) which take counterfactuals into account. All the learning models we estimate do take counterfactuals fully into account. Hence our estimates are not affected by this particular issue.[16] To study whether similar issue could plague our parameters of interest, most importantly α, we conduct extensive Monte Carlo studies which we report in detail in Appendix B, where we also derive the maximum likelihood estimator of α. Those studies show that heterogeneity in λ can indeed cause downward biases of λ of up to 5% for the most relevant case when λ is assumed to be in the range found in Table 3. Most importantly, they also show that estimates of α are unbiased both when heterogeneity in λ is introduced as well as when heterogeneity in α is assumed in the data generating process. For details see Appendix B.

Estimation Results

Table 2 shows the parameter estimates for the basic model. Two patterns emerge. First, in terms of the difference between initial propensities we see that participants favour AD over GT ($(\psi^{i,0}_{GT} - \psi^{i,0}_{AD}) < 0$) especially when the horizon is short ($\delta = 0.5$). This is in line with the intuition that a longer horizon makes more cooperative strategies more attractive. Second, there is a relatively large degree of noise with $\lambda \approx 0.027$ on average. Noise decreases somewhat as δ increases. To provide some context to these numbers we note that given the average values of $\Delta\Pi^m$ these estimates together with choice rule (2) imply an $\approx$ 59% chance of picking the "correct" strategy, i.e. the strategy with the higher value of $\Delta\Pi^m$, on average across all matches m. If we focus on the second half of matches in a session, then this value increases to about 64%, but it is well below 100% in all cases.

Note that the accumulated payoff differences $\Delta\Pi^m$ will tend to be strongly correlated with `SizeBAD` (across all supergames $\rho = -0.7213^{***}$). Hence one might wonder whether controlling for `SizeBAD` might reduce bias in the estimation and/or improve the precision of estimates. Of course doing so makes the interpretation of our parameters of interest in terms of the learning

[16] We are not interested in comparing learning models with and without counterfactuals. Still, simulations reported in Appendix A show that these types of models perform very similarly with our data.

	Parameter Estimates Basic Model			
	All Data	$\delta = 0.5$	$\delta = 0.75$	$\delta > 0.5$
λ	0.027 [0.026, 0.028]	0.020 [0.020,0.021]	0.034 [0.033,0.035]	0.033 [0.033, 0.034]
$(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$	-1.460 [-1.576, -1.341]	-8.700 [-9.459, -7.941]	-0.676 [-0.700, -0.652]	-1.242 [-1.366, -1.118]
N	37394	16088	18136	21306

Table 2: Parameter estimates basic model for all data and separately for $\delta = 0.5$, $\delta = 0.75$ and $\delta > 0.5$. 95% confidence interval in brackets.

model less straightforward. If we nevertheless control for sizebad we get an estimated value of $\lambda = 0.019$ (confidence interval $[0.019, 0.020]$) when using all data as in column (1) of Table 2.[17] This corresponds to an about 55% chance of picking the correct strategy across all matches which increases to 60% in the second half of matches.

	Parameter Estimates Extended Model			
	All Data	$\delta = 0.5$	$\delta = 0.75$	$\delta > 0.5$
α	0.432 [0.423, 0.440]	0.486 [0.484,0.488]	0.589 [0.573, 0.605]	0.631 [0.605, 0.657]
λ	0.186 [0.165, 0.207]	0.603 [0.533, 0.673]	0.172 [0.145,0.200]	0.117 [0.096, 0.138]
$(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$	-0.768 [-0.913, -0.623]	-0.486 [-0.578, -0.394]	-0.191 [-0.384, 0.001]	-0.708 [-0.989, -0.426]
N	37394	16088	18136	21306

Table 3: Parameter estimates extended model for all data and separately for $\delta = 0.5$, $\delta = 0.75$ and $\delta > 0.5$. 95% confidence interval in brackets.

Table 3 shows the estimates for the extended model. We see again that participants somewhat favor AD over GT in terms of the initial propensities. The model has much less noise than the basic model. We can ask again what our estimates imply in terms of the probability to choose the "correct" strategy - i.e. the strategy with the higher weighted average of $\Delta\Pi^m$ and $\Delta\Pi^m_{\mathbb{E}[T]}$ where the weight is determined by the estimated α. The estimated values of λ imply a probability of choosing this strategy between $\approx$ 92% when $\delta = 0.5$ and $\approx$ 76% when $\delta > 0.5$.

Our main parameter of interest is α. In all cases α substantially exceeds zero, showing that participants do learn from match length realizations. This explains why match length realization affects cooperation in subsequent matches. However, it is not the case that only realized match length matters. The estimates clearly suggest that both expected match length and realized match length play a role for participants' learning. Learning from match length realizations is somewhat more important when the horizon is longer ($\delta > 0.5$), which is intuitive as in these cases we can expect more variation in match lengths.[18]

How well do these models predict actual cooperation? Appendix Figure E.3 shows observed cooperation in stage 1 of a match depending on the predicted probability ($p^{i,m}$) of using

[17]When restricting to $\delta = 0.5$ we obtain $\lambda = 0.010$, for $\delta = 0.75$ we get $\lambda = 0.028$ and for $\delta > 0.5$ we obtain $\lambda = 0.028$ in this case.

[18]We can again control for sizebad. Doing so yields values of $\alpha = 0.437$ and $\lambda = 0.151$ when all data are used. For $\delta = 0.5$ we get $(\alpha, \lambda) = (0.486, 0.434)$, for $\delta = 0.75$ we get $(\alpha, \lambda) = (0.415, 0.151)$ and for $\delta > 0.5$ we get $(\alpha, \lambda) = (0.631, 0.106)$.

Grim Trigger according to the basic model (left panel) and the extended model (right panel).[19] The figure shows that there is some prediction error in both models with observed cooperation differing from predicted cooperation by up to twenty percentage points. Prediction errors are generally lower with the extended model.

In sum this section has shown that simple learning models can explain the data reasonably well. Estimates of our extended model clearly show that participants do learn from match length realization. Because participants learn from match length realization, treatment comparisons can be affected by "unusually" long or short match length realizations. And, because learning displays the "power law of practice" early match length realizations will be particularly important.

3 Case Studies

We will discuss three applications to illustrate how match length realizations can affect treatment comparisons when indefinitely repeated games are compared with finitely repeated games (subsections 3.1 and 3.2) or when indefinitely repeated games are compared with other indefinitely repeated games (subsection 3.3). The three cases highlighted are not part of the Dal Bó and Fréchette (2018) meta-study data used in Section 2.2 and feature a continuous time prisoner's dilemma (subsection 3.1), a public good game (3.2) and oligopoly games (3.3).[20]

3.1 Cooperation in Continuous Time

Our first case study is the paper "Time Horizon and Cooperation in Continuous Time" by Bigoni et al. (2015) published in *Econometrica*. Bigoni et al. (2015) compare cooperation rates in a prisoner's dilemma played in deterministic and stochastic continuous time.[21] They consider games of short (20 seconds) and long (60 seconds) expected length, where here we focus on the short games (which is where they find a treatment effect). The deterministic short game lasts 20 seconds. The stochastic short game has a continuation probability of $\delta = \frac{992}{1000}$ and every 0.16 seconds it ends with probability $1 - \delta$. This means that the expected match length in the continuous game is 20 seconds just as in the deterministic game. The expected median length is 13.86 seconds. Bigoni et al. (2015) focus on average cooperation rates in a match. They find that in short games cooperation is higher under deterministic than under the stochastic horizon.

We now study how this result might be affected by match length realizations. Bigoni et al. (2015) conduct two sessions for each treatment condition. In each session there are 24 participants who play 23 matches. Match length is drawn at the session level, i.e. all participants face the same sequence of match lengths. Figure 3 shows kernel density estimates of the difference between theoretical median match length and realized match length for the two sessions. The

[19]In the setting with only two strategies GT and AD cooperation in stage 1 uniquely identifies strategy GT. Fudenberg and Karreskog (2020) show that initial play is indeed highly predictive about average cooperation in a match using the same data.

[20]Our selection of case studies followed four criteria: (i) the paper should *not* be already included in the meta-study used in Section 2; (ii) it should be on an indefinitely repeated social dilemma; (iii) it has to feature different match length realizations across sessions and (iv) data are publicly available or were made available to us.

[21]This important research programm combines elements from Dal Bó (2005) studying the role of deterministic vs. stochastic horizon in discrete time and Friedman and Oprea's (2012) study of discrete vs. continuous time under a deterministic horizon.

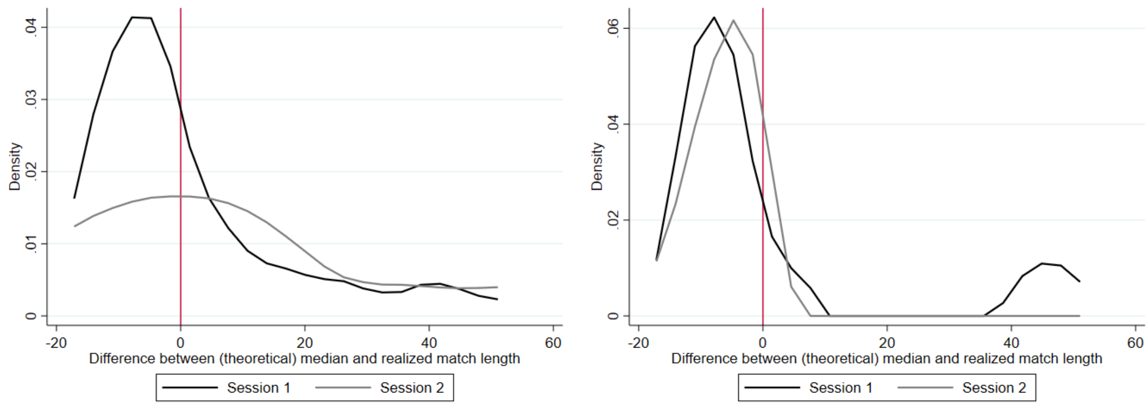

Figure 3: Kernel density estimates of the difference between theoretical median match length and realized match length overall (left panel) and for the 1st third of matches. 58% of all matches and 81% of matches in the 1st third were shorter than theoretical median match length.

left panel shows the entire session and the right panel only the 1st third of the experiment, specifically the first 8 matches (out of 23). It can be seen that in both sessions the vast majority of matches (81%) at the beginning of the experiment (right panel) were shorter than theoretical median length.

	S	N	*Avg ML*	*Avg ML 1st third*	*Median ML*	*Median ML 1st third*	*Avg Coop*	*Avg Coop Initial*
Deterministic	4	2208	20	20	20	20	54.04	73.95
Replication	2	1104	22.94	17.6	11.04	8.48	39.58***	51.90***
Inverse	2	1104	17.86	24.0	17.56	21.66	50.90	73.36
Match Stoch	4	2208	19.97	19.57	13.44	13.52	47.40	61.18**

Table 4: Summary Statistics of the different treatments conducted to replicate Bigoni et al. (2015) . Number of Sessions (S) and observations (N) in the different conditions. Average Match Length (Avg ML), average match length in the 1st third of the experiment (Avg ML 1st third), median match length and median match length in the 1st third, average cooperation rate (avg coop) and average initial cooperation rate (avg coop initial). Stars indicate statistical significance (*** 1%, ** 5%, * 10 %) of the difference to the deterministic case in random effects OLS regression with standard errors clustered at session level (see Appendix Tables D.9 and D.10).

To study whether this realization of match lengths could have affected the treatment effect we first replicated Bigoni et al. (2015)'s experiment. We conducted four sessions of the deterministic condition and then two sessions with the same match-length realizations as Bigoni et al. (2015) ("Replication"). Those sessions were conducted as exact replications of their study. See Appendix C for further details. We further conducted two sessions with inverse match length realizations ("Inverse"). For the inverse sessions we determined a sequence of match lengths $(T^m)_{m=1}^{23}$ as follows. For each realized match length T^m in the Replication we compute $\Pr(x \leq T^m)$ and then replace the m-th entry in the sequence by the value T' that satisfies $\Pr(x \leq T') = 1 - \Pr(x \leq T^m)$. Appendix Figure E.4 illustrates how the "inverse" match length sequences are constructed. Last, we conducted 4 sessions where we randomize the sequence of match lengths at the match level ("Match Stoch"). Hence in this treatment we have 96 different realized match length sequences as opposed to just two.

Table 4 gives an overview of the different treatments we conducted as well as the average and median match lengths. The table shows that - compared to the deterministic case - both average and median match length are short in the replication treatments, particularly in the 1st third of the experiment. There the median match length is only 8.48 seconds, much shorter than the 20 seconds in the deterministic case or than the theoretical median of 13.86 seconds. In the inverse condition these match length realizations are naturally longer with the median match length in the 1st third being 21.66, just above the deterministic condition. Last, as expected, when match lengths are drawn at the match level, then, by the law of large numbers, both average and median lengths are close to the theoretical averages and medians.

How does match length realization affect average cooperation rates and the treatment comparison? First, it should be noted that we manage to replicate Bigoni et al. (2015)'s result quite closely. Between the deterministic and replication treatment there is a 14.46 percentage point difference in average cooperation (Table 4) compared to Bigoni et al. (2015)'s 10.9 percentage point difference (Table II in Bigoni et al. (2015)). We fail to replicate the result, though, when we use inverse match lengths. Here the difference in average cooperation rates to the deterministic case is only 3.14 percentage points and not statistically different from zero. With match level draws (Match Stoch) we find a difference to the deterministic case of 6.64 percentage points which is less than half of the effect size than in the replication, but more than twice the effect size of the inverse condition. The difference between the match stochastic condition and the deterministic case is not statistically significant at the 10% level.[22] Having a closer look at the data, we do find, however, that average *initial* cooperation rates (in the first stage of each match) do differ significantly between the Match Stoch and deterministic environments with an effect size of about half of that found in the replication.

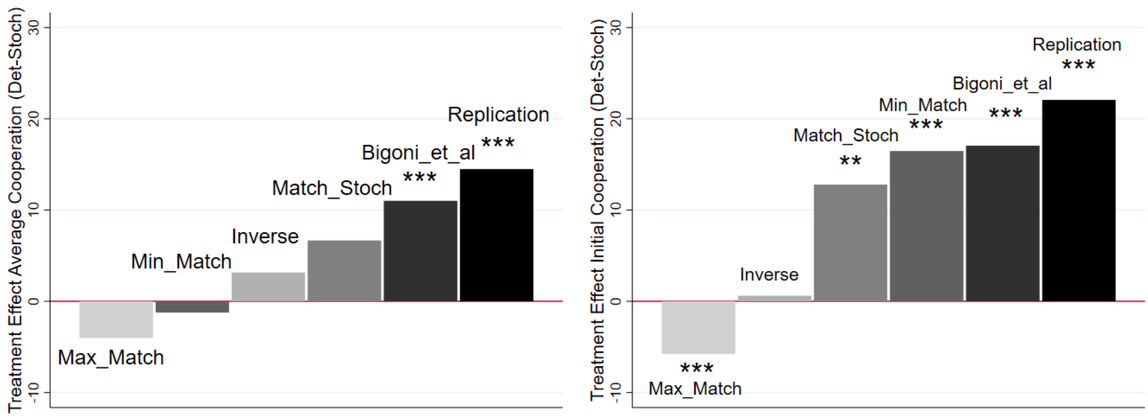

Figure 4: Different effect sizes obtained for the treatment differences between deterministic and stochastic treatments in average cooperation rates (left panel) and initial cooperation rates (right panel) depending on match length realizations. Bigoni et al: original effect size in Bigoni et al. (2015); Replication: replication treatment; Inverse: inverse treatment; MatchStoch: treatment with match length realization drawn at match level; MinMatch: only individuals from MatchStoch who had the smallest number of matches 1-7 above median (specifically 1) in this treatment; MaxMatch: only individuals from MatchStoch who had the largest number of matches 1-7 above median (specifically 6) in this treatment;

[22]In this treatment we use twice the number of observations as in Bigoni et al. (2015). It is still possible that this effect becomes statistically significant with a larger sample size.

Figure 4 illustrates the different effect sizes that can be obtained for the comparison between the deterministic and stochastic game depending on match length realizations. The largest effect size is obtained in our replication of Bigoni et al. (2015)'s original study where we use the same match length realizations as them. This is true for both average cooperation rates (left panel) and initial cooperation rates (right panel). It can also be seen, though, that the treatment difference for average cooperation rates is not statistically significant for any other match length realization. For initial cooperation rates the treatment difference is smaller in the treatment with match-level randomization and statistically not different from zero for the inverse treatment. We also analyzed two sub-groups from the match-stochastic treatments: those with the smallest and those with the largest share of early matches with above median length. For the latter (MaxMatch) we even find a statistically significant negative treatment effect, specifically higher initial cooperation rates in the stochastic game. This exercise illustrates how treatment comparisons can yield entirely different conclusions depending on match length realization. We should also note that - despite the fact that treatment effects can be strongly impacted by match length realizations - we do not consider this an overall unsuccessful replication. Our exact replication was very successful and in the treatment with "many match length realizations" (MatchStoch) the direction of the effect goes in the same direction as in the original study, even though not always being statistically significant.[23]

We will get back to the question of how to measure the "correct" treatment effect in Section 5. Before we do so, we study two more applications showing how treatment comparisons can be affected by match length realizations.

3.2 Finite and Indefinitely Repeated Linear Public Good Games

Our second case study is the paper "An experimental study of finitely and infinitely repeated linear public goods games" by Lugovskyy et al. (2017) published in *Games and Economic Behavior*. The finitely repeated games they study all have a match length of 5 rounds, while for the indefinitely repeated public good games they draw three sequences of match lengths (using discount factor $\delta = 0.8$). Average match length in the 1st third of sequence 1 is below the mean of 5 used in the finite sessions, specifically 4.4 rounds. By contrast, in sequences 2 and 3 it is above, specifically 6 and 6.6, respectively. Hence initial matches are substantially shorter in sequence 1 compared to the other sequences. Overall, however, the three sequences are very similar with average match length across all 15 matches equalling 5.3, 5.4 and 5.7, respectively. In both the finite and indefinitely repeated sessions participants play 15 matches.

The first hypothesis Lugovskyy et al. (2017) test is that "contributions in repeated games with sequences that have probabilistic end rounds will be greater than or equal to those in repeated games with sequences that have known end rounds". They evaluate this hypotheses by comparing behavior in finite and probabilistic settings for four different pairs of treatments which differ in group size, MPCR and whether participants make a binary contribution choice or not.

Table 5 shows the results of this analysis. The first two columns ("Finite" and "Prob All")

[23] Appendix Figure E.5 shows different treatment effects when median cooperation frequencies are compared. Here again, the replication shows the biggest effect and the treatment difference is not statistically significant for the Inverse and MaxMatch condition.

Decision Setting	*All Rounds Cooperation Rate* Finite	Prob All	Prob S1	Prob S23	S1 vs S23
$N = 4$, MPCR$= 0.3$	15.0	22.4	10.36	28.41	
		$<^{**}$	$>$	$<^{***}$	$p = 0.000$
$N = 4$, MPCR$= 0.6$	39.4	44.3	33.52	48.39	
		$<$	$>$	$<$	$p = 0.088$
$N = 2$, MPCR$= 0.6$	41.1	38.3	31.71	42.64	
		$>$	$>$	$<$	$p = 0.084$
$N = 2$, MPCR$= 0.6$, Binary	54.5	41.2	36.84	42.65	
		$>^{***}$	$>^{**}$	$>^{**}$	$p = 0.398$

Table 5: Average cooperation rates across all rounds in the finite sessions (column (1)) and across all sessions with a probabilistic ending (column (2)) as in Table 3 in Lugovskyy et al. (2017). We further split the sessions with probabilistic ending in those with initially short matches (S1) and those with initially long matches (S23). Below each cooperation rate we show how the finite setting compares to the rate in question (as in Lugovskyy et al. (2017)). The last column shows the p-value when comparing initially short and long sequences. Following Lugovskyy et al. (2017) standard errors are clustered at the participant level in all regressions.

reproduce the analysis in Table 3 in Lugovskyy et al. (2017). The analysis shows that in two of the four treatments cooperation is higher in the finitely repeated game and in the other two it is higher in the indefinitely repeated (probabilistic) game. One each of these comparisons is statistically significant. These and other analysis lead Lugovskyy et al. (2017) to conclude "We do not, however, find consistent evidence that overall cooperation rates are affected by whether the number of decision rounds is finite or determined probabilistically."

When we split out the sessions in those with initially short and those with initially long matches, though, we might have reached a different conclusion. The column "Prob S1" shows cooperation rates as well as comparisons in the session with initially shorter matches. In this case all four comparisons point into the same direction: more cooperation in the finitely repeated game. Only one of the comparisons is statistically significant. It should be noted, however, that the first ($N = 4$, MPCR$= 0.3$) and third comparisons ($N = 2$, MPCR$= 0.6$) are both just outside 10 percent statistical significance ($p = 0.150$, $p = 0.102$) in a comparison that is somewhat underpowered.

In the sessions with initially long matches the picture is very different. In this case three out of four comparisons point towards less cooperation in the finitely repeated game. Out of the statistically significant comparisons one each is pointing towards more and one towards less cooperation in the finitely repeated game. Hence while the sessions with initially long matches show more of a similar picture than the overall sample, the sessions with initially short matches behave quite differently and would lead to a different conclusion. It should also be noted that, except for the last comparison ($N = 2$, MPCR$= 0.6$, Binary), the differences in average cooperation rates across the initially short and long sessions are always statistically significant.[24] Appendix Table D.11 shows that similar conclusions hold when we consider first round cooperation rates only.

[24]This difference is not driven by the shorter matches themselves. If we restrict attention to cooperation rates in the last third of matches only we find a difference (S1-S23) of -18.03, $p = 0.000$ for treatment $N = 4$, MPCR$= 0.3$, of -17.55, $p = 0.082$ for treatment $N = 4$, MPCR$= 0.6$, of -7.66, $p = 0.260$ for treatment $N = 2$, MPCR$= 0.6$ and of $-16, 42$, $p = 0.133$ in treatment $N = 2$, MPCR$= 0.6$, Binary.

While on balance it seems to us that Lugovskyy et al. (2017)'s overall conclusion is likely to be robust once "many" match length realizations are considered, the case study shows again how easily different conclusions could have been reached with different match length realizations.

3.3 Strategy Revision Opportunities

Our third case study is the paper "Strategy Revision Opportunities and Collusion" by Embrey et al. (2019) published in *Experimental Economics*. Embrey et al. (2019) explore how the possibility of being able to change a repeated game strategy during the course of play (i.e. to use "behaviour" strategies) affects cooperative behaviour in stylized oligopoly experiments. Their main treatment variations compare games of strategic substitutes and strategic complements with and without revision opportunities (RO). They find that without RO (when strategies have to be encoded upfront) there is more cooperation in games of substitutes than in games of complements. With RO there is more cooperation in games of strategic complements than with substitutes, but the latter difference is not statistically significant. The column "All groups" in Table 6 shows their main treatment effects in terms of efficiency, i.e. in terms of the percentage of the difference between joint profit maximizing payoff and Nash equilibrium payoff realized in the stage game.

with RO	All groups	Short First Matches	Other Groups	Short vs Others
Substitutes	21.0	7.7	28.2	
Complements	26.4	44.2	16.9	
Treatment Effect	- 5.4	-36.4***	11.3**	$p = 0.0030$
without RO	All groups	Short First Matches	Other Groups	Short vs Others
Substitutes	22.7	30.9	18.2	
Complements	9.4	13.8	7.0	
Treatment Effect	13.3**	17.1	11.2	$p = 0.6888$

Table 6: Efficiency measure from Embrey et al. (2019). Treatment Effect is the difference between substitutes and complements. The two rightmost columns split out the groups with short initial matches from the rest. As in Embrey et al. (2019)'s main analysis matches 7-10 are considered. Stars are from t-tests with standard errors clustered at group level.$^{***}p < 0.01$, $^{**}, p < 0.05$, $^{*}p < 0.1$.

Embrey et al. (2019) use a discount factor of 0.875 implying a median match length of 7 stages. There are six different matching groups with different match length realizations. We will group those into two categories: (i) those where first matches are short, specifically where the first three (out of ten) matches all have a below median match length and (ii) all other matches. Table 1 in Embrey et al. (2019) shows that two matching groups (groups 2 and 6) fall into category (i). It should be noted that these matching groups do not have fewer stages overall than others. In fact group 6 has the most stages overall of all groups.[25]

Table 6 shows the average efficiency, defined as the percentage of the difference between

[25] Alternatively we could split the sample into two equal sized categories by focusing on the length of the first two matches only, as there are three groups were both the first and second match are below median length. In this case qualitatively the same conclusions do hold.

joint profit maximizing payoff and Nash equilibrium payoff realized in the stage game. In all cases we follow Embrey et al. (2019) and focus on the average efficiency across matches 7-10.[26] We then compare average efficiency in games of strategic substitutes and complements separately for games with and without revision opportunities (RO).

The table shows that treatment effects depend on match length realizations. The overall negative, but statistically insignificant treatment effect (difference between substitutes and complements) with revision opportunities seems driven by the groups with short initial matches, where the effect is almost seven times larger and statistically highly significant. In the other groups the treatment effect reverses sign and is also statistically significant. The difference between groups with short initial matches and other groups is statistically significant at the 1% level. In this case, hence, diametrically opposite conclusions could be reached when matches with short or long initial realizations are studied. This is also illustrated in the left panel Appendix Figure E.6 which shows the treatment effect for all possible selections of two match length realizations.

Without revision opportunities we see a slightly different pattern. There is an overall positive and statistically significant treatment effect. The effect is somewhat larger with short first matches and somewhat smaller in the other groups, but both are positive. The effect is not significant in either of the subgroups, presumably due to lower sample size. Hence in this case the treatment comparison seems robust to match length realizations as is also illustrated in the right panel of Appendix Figure E.6.

Note that when revision opportunities are ruled out then all updating has to take place across (as opposed to within) matches. The analysis in this Section and the fact that match length realization seems to have a stronger impact with revision opportunities could suggest that within match learning might also play an important role for the effect of match length realizations. This could be an additional reason why the effect sizes observed with human players in the Dal Bó and Fréchette (2018) meta study are larger than with computer simulated learners who learn only across matches (see Section 2.3). In the next section we will discuss implications for experimental design in more detail.

4 Discussion

In this section we discuss some implications of our findings as well as potential ways to diagnose and solve potential problems stemming from match length realizations in experimental studies of indefinitely repeated games. Before we go deeper into the different possibilities, it is useful to define the problem. Assume a researcher is interested in measuring how the probability of cooperation depends on the decision environment $\omega \in \Omega$. Here ω can capture things like subject characteristics (age, gender,...), game parameters (sizeBAD, δ,...), decision settings (lab, field,...) or experimental treatments that affect the probability of cooperation $\Pr(C) = f(\omega)$.

Say the experimenter can conduct two treatments ω_0 and ω_1 which cleanly manipulate a dimension of interest in the decision environment and is interested in the effect of these treat-

[26] Appendix Table D.12 shows all matches

ments on the probability of cooperation. The experimenter is interested in the treatment effect

$$f(\omega_0) - f(\omega_1).$$

The problem arises because the effect of ω can only be measured together with a match length realization. The probability of cooperation is hence given by $\Pr(C) = f(\omega, T)$, where T is the match length realization drawn from a geometric distribution with mean $\frac{1}{1-\delta}$ and standard deviation $\sqrt{\frac{\delta}{(1-\delta)^2}}$.

Say the experimenter conducts K sessions for each treatment, indexed $k = 1, ..., K$, where each session has N participants, indexed $i = 1, ..., N$ who play each M matches, indexed $m = 1, ..., M$. Hence, for each treatment ω_0 and ω_1 the experimenter observes $K \cdot N \cdot M$ individual decisions to cooperate $f(\omega_0, T^m_{0,k})$ and $f(\omega_1, T^m_{1,k})$, which in the usual design depend on K realizations of sequences of match lengths $(T^m_{0,k})^M_{m=1}$ and $(T^m_{1,k})^M_{m=1}$. It is common to focus on average cooperation across the $K \cdot N \cdot M$ observations.[27] This gives the following observed (sample average) treatment effect

$$\frac{1}{K \cdot N \cdot M} \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{m=1}^{M} \left(f(\omega_0, T^m_{0,k}) - f(\omega_1, T^m_{1,k}) \right), \tag{8}$$

where we have abstracted from individual heterogeneity in the treatment effect as our focus is on match length realization. In general the observed treatment effect will be biased, i.e. will not equal $f(\omega_0) - f(\omega_1)$. A common practice in experimental research is to draw the same sequence of match length realizations for all treatments, i.e. to ensure that $(T^m_{0,k})^M_{m=1} = (T^m_{1,k})^M_{m=1}, \forall k$. Both Lugovskyy et al. (2017) and Embrey et al. (2019) do exactly that, i.e. draw one set of sequences and use it for all treatments. If the influence of match length realization is orthogonal to the treatment, then the resulting treatment effect will *not* be biased. To see this note that in this case the observed treatment effect (8) can be written as follows

$$\begin{aligned} & \frac{1}{K \cdot N \cdot M} \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{m=1}^{M} \left(f(\omega_0) + f(T^m_{0,k}) - f(\omega_1) - f(T^m_{1,k}) \right) \\ = & \frac{1}{K \cdot N \cdot M} \sum_{k=1}^{K} \sum_{i=1}^{N} \sum_{m=1}^{M} \left(f(\omega_0) - f(\omega_1) \right) \\ = & f(\omega_0) - f(\omega_1). \end{aligned}$$

However, in practice, this is often not the case and there will be interactions between the treatment and the effect of match length realization. Both case studies (Lugovskyy et al. (2017) and Embrey et al. (2019)) have illustrated that point. What can researchers do if this orthogonality condition is not given and if - as a consequence - fixing the sequence of match length realizations is not enough?

[27] Sometimes the focus is only on a subset of matches. This could easily be incorporated in the arguments below. It is more common to focus on initial cooperation in each match, but often people also consider average cooperation across the T^m stages of match m. The arguments below apply irrespective of which of these cases is considered.

Diagnosing the Problem A first step towards dealing with potential bias in treatment effects induced by match length realizations is to diagnose it. As we have seen in Section 2 there are some regularities in the effect of match length realizations on cooperation that allow identifying and bounding the effect. The most important regularity is that cooperation rates seem to be monotonically increasing in match length, i.e. $f(\omega, T-x) \leq f(\omega, T) \leq f(\omega, T+x), \forall x \in \mathbb{N}$. If monotonicity is indeed given, then it is possible for the researcher to get an idea of how strong the relationship between match length realization and cooperation is in the environment(s) they consider. This can be done by using "very different" match length realizations (across sessions or matching silos) and compare the effect sizes observed for these. Note that in order to do this, it is best if the same sequence of match length realizations is drawn across treatments. More precisely, under monotonicity and equal match length realizations across treatments it is possible to bound the treatment effect by comparing the average treatment effect under the (initially) shortest realizations (ATE^{-}) and the treatment effect under the (initially) longest realizations (ATE^{+}). As we know the distribution from which match lengths are drawn, we can further compute the probability that match length realizations are even shorter (longer) than the shortest (longest) observed match lengths. Denote this probability by q^{-} (q^{+}) and define $q = q^{-} + q^{+}$. The researcher can then make statements of the form "With probability $(1-q)$ the average treatment effect is between ATE^{-} and ATE^{+}". See Imbens and Angrist (2004) or De Quidt, Haushofer, and Roth (2018) for a more detailed discussion.[28]

Fixing the Problem We now outline some potential ways to "fix" the problem. It is important to note that while none of the proposed solutions is a perfect fix for any possible environment, many of them work in specific environments. Maybe even more importantly many of the solutions discussed can help with diagnosing the problem even if the treatment effect cannot be fully de-biased.

The most immediate solution to the problem is probably to simply use as many match length realizations as possible. There are three basic ways to do so: (i) increase K, (ii) increase M or (iii) change the level at which randomization takes place. (i) is a great solution whenever it is feasible. If it is possible to obtain information on the expected size of the effect of match length realization (e.g. from prior literature) then power analysis can be conducted to determine how many sessions K are needed to obtain unbiased treatment effects. (ii) seems less promising. Increasing M works only to the extent that the effect of early match length realization washes out over time. As our analysis in Section 2.3 has shown this does not seem to be the case for match numbers that can reasonably fit in a two-hour experimental session. A possibility that we explored in this paper is (iii) when we randomized match length at the match level in our "MatchStoch treatment". A downside of this approach is that it can induce waiting times as all participants in a matching silo have to wait for the longest match to end before being rematched. This concern is especially important if δ is high and restricts the total number of matches that can be played.

There are some alternatives to the standard random termination method. In a method used

[28]This can be done even if match length realizations differ across treatments as long as the intervals defined by the two sets of match length realizations do have some overlap, but the computation is not as straightforward in this case.

e.g. by Sabater-Grande and Georgantzis (2002), Cabral, Ozbay, and Schotter (2014) or Vespa (2019), a fixed (known) number of rounds are played with certainty, and payoffs in these rounds are discounted at a known rate δ. Afterwards, there is a fixed known probability δ that the match continues for an additional round, and payoffs in these rounds are no longer discounted. Andersson and Wengstroem (2012) and Cooper and Kuehn (2014) use a similar method that also starts with a fixed number of rounds with payoff discounting, but is then followed by the coordination game induced by considering only two strategies, "Grim Trigger" and "Always Defect". The first method avoids very short matches, but, as overall match length is still random, does not eliminate the problem. The second method does not have an uncertain match length, but it has the downside that the number of repeated game strategies allowed needs to be restricted ex ante. Further, Fréchette and Yuksel (2017) found that neither of these methods induces behaviour that is consistent with the presence of dynamic incentives.

Fréchette and Yuksel (2017) propose a promising method called block random termination. Participants play as in the standard method, but in blocks of a pre-announced fixed number of rounds. Within a block they receive no feedback about whether or not the match has continued until that round, but they make choices that will be payoff-relevant if it has. Once the end of a block is reached, subjects are told whether the match ended within that block and, if so, in what round; otherwise, they are told that the match has not ended yet, and they start a new block. Subjects are paid for rounds only up to the end of a match, and all decisions for subsequent rounds within that block are void. With block random termination the length of blocks is a crucial parameter which has to be set carefully.[29] An open question is whether participants learn mainly from the number of stages played or from the number of stages that are payoff relevant.

Other possibilities could include re-sampling approaches by selecting subsets of sessions or matching silos with different match length realizations. Another possibility is to use "inverse" designs (see Section 3.1) more systematically to pair each session with its inverse. Or one could start the experiment with a "training phase", where participants are given the chance to learn about the distribution of match lengths by observing several realizations.[30] Last, one could use constraints on the realized empirical distribution by e.g. imposing that the mean match length of early matches cannot be more than one standard deviation away from the expected length.[31] In sum, there are many possibilities to deal with the problem of match length realizations. Which one is the most suitable will depend on the specific environment researchers are interested in studying.

[29] If blocks are of length one the method is the same as the standard random termination method. With very long blocks a downside is that the experiment lasts long and fewer matches can be played.

[30] One advantage of such a training phase is that - since there is no strategic interaction - match lengths can be randomized at the individual level for the training phase. This means that many match length realizations can be observed a least for this phase.

[31] Which constraints are effective will depend on the specific treatment comparisons the researcher is interested in. In Section 3.1 we have, for example, seen that cooperation rates are similar for all sequences where mean match length realizations in early matches are within one standard deviation of what we should expect theoretically. However, results in Section 3.2 differ even across sessions where this is the case.

5 Conclusions

We have seen that the realized length of early matches in indefinitely repeated games has a substantial impact on cooperation rates in subsequent matches. Using three cases from the literature we also demonstrated a strong impact on treatment comparisons, both in terms of the size and the direction of the treatment effect. Our results have important implications for our understanding of how people learn in infinitely repeated games, for the interpretation of treatment effects when there are stochastic elements, as well as for experimental design.

Theories of learning should take into account how agents learn from sequences of realizations of random variables. One interesting question in this context is which moment of the distribution of match length realizations is most important for learning. In empirical research indefinitely repeated games are implemented using a mean expected match length that derives from the discount factor in the infinitely repeated game considered. Appendix Table D.13 shows, however, that the median match length realization seems a more important determinant for participants' behaviour than the mean.[32] This raises the question of which sequence of indefinitely repeated games "correctly" represents the infinitely repeated game one ultimately has in mind. This question has been answered theoretically under standard game theoretic assumptions. But, given how people seem to learn in these games, it might be necessary to rethink this question. For example, in our first case study (subsection 3.1), it is the match stochastic treatment, which, as expected, closely matches the mean match length of the deterministic case. The median match length of the deterministic case is better matched by the inverse treatment, though. Which of those is the more relevant comparison depends on which of these moments is more important for how people learn. If it indeed turns out that median match length is the key statistic determining learning, then future research in both theory and experiments is needed to build and test new models of learning which can accommodate this fact.

For applied work it is important to know that different learning experiences may lead to different behaviours and may be confounded with treatment effects. Not accounting for potential differences in learning experience may lead to falsely claiming effects when there are none or to not finding effects when there are. In Section 4 we have discussed at length which design features empirical studies might use to ensure their conclusions are less vulnerable to match length realization effects. Future research could select the most promising among these and systematically assess how well they work in practice.

References

Aldrovandi, S., Brown, G., & Wood, A. (2015). Social norms and rank-based nudging: Changing willingness to pay for healthy food. *Journal of Experimental Psychology: Applied*, *21*(3), 242-254.

Andersson, O., & Wengstroem, E. (2012). Credible communication and cooperation: Experimental evidence from multi-stage games. *Journal of Economic Behavior and Organization*, *81*(1), 207-219.

[32]There is research in other contexts suggesting that the median experience might be relevant and that people understand information based on median/rank better especially when there is a lot of skewness in the distribution, which is also the case with match length realizations (Aldrovandi, Brown, & Wood, 2015; Wood, Brown, & Maltby, 2012). One example is consumption of alcohol, where people seem to have a good sense of how their consumption compares to the median, but not to the mean (Wood et al., 2012).

Bernard, M., Fanning, J., & Yuksel, S. (2018). Finding cooperators: Sorting through repeated interaction. *Journal of Economic Behavior and Organization*, *147*, 76–94.

Bigoni, M., Casari, M., Skrzypacz, A., & Spagnolo, G. (2015). Time horizon and cooperation in continuous time. *Econometrica*, *83*, 587-616.

Blackburn, J. (1936). Acquisition of skill: An analysis of learning curves. *IHRB Report*, *73*.

Blonski, M., Ockenfels, P., & Spagnolo, G. (2011). Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, *3(3)*, 164-192.

Blonski, M., & Spagnolo, G. (2015). Prisoners' other dilemma. *International Journal of Game Theory*, *44*, 61-81.

Boergers, T., & Sarin, R. (1997). Learning through reinforcement and replicator dynamics. *Journal of Economic Theory*, *77*, 1-14.

Cabral, L., Ozbay, E., & Schotter, A. (2014). Intrinsic and instrumental reciprocity:an experimental study. *Games and Economic Behavior*, *87*, 100-121.

Cabrales, A., & Garcia-Fontes, W. (2000). Estimating learning models from experimental data. *UPF Economics and Business Working Paper 501*.

Camera, G., & Casari, M. (2009). Cooperation among strangers under the shadow of the future. *American Economic Review*, *99*(3), 979–1005.

Camerer, C., & Ho, T. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, *67*(4), 827-874.

Cheung, Y.-W., & Friedman, D. (1997). Individual learning in normal form games: Some laboratory results. *Games and Economic Behavior*, *19*(1), 46–76.

Cooper, D., & Kuehn, K. (2014). Communication, renegotiation, and the scope for collusion. *American Economic Journal: Microeconomics*, *6*(2), 247-278.

Dal Bó, P. (2005). Cooperation under the shadow of the future: experimental evidence from infinitely repeated games. *American Economic Review*, *95*(5), 1591–1604.

Dal Bó, P., & Fréchette, G. (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, *101(1)*, 411-429.

Dal Bó, P., & Fréchette, G. (2018). On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, *56*(1), 60-114.

De Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and bounding experimenter demand. *American Economic Review*, *108*(11), 3266-3302.

Embrey, M., Fréchette, G., & Yuksel, S. (2018). Cooperation in the finitely repeated prisoner's dilemma. *Quarterly Journal of Economics*, 301-355.

Embrey, M., Fréchette, G. R., & Yuksel, S. (2017). Cooperation in the finitely repeated prisoner's dilemma. *The Quarterly Journal of Economics*, *133*(1), 509–551.

Embrey, M., Mengel, F., & Peeters, R. (2019). Strategy revision opportunities and collusion. *Experimental Economics*, *22 (4)*, 834-856.

Engle-Warnick, J., & Slonim, R. L. (2006). Learning to trust in indefinitely repeated games. *Games and Economic Behavior*, *54*(1), 95–114.

Erev, I., & Roth, A. (1998). Predicting how people play in games: Reinforcement learning in experimental games with unique mixed strategy equiloibria. *The American Economic Review*, *88*(4), 848-881.

Erev, I., & Roth, A. (2001). Simple reinforcement learning models and reciprocation in the prisoner's dilemma game. In G. Gigerenzer & R. Selten (Eds.), *Bounded rationality: The adaptive toolbox* (p. 215-231). MIT Press.

Fréchette, G., & Yuksel, S. (2017). Infinitely repeated games in the laboratory: Four perspectives on discounting and random termination. *Experimental Economics*, *20*(2), 279-308.

Friedman, D., & Oprea, R. (2012). A continuous dilemma. *American Economic Review*, *102*(1), 337–63.

Fudenberg, D., & Karreskog, G. (2020). Learning about initial play determines average cooperation in repeated games. *mimeo*.

Fudenberg, D., & Kreps, D. M. (1993). Learning mixed equilibria. *Games and Economic Behavior*, *5*(3), 320–367.

Fudenberg, D., & Levine, D. K. (1998). *The theory of learning in games* (Vol. 2). MIT press.

Hanaki, N., Sethi, R., Erev, I., & Peterhansl, A. (2005). Learning strategies. *Journal of Economic Behavior and Organization*, *56*(4), 523-542.

Hopkins, E. (2002). Two competing models of how people learn in games. *Econometrica*, *70*(6), 2141-2166.

Imbens, G., & Angrist, J. (2004). Confidence intervals for partially identified parameters. *Econometrica*, *72*(6), 1845-1857.
Ioannou, C., & Romero, J. (2014). Learning with repeated-game strategies. *Frontiers in Neuroscience*.
Kagel, J. H., & Levin, D. (1986). The winner's curse and public information in common value auctions. *The American economic review*, 894–920.
Lugovskyy, V., Puzzello, D., Sorensen, A., Walker, J., & Williams, A. (2017). An experimental study of finitely and infinitely repeated linear public goods games. *Games and Economic Behavior*, *102*, 286-302.
Mengel, F. (2018). Risk and temptation: A meta-study on prisoner's dilemma games. *Economic Journal*, *128*, 3182-3209.
Mookherjee, D., & Sopher, B. (1997). Learning and decision costs in experimental constant sum games. *Games and Economic Behavior*, *19*, 97-132.
Plott, C. R., & Sunder, S. (1982). Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of political economy*, *90*(4), 663–698.
Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, *8*(1), 164–212.
Roth, A. E., & Murnighan, J. K. (1978). Equilibrium behavior and repeated play of the prisoner's dilemma. *Journal of Mathematical Pyschology*, *17*(2), 189-198.
Rustichini, A. (1999). Optimal properties of stimulus—response learning models. *Games and Economic Behavior*, *29*(1-2), 244–273.
Sabater-Grande, G., & Georgantzis, N. (2002). Accounting for risk aversion in repeated prisoners' dilemma games: an experimental test. *Journal of Economic Behavior and Organization*, *48*(1), 37-50.
Vespa, E. (2019). An experimental investigation of cooperation in the dynamic common pool game. *International Economic Review*, *61*(1), 417-440.
Vriend, N. J. (1997). Will reasoning improve learning? *Economics Letters*, *55*(1), 9–18.
Wilcox, N. (2006). Theories of learning in games and heterogeneity bias. *Econometrica*, *74*(5), 1271-1292.
Wood, A., Brown, G. D., & Maltby, J. (2012). Social norm influences on evaluations of the risks associated with alcohol consumption: Applying the rank-based decision by sampling model to health judgments. *Alcohol and Alcoholism*, *47*(1), 57-62.

Online Appendix for "Match Length Realizations and Cooperation in Indefinitely Repeated Games"

F. Mengel, L. Orlandi, S. Weidenholzer

Contents

A Simulation Studies of Alternative Learning Models

We simulate three different learning models. Under *reinforcement learning without counterfactuals* (see e.g. Erev and Roth (1998) and Roth and Erev (1995)) agents increase the propensity of the strategy chosen by the payoff received, i.e.

$$\psi_s^{i,m+1} = \psi_s^{i,m} + \mathbb{1}(s^{i,m} = s)\pi(s, s^{-i,m}, T^m),$$

where $\mathbb{1}(s^{i,m} = s)$ indicates whether agent i uses strategy s in match m or not. Note that the propensity for strategies not chosen does not change. This can be seen as a fairly simplistic way of learning which is solely driven by ones' own experience. Match length realizations influence learning through their role in determining payoffs in the game. Note that this form of reinforcement learning features what Erev and Roth (1998) call "force of habit" where frequently chosen actions are reinforced more frequently. This is not the case under *reinforcement learning with counterfactuals* (as discussed in the main part of the paper) where also strategies that have not been played are reinforced and propensities for all strategies evolve according to

$$\psi_s^{i,m+1} = \psi_s^{i,m} + \pi(s, s^{-i,m}, T^m).$$

In order to specify (smooth) *fictitious play* in the present environment we need to specify beliefs about play of the others as well as beliefs about match length realizations. Here we show a version where agent i simply uses the average previous match length realization, given by $\bar{T}^m = \frac{1}{m}\sum_{k=1}^{m} T^k$, and the share of her opponents previously choosing grim trigger, given by $\sigma^{-i,m} = \frac{1}{m}\sum_{k=1}^{m} \mathbb{1}(s^{-i,k} = GT)$, where $\mathbb{1}(s^{-i,k} = GT)$ equals one if agent i's opponent had chosen GT in period k and is zero otherwise. We adopt a propensity based formulation of fictitious play as in Camerer and Ho (1999) and Hopkins (2002) where propensities are simply given by the expected payoffs $\hat{\pi}$ given these beliefs, i.e.

$$\psi_s^{i,m+1} = \hat{\pi}(s, \sigma^{-i,m}, \bar{T}^m).$$

For our simulation exercise we use Luce's linear probability choice rule where choice probabilities are linearly proportional to propensities.

To quantify the effect of match length realizations on cooperation rates in these models, we have run several simulations for each of these three learning models, using key game parameters from the meta dataset of Dal Bó and Fréchette (2018)'s data. The payoff parameters of the underlying PD games were chosen to correspond to the 25th, 50th and 75th percentile of the distribution of `SizeBAD`. We follow Erev and Roth (1998) and have normalized payoffs of the underlying PD games to ensure payoffs and propensities are positive. Table A.1 reports the corresponding parameter values of Dal Bó and Fréchette (2018) alongside the values for `R` and `P` when `S` and `T` were normalized to 0 and 1, respectively. In each simulation run, 16 agents were matched against each other in a round robin tournament and match length realizations were drawn at the session level using the discount factor $\delta = 0.75$, thus targeting the median group size and the most frequent discount factor in Dal Bó and Fréchette (2018). Our exercise contains 4000 simulated experiments played over 15 matches for each of the three payoff configuration leading to 960,000 indefinitely repeated PD games for each of the three learning models.

percentile	SizeBAD	ℓ	g	R	P
25	0.1625	0.5	0.4	0.789	0.263
50	0.2	0.56	1	0.609	0.219
75	0.667	1.85	2	0.588	0.381

Table A.1: Payoff parameters used for simulations.

In the reinforcement learning models, initial propensities for both strategies were set to the expected payoff when the match length corresponds to its mean and the opponent randomizes with equal probability among the two strategies, averaged across the two strategies, i.e. $\hat{\psi}_s^{i,0} = (\hat{\pi}(GT, \frac{1}{2}, T) + \hat{\pi}(AD, \frac{1}{2}, T))/2$. Note that for the reinforcement models this implies initial choice probabilities of $p_{GT}^{i,m} = p_{AD}^{i,m} = \frac{1}{2}$. Correspondingly, in the fictitious play model we started with initial belief on the strategy of the opponent of $\sigma^{-i,0} = 1/2$ and with initial belief on the match length of $\bar{T}^0 = \mathbb{E}[T]$.

Figure A.1: Effect Size for Δ_{above}^{1st} in simulated data for the three models (bars) compared to effect size in Dal Bó and Fréchette (2018) data (dashed line) in two samples: below median values of `SizeBAD` (left three bars and thick dashed line) and above median values of `SizeBAD` (right three bars and thick dashed line).

For each of these three datasets we run regressions identical to the one shown in column (1) of Table 1, thus measuring the impact of long match length realizations in the first third of matches on initial cooperation rates in the remainder of matches (see Tables A.2-A.4). Standard errors are clustered at the run level. We then compare the effect size obtained in these regressions to the empirically observed one. As we have seen that there is an interaction effect between `SizeBAD` and the effect of early match length realization (Table 1) we also split the sample in below and above median `SizeBAD`.

Figure A.1 shows the results. Reinforcement learning (without counterfactuals) shows a slightly larger effect compared to the other models.[1] For all models the effect size does not differ across the two subsamples of low and high `SizeBAD` ($p > 0.1$). This differs from the human data, where for small values of SizeBAD early match length realization has a much larger effect size, though the difference is just outside of conventional levels of statistical significance ($p = 0.1753$). In this sample the effect size obtained purely through the learning models is about

[1]Given the substantial sample size in the simulations all the effects obtained are highly statistically significant ($p < 0.0001$).

	Simulated Reinforcement Learning					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ^{1st}_{above}	0.040***	0.040***	0.042***	0.039***	0.039***	0.039***
	(0.002)	(0.003)	(0.004)	(0.002)	(0.002)	(0.005)
Δ^{2nd}_{above}				0.027***	0.026***	0.035***
				(0.003)	(0.002)	(0.004)
`SizeBAD`		-0.139***	-0.138***		-0.161***	-0.159***
		(0.003)	(0.003)		(0.003)	(0.004)
$\Delta^{1st}_{above}\times$ `SizeBAD`			-0.004			-0.000
			(0.011)			(0.012)
$\Delta^{2nd}_{above}\times$ `SizeBAD`						-0.025**
						(0.010)
Constant	0.415***	0.463***	0.463***	0.419***	0.475***	0.474***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Test $\Delta^{1st}_{above} = \Delta^{2nd}_{above}$	-	-	-	0.0063	0.0003	0.0989
Observations	1,920,000	1,920,000	1,920,000	960,000	960,000	960,000
R-squared	0.002	0.013	0.013	0.002	0.017	0.017

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.2: Simulated Reinforcement Learning Model. Columns (1)-(3): Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by dummy Δ^{1st}_{above} indicating whether more than $\frac{2}{3}$ of matches in the 1st third of the experiment were longer than the theoretical median match length. Columns (4)-(6): Initial (first stage) cooperation rate in the 3rd third of matches explained by dummies Δ^{1st}_{above} and Δ^{2nd}_{above}. Standard errors clustered at session level.

	Simulated Reinforcement Learning with Counterfactuals					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ^{1st}_{above}	0.030***	0.029***	0.037***	0.025***	0.024***	0.029***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)
Δ^{2nd}_{above}				0.016***	0.018***	0.024***
				(0.001)	(0.001)	(0.002)
`SizeBAD`		-0.181***	-0.178***		-0.184***	-0.180***
		(0.001)	(0.001)		(0.002)	(0.001)
$\Delta^{1st}_{above}\times$ `SizeBAD`			-0.023***			-0.017***
			(0.003)			(0.003)
$\Delta^{2nd}_{above}\times$ `SizeBAD`						-0.017***
						(0.005)
Constant	0.399***	0.461***	0.460***	0.400***	0.463***	0.462***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Test $\Delta^{1st}_{above} = \Delta^{2nd}_{above}$	-	-	-	0.0001	0.0000	0.0000
Observations	1,920,000	1,920,000	1,920,000	960,000	960,000	960,000
R-squared	0.010	0.200	0.201	0.010	0.207	0.207

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.3: Simulated Reinforcement Learning Model with Counterfactuals. Columns (1)-(3): Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by dummy Δ^{1st}_{above} indicating whether more than $\frac{2}{3}$ of matches in the 1st third of the experiment were longer than the theoretical median match length. Columns (4)-(6): Initial (first stage) cooperation rate in the 3rd third of matches explained by dummies Δ^{1st}_{above} and Δ^{2nd}_{above}. Standard errors clustered at session level.

a third of the overall effect size. In the sample where defection is relatively attractive (high `SizeBAD`), the simulated effect size is about 80% of the empirical effect size with human players.

We have seen that all three learning models can explain the direction of the empirically observed effect of match length realization and that effect sizes are similar across the three

	Simulated Fictitious Play Learning					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ^{1st}_{above}	0.027***	0.029***	0.038***	0.021***	0.023***	0.030***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Δ^{2nd}_{above}				0.023***	0.022***	0.029***
				(0.001)	(0.001)	(0.002)
`SizeBAD`		-0.178***	-0.175***		-0.189***	-0.185***
		(0.001)	(0.001)		(0.001)	(0.002)
$\Delta^{1st}_{above}\times$ `SizeBAD`			-0.025***			-0.019***
			(0.003)			(0.005)
$\Delta^{2nd}_{above}\times$ `SizeBAD`						-0.022***
						(0.005)
Constant	0.392***	0.449***	0.448***	0.395***	0.456***	0.454***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Test $\Delta^{1st}_{above} = \Delta^{2nd}_{above}$	-	-	-	0.2481	0.4042	0.9592
Observations	1,920,000	1,920,000	1,920,000	960,000	960,000	960,000
R-squared	0.007	0.128	0.128	0.009	0.149	0.149

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table A.4: Simulated Fictitious Play Learning Model. Columns (1)-(3): Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by dummy Δ^{1st}_{above} indicating whether more than $\frac{2}{3}$ of matches in the 1st third of the experiment were longer than the theoretical median match length. Columns (4)-(6): Initial (first stage) cooperation rate in the 3rd third of matches explained by dummies Δ^{1st}_{above} and Δ^{2nd}_{above}. Standard errors clustered at session level.

models. However, we have also seen that effect sizes are larger empirically than what we would expect from the models. This is particularly the case when `sizeBAD` is small, i.e. when cooperation is relatively attractive. One possible explanation why effect sizes are larger with human players is that they may stop learning after some rounds, while the simulated learning models keep learning. Hence for the simulated learners the effect of initially long matches is (partially) corrected when later matches are shorter.

B Population Estimates of Learning Models

This section reports the results of Monte Carlo studies designed to investigate the presence and direction of bias in the estimation of α, λ and $(\psi_{GT}^{i,0} - \psi_{AD}^{i,0})$. As explained in Section 2.3 we estimate the following likelihood function.

$$L(\psi_{GT}^{i,0}, \psi_{AD}^{i,0}, \lambda, \alpha) = \prod_{i=1}^{N} \prod_{m=1}^{M} \left(p^{i,m}\right)^{y_{GT}^{i,m}} \left(1 - p^{i,m}\right)^{1-y_{GT}^{i,m}}$$

with

$$p^{i,m} = \frac{e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \alpha\Delta\Pi^m + (1-\alpha)\Delta\Pi^m_{\mathbb{E}[T]}]}}{1 + e^{\lambda[(\psi_{GT}^{i,0} - \psi_{AD}^{i,0}) + \alpha\Delta\Pi^m + (1-\alpha)\Delta\Pi^m_{\mathbb{E}[T]}]}}$$

The log-likelihood is given by

$$LL(\psi_{GT}^{i,0}, \psi_{AD}^{i,0}, \lambda, \alpha) = \sum_{i=1}^{N} \sum_{m=1}^{M} y_{GT}^{i,m} \ln\left(p^{i,m}\right) + (1 - y_{GT}^{i,m}) \ln\left(1 - p^{i,m}\right).$$

Wilcox (2006) has made the point that in pooled estimations of models that contain lagged variables (such as the learning model we are interested in), the presence of heterogeneity in the parameter λ can cause biases in the estimated parameters (see also Cabrales and Garcia-Fontes (2000)). To study whether our estimations could potentially be affected by a similar issue we assume that α, our parameter of interest, is constant across players and has an unknown true mean α_0. We assume, however, that there is heterogeneity in λ which has known true mean λ_0. For simplicity we assume that $\psi_{GT}^{i,0} - \psi_{AD}^{i,0} = 0$. Without heterogeneity (i.e. if $\lambda^i = \lambda_0, \forall i$), the maximum likelihood estimator $\hat{\alpha}$ of α_0 solves

$$m(\hat{\alpha}) = (N(M-1))^{-1} \sum_{i=1}^{N} \sum_{m=1}^{M} \left(y_{GT}^{i,m} - p^{i,m}\right) [\Delta\Pi^m - \Delta\Pi^m_{\mathbb{E}(T)}]\lambda.$$

With heterogeneity in λ the MLE converges to

$$m(\alpha_0) = \left(\mathbb{E}^\lambda[\mathbb{E}^i(y_{GT}^{i,m})] - \mathbb{E}^\lambda[\mathbb{E}^i(p^{i,m})]\right) \mathbb{E}^\lambda[\mathbb{E}^i(\Delta\Pi^m - \Delta\Pi^m_{\mathbb{E}(T)})\lambda] + W,$$

where $\mathbb{E}^\lambda$ is the expectation with respect to the distribution of λ^i and W consists of two potentially non-zero terms (i) $\mathbb{E}^\lambda[\mathrm{Cov}^i[(y_{GT}^{i,m} - p^{i,m}), \Delta\Pi^m - \Delta\Pi^m_{\mathbb{E}(T)}]]$ and (ii) $\mathrm{Cov}^\lambda[\mathbb{E}^i(y_{GT}^{i,m}) - \mathbb{E}^i(p^{i,m}), \mathbb{E}^i(\Delta\Pi^m - \Delta\Pi^m_{\mathbb{E}(T)})]$, where Cov^i denotes the covariance given λ^i and Cov^λ the covariance with respect to the distribution of λ^i. See Wilcox (2006) for the derivation of this expression. Intuitively the first term measures the expected covariance between the reinforcement values $(\Delta\Pi^m - \Delta\Pi^m_{\mathbb{E}(T)})$ and the model prediction errors $(y_{GT}^{i,m} - p^{i,m})$. The second term measures the between players steady-state covariance of the expected pooled model prediction errors given λ^i and the expected counterfactual reinforcement given λ^i. Wilcox (2006) has shown ("Second Argument" in the Appendix) that the first term approaches zero as the variance of $(\Delta\Pi^m - \Delta\Pi^m_{\mathbb{E}(T)})$ approaches zero.[2] The potential source of the bias is the second term $(\mathrm{Cov}^\lambda[\mathbb{E}^i(y_{GT}^{i,m}) - \mathbb{E}^i(p^{i,m}), \mathbb{E}^i(\Delta\Pi^m - \Delta\Pi^m_{\mathbb{E}(T)})])$.

[2]This is true because our model takes counterfactuals fully into account and hence corresponds to the case $\delta_0 = 1$ in Wilcox (2006).

We conduct Monte Carlo studies to answer the following questions: (i) Do we obtain biases and if so in which of our parameters of interest? and (ii) how big are these biases and can we systematically sign them? In the simulations the data generating process is given by equations (6) and (7). We assume $\delta = 0.75$, $M = 15$ matches and $N = 16$ players who are randomly matched in pairs to play an indefinitely repeated game. We conduct two sets of simulations: (i) without heterogeneity in the underlying parameters and (ii) with heterogeneity. For each case we consider values of $\alpha_0 \in \{0, 0.25, 0.5, 0.75, 1\}$. Without heterogeneity we consider two values of $\lambda_0 \in \{0.2, 1\}$. The value 0.2 is close to the estimate we obtain with the Dal Bó and Fréchette (2018) data and 1 is used as an alternative case of substantially higher λ to get a sense of how significant potential downward biases could be. With heterogeneity we assume that λ is uniformly distributed either in $[0.1, 0.2]$ or $[0.5, 1.5]$. We also consider two cases with heterogeneity in α in which case we assume α is uniformly distributed in $[0, 1]$. The term $(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$ is of secondary interest and we always assume $(\psi^{i,0}_{GT} - \psi^{i,0}_{AD}) = 0$. For each of these parameter combinations we conduct 1000 such exercises leading to a total of 240 000 observations of individual choices between GT and AD for each case.

We first discuss the results when there is no heterogeneity. Without heterogeneity Cabrales and Garcia-Fontes (2000) (and Wilcox (2006)) have shown that the maximum likelihood estimator is consistent in the limit as the number of observations approaches infinity.[3] For a reasonably large sample as in our simulations and as in Section 2.3 we would hence not expect substantial biases in this case. Indeed Table B.1 shows that in this case all three parameters are estimated very precisely and without bias.

As an aside, it should also be noted that for individual level estimates the sample size in the our data is rather small and, as expected, we do see substantial biases emerge. For example, in the case where $\alpha_0 = 0.5$ and $\lambda_0 = 1$ shown in Table B.1, individual level estimates fail to converge in almost half of the cases. When they do the average estimated $\alpha = 0.453$, a substantial downward bias, and the range of estimated α reaches from 0.1 to 1. It is clearly advisable, hence, to estimate α at the population level given the structure of the data available.

Table B.2 shows the results of our Monte Carlo simulations with heterogeneity. The table shows that when there is heterogeneity in λ we obtain systematic downward biases in our estimates of λ. The maximal size of the bias we detect is a 15% underestimation of the true mean in the case where $\lambda \sim [0.5, 1.5]$ and a 5% underestimation in the case where $\lambda \sim [0.1, 0.3]$. When $\lambda \sim [0.5, 1.5]$ we detect biases in α in two out of the five cases. For $\alpha_0 = 0.5$, closest to the estimated values in Section 2.3, we detect no statistically significant bias. If at all, there is slight overestimation by about 5%. Crucially, whenever $\lambda \sim [0.1, 0.3]$ we do not detect any biases in our estimates neither of α nor of $(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$. Last, we explore also a case where heterogeneity is not in λ as in Wilcox (2006), but instead in α (bottom panel in Table B.2) . Specifically we assume $\alpha \sim [0, 1]$. In this case we do not find systematic biases.

[3]Cabrales and Garcia-Fontes (2000) and Wilcox (2006) discuss estimating parameters from EWA (experience weighted attraction learning), in particular a parameter that measures how much weight participants give to counterfactual information in the reinforcement values.

	Without heterogeneity					
	α	λ	$(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$	α	λ	$(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$
True	0	1	0	0	0.2	0
Estimate	0.005 (0.010)	0.998 (0.006)	-0.006 (0.010)	-0.012 (0.039)	0.202 (0.002)	0.011 (0.036)
Bias	NO	NO	NO	NO	NO	NO
True	0.25	1	0	0.25	0.2	0
Estimate	0.248 (0.011)	0.999 (0.007)	-0.006 (0.009)	0.233 (0.030)	0.198 (0.002)	0.056 (0.056)
Bias	NO	NO	NO	NO	NO	NO
True	0.5	1	0	0.5	0.2	0
Estimate	0.505 (0.015)	0.996 (0.013)	0.004 (0.004)	0.484 (0.039)	0.197 (0.002)	0.058 (0.032)
Bias	NO	NO	NO	NO	NO	NO
True	0.75	1	0	0.75	0.2	0
Estimate	0.768 (0.007)	1.003 (0.005)	0.007 (0.007)	0.750 (0.001)	0.200 (0.001)	0.048 (0.048)
Bias	0.018	NO	NO	NO	NO	NO
True	1	1	0	1	0.2	0
Estimate	1.000 (0.007)	1.009 (0.005)	0.000 (0.007)	0.979 (0.010)	0.200 (0.002)	-0.020 (0.026)
Bias	NO	NO	NO	NO	NO	NO
N	240 000					

Table B.1: Simulations of the extended learning model without heterogeneity in the underlying true parameters. "True" is the parameter value used to generate the data. "Estimate" is our estimate with standard deviation in brackets. "Bias" indicates whether there is bias, i.e. whether the estimates are statistically different from the true value at the 5% level and if so the sign and size of the bias.

	With heterogeneity					
	α	λ	$(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$	α	λ	$(\psi^{i,0}_{GT} - \psi^{i,0}_{AD})$
True	0	$\sim [0.5, 1.5]$	0	0	$\sim [0.1, 0.3]$	0
Estimate	0.027 (0.012)	0.854 (0.006)	0.076 (0.012)	0.062 (0.041)	0.192 (0.002)	0.043 (0.038)
Bias	0.027	-0.146	0.076	NO	-0.008	NO
True	0.25	$\sim [0.5, 1.5]$	0	0.25	$\sim [0.1, 0.3]$	0
Estimate	0.296 (0.012)	0.852 (0.005)	0.076 (0.011)	0.273 (0.034)	0.193 (0.002)	-0.020 (0.034)
Bias	0.046	-0.148	0.076	NO	-0.006	NO
True	0.5	$\sim [0.5, 1.5]$	0	0.5	$\sim [0.1, 0.3]$	0
Estimate	0.524 (0.016)	0.875 (0.005)	0.022 (0.009)	0.502 (0.039)	0.192 (0.001)	0.043 (0.032)
Bias	NO	-0.125	0.022	NO	-0.007	NO
True	0.75	$\sim [0.5, 1.5]$	0	0.75	$\sim [0.1, 0.3]$	0
Estimate	0.752 (0.006)	0.876 (0.005)	0.038 (0.008)	0.750 (0.001)	0.191 (0.002)	0.036 (0.030)
Bias	NO	-0.124	0.038	NO	-0.008	NO
True	1	$\sim [0.5, 1.5]$	0	1	$\sim [0.1, 0.3]$	0
Estimate	1.002 (0.001)	0.907 (0.005)	-0.020 (0.006)	1.006 (0.002)	0.195 (0.027)	0.002
Bias	NO	-0.093	-0.020	NO	-0.005	NO
True	$\sim [0, 1]$	1	0	$\sim [0, 1]$	0.2	0
Estimate	0.539 (0.018)	0.986 (0.006)	-0.002 (0.008)	0.505 (0.009)	0.199 (0.006)	0.004 (0.006)
Bias	NO	-0.014	NO	NO	NO	NO
N	240 000					

Table B.2: Simulations of the extended learning model with heterogeneity in the underlying true parameters. “True” is the parameter value used to generate the data. When there is heterogeneity the individual parameter values are always drawn from a uniform distribution on the interval indicated. The true mean is hence the midpoint of the interval. “Estimate” is our estimate with standard deviation in brackets. “Bias” indicates whether there is bias, i.e. whether the estimates are statistically different from the true value at the 5% level and if so the sign and size of the bias.

C Additional Information Experiments

We conducted our own experiment to complete Case Study I (Section 3.1). In this Appendix we provide additional information on these experiments. Our experiments were conducted at Essex Lab at the University of Essex. The deterministic, replication and inverse treatments were conducted in February/March 2017 and the MatchStoch treatment in May 2019. Participants were students (and some non-students) who signed up for lab experiment at EssexLab at the University of Essex. They were recruited using recruitment software hroot. Table C.1 shows some demographics of our participants as well their answers to a post experimental questionnaire. We used exactly the same questionnaire as Bigoni et al. (2015), but replaced the question about whether people were born in Italy with whether they were born in the UK.

	Sample Characteristics			
	Rep	Inverse	MatchStoch	Det
Age	26.54	24.25	22.76	25.27
From UK (1=yes)	0.38	0.23	0.37	0.44
Gender (1=female)	0.47	0.44	0.42	0.50
Risk Attitude (0-10)	6.54	5.89	6.16	6.23
Trust (0-1)	0.38	0.39	0.33	0.42
Logic1 (0-1)	0.64	0.58	0.70	0.67
Logic2 (0-1)	0.33	0.31	0.52	0.40

Table C.1: Basic Sample Characteristics of Lab Experimental Sample used in Section 3.1. Mean Age, fraction of participants born in UK, fraction female, mean risk attitude (0, most risk averse, 10 least risk averse), fraction displaying high trust, fraction answering Logic 1 question correctly and fraction answering Logic 2 question correctly.

In all the experiments we followed exactly the same procedures used by Bigoni et al. (2015) including using the exact same Instructions and software (translated from Italian to English). The only change made to the software was to change the match length realizations (i) to be drawn at the match level in session MatchStoch and (ii) inversed in treatment "Inverse" (see Figure E.4). We use the draws generated by Bigoni et al. (2015) in the "Replication" and "Inverse" treatments. As the interest of the study is match length realization, there was no other way to conduct an exact replication. The procedure are in line with what was communicated to participants which is the following: "*How is a period duration established? The period may stop at every tick of 0.16 seconds. This event depends on the results of a random draw...*". No further information was given about when this draw took/will take place.

D Additional Tables

This Appendix contains additional tables.

D.1 Additional Tables for Section 2.2

In Table D.1 we consider separately the effect of match length realization of the 1st tenth, ninth, eight,..., half of matches on first-stage cooperation rates in the remaining matches. The table shows that there is a positive effect of match length realization in all these cases. However, the effect is smaller for the first tenth, ninth,...,sixth of matches compared to our baseline specification using the first third of matches. This shows that, while the very first matches are very important there is still learning and match length realizations become more important as more early matches are aggregated. Using the 1st half of matches, however, does not lead to a larger effect than using the 1st third.

1st...	(1) tenth	(2) ninth	(3) eight	(4) seventh	(5) sixth	(6) fifth	(7) fourth	(8) third	(9) half
Δ^{1st}_{above}	0.125**	0.081	0.020	0.055	0.023	0.078	0.181***	0.226***	0.165***
	(0.057)	(0.059)	(0.058)	(0.061)	(0.060)	(0.063)	(0.061)	(0.056)	(0.073)
SizeBAD	-0.592***	-0.692***	-0.781***	-0.815***	-0.829***	-0.744***	-0.628***	-0.539***	-0.688***
	(0.0726)	(0.0769)	(0.0967)	(0.0964)	(0.0867)	(0.116)	(0.120)	(0.0945)	(0.152)
SizeBAD $\times$ Δ^{1st}_{above}	-0.189**	-0.056	0.053	-0.066	-0.105	-0.028	-0.212***	-0.296***	-0.103
	(0.081)	(0.082)	(0.093)	(0.077)	(0.084)	(0.116)	(0.073)	(0.094)	(0.163)
Constant	0.852***	0.916***	1.003***	1.038***	1.051***	0.965***	0.843***	0.747***	0.886***
	(0.0574)	(0.0750)	(0.0977)	(0.0992)	(0.0901)	(0.118)	(0.123)	(0.0987)	(0.155)
Observations	45,869	45,469	44,773	43,955	43,481	41,873	38,778	34,319	25,795
R-squared	0.188	0.188	0.189	0.195	0.196	0.199	0.212	0.223	0.237

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.1: Initial (first stage) cooperation rate in remaining matches depending on whether at least $\frac{2}{3}$ of first X-th of matches where above theoretical median length.

Table D.2 compares the impact of match length realization in the first X-th, second X-th, third X-th,... of matches on cooperation in the last X-th of matches, where X ranges from 2,...,10. The coefficient Δ^{1st}_{above} is statistically significant in seven out of these nine different splits while all other dummies are statistically significant in fewer than 50 percent of the cases where they are included. Δ^{1st}_{above} also has the biggest coefficient size in all cases $X = 2, ..., 6$ and the second-biggest coefficient size when $X = 7, .., 10$, though it should be noted that these differences are not always statistically significant. Furthermore for all $X = 2, ..., 9$ the coefficient corresponding to the first X-th of matches is bigger than the one corresponding to the most recent (the $X-1$th) X-th of matches. As the finer splits often involve only very few matches in each group we consider the cases $X = 2, ..., 6$ to be more meaningful. The table hence emphasizes the point that early match length realization is at least as important or more important than match length realization in more recent matches.

Table D.3 shows the results of a placebo test, where we regress cooperation in the 1st third of matches on Δ^{3rd}_{above} (as well as `SizeBAD` and an interaction). We would not expect realized match length of final matches, which have not yet been played, to affect cooperation in the beginning of the experiment. Hence we would expect zero coefficients on Δ^{3rd}_{above} and the corresponding interaction term. We do indeed find that these coefficients are close to zero and statistically

1st...	(1) tenth	(2) ninth	(3) eight	(4) seventh	(5) sixth	(6) fifth	(7) fourth	(8) third	(9) half
Δ^{1st}_{above}	0.010	0.057*	0.048	0.082**	0.065*	0.077**	0.105***	0.101**	0.123***
	(0.037)	(0.031)	(0.033)	(0.035)	(0.034)	(0.039)	(0.036)	(0.041)	(0.044)
Δ^{2nd}_{above}	0.051	0.026	0.022	0.037	0.017	0.075**	0.097**	0.069*	
	(0.03)	(0.041)	(0.037)	(0.034)	(0.045)	(0.036)	(0.038)	(0.039)	
Δ^{3rd}_{above}	-0.014	-0.015	0.019	0.087**	0.048	0.076*	0.074**		
	(0.039)	(0.039)	(0.038)	(0.037)	(0.036)	(0.040)	(0.035)		
Δ^{4th}_{above}	0.070*	0.103***	0.017	0.003	0.050	0.075**			
	(0.038)	(0.035)	(0.037)	(0.034)	(0.037)	(0.036)			
Δ^{5th}_{above}	-0.000	0.001	0.091**	0.006	0.014				
	(0.035)	(0.032)	(0.040)	(0.045)	(0.044)				
Δ^{6th}_{above}	0.033	0.044	0.047	0.001					
	(0.033)	(0.043)	(0.039)	(0.039)					
Δ^{7th}_{above}	-0.004	0.012	0.022						
	(0.041)	(0.040)	(0.034)						
Δ^{8th}_{above}	-0.025	0.030							
	(0.033)	(0.037)							
Δ^{9th}_{above}	0.050								
	(0.037)								
SizeBAD	-0.824***	-0.814***	-0.806***	-0.806***	-0.809***	-0.805***	-0.805***	-0.809***	-0.769***
	(0.0808)	(0.0729)	(0.0772)	(0.0875)	(0.0796)	(0.0777)	(0.0734)	(0.0681)	(0.0670)
Constant	0.989***	0.966***	0.965***	0.969***	0.977***	0.973***	0.983***	0.994***	0.967***
	(0.0940)	(0.0789)	(0.0836)	(0.0926)	(0.0852)	(0.0823)	(0.0798)	(0.0747)	(0.0736)
Observations	5,905	7,338	7,506	8,296	9,148	11,721	13,573	18,536	25,795
R-squared	0.257	0.273	0.266	0.269	0.262	0.268	0.265	0.251	0.236

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.2: Initial (first stage) cooperation rate in remaining matches depending on whether at least $\frac{2}{3}$ of first X-th, second X-th, third X-th,... of matches where above theoretical median length.

not significant. This shows that the results in Table 1 are fundamental and not e.g. driven by correlations of match lengths within sessions or observed or unobserved heterogeneity across papers or treatments, e.g. caused by different ways researchers implement match length draws.

	Placebo test (1)	(2)	(3)
Δ^{3rd}_{above}	0.038	-0.054	-0.053
	(0.055)	(0.061)	(0.060)
SizeBAD		-0.643***	-0.650***
		(0.061)	(0.078)
SizeBAD × Δ^{3rd}_{above}			0.009
			(0.055)
Constant	0.365***	0.908***	0.915***
	(0.045)	(0.070)	(0.085)
δ f.e.	NO	YES	YES
Observations	29,467	29,467	29,467
R-squared	0.002	0.141	0.141

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.3: Placebo test. Initial (first stage) cooperation rate in the *1st* third of matches explained by dummy Δ^{3rd}_{above} indicating whether more than $\frac{2}{3}$ of matches in the 3rd third of the experiment were longer than the theoretical median match length.

Appendix Table D.4 reproduces our main results reported in Table 1 using paper fixed effects.

In Appendix Table D.5 we use the share of matches above theoretical median length instead of a dummy variable.

	Main Result with paper fixed effects					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ^{1st}_{above}	0.115**	0.120***	0.267***	0.120*	0.119***	0.247***
	(0.055)	(0.040)	(0.053)	(0.060)	(0.040)	(0.058)
Δ^{2nd}_{above}				0.025	0.068	0.058
				(0.067)	(0.042)	(0.061)
`SizeBAD`		-0.746***	-0.497***		-0.808***	-0.573***
		(0.075)	(0.084)		(0.075)	(0.148)
`SizeBAD` $\times \Delta^{1st}_{above}$			-0.328***			-0.291***
			(0.084)			(0.094)
`SizeBAD` $\times \Delta^{2nd}_{above}$						-0.015
						(0.132)
Constant	0.826***	1.006***	0.757***	0.227***	1.101***	0.946***
	(0.0147)	(0.0782)	(0.0869)	(0.0215)	(0.120)	(0.161)
δ f.e.	NO	YES	YES	NO	YES	YES
paper f.e.	YES	YES	YES	YES	YES	YES
Test Δ^{1st}_{above}=Δ^{2nd}_{above}	-	-	-	0.3477	0.4220	0.0508
Observations	34,319	34,319	34,319	18,536	18,536	18,536
R-squared	0.104	0.239	0.245	0.122	0.274	0.278

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.4: Columns (1)-(3): Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by dummy Δ^{1st}_{above} indicating whether more than $\frac{2}{3}$ of matches in the 1st third of the experiment were longer than the theoretical median match length. Columns (4)-(6): Initial (first stage) cooperation rate in the 3rd third of matches explained by dummies Δ^{1st}_{above} and Δ^{2nd}_{above}. Standard errors clustered at session level. Observations stem from 103 sessions spread across 11 papers. With paper fixed effects.

	Share of matches above median					
	(1)	(2)	(3)	(4)	(5)	(6)
$Share^{1st}_{above}$	0.218***	0.206**	0.309***	0.072	0.210**	0.310**
	(0.068)	(0.079)	(0.113)	(0.112)	(0.080)	(0.120)
$Share^{2nd}_{above}$				0.235**	0.151	0.111
				(0.112)	(0.094)	(0.130)
`SizeBAD`		-0.764***	-0.446		-0.813***	-0.573
		(0.0759)	(0.383)		(0.0754)	(0.424)
`SizeBAD` $\times$ $Share^{1st}_{above}$			-0.441			-0.434
			(0.490)			(0.498)
`SizeBAD` $\times$ $Share^{2nd}_{above}$						0.104
						(0.281)
Constant	0.290***	1.023***	0.706*	0.250***	1.041***	0.801*
	(0.031)	(0.079)	(0.384)	(0.030)	(0.078)	(0.425)
δ f.e.	NO	YES	YES	NO	YES	YES
Test $Share^{1st}_{above}$=$Share^{2nd}_{above}$	-	-	-	0.4465	0.6338	0.2631
Observations	34,319	34,319	34,319	18,536	18,536	18,536
R-squared	0.023	0.236	0.237	0.044	0.271	0.271

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.5: Columns (1)-(3): Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by share of matches in the 1st third of the experiment were longer than the theoretical median match length. Columns (4)-(6): Initial (first stage) cooperation rate in the 3rd third of matches explained by share of matches longer than theoretical median length in 1st and 2nd third of matches. Standard errors clustered at session level. Observations stem from 103 sessions spread across 11 papers.

In Appendix Table D.6 we conduct the same analysis as Table 1 with the only difference that Δ^{1st}_{above} takes the value 1 if more than half (rather than $\frac{2}{3}$) of early matches are above median. It is hence a much weaker test as match length realizations with more or less than half of matches above median need not differ by much. In the most extreme case they might differ by one round only, which is not much if matches are long, i.e. δ is high. On the other hand it should allow us to better identify the effect of shorter match length when $\delta = 0.5$, as there will be more variation in the dummy in this case. Indeed we do find that the dummy defined in this way is more effective in capturing differences when $\delta = 0.5$. Overall the effect of shorter matches reported in column (1) is now a 38% increase in cooperation rates compared to the 43% increase identified in Table 1. Table D.7 goes the opposite direction and shows results when a dummy is used indicating whether more than $\frac{3}{4}$ of early matches were above median. Again results are similar, even though this dummy takes the value 1 much less often.

	Reduced Threshold					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ^{1st}_{above}	0.120**	0.142***	-0.095	0.051	0.143**	-0.102
	(0.050)	(0.051)	(0.085)	(0.058)	(0.057)	(0.091)
Δ^{2nd}_{above}				0.157**	0.064	0.076
				(0.062)	(0.054)	(0.094)
`SizeBAD`		-0.792***	-1.967***		-0.826***	-2.039***
		(0.066)	(0.286)		(0.067)	(0.322)
`SizeBAD` $\times$ Δ^{1st}_{above}			1.184***			1.239***
			(0.279)			(0.321)
`SizeBAD` $\times$ Δ^{2nd}_{above}						-0.018
						(0.163)
Constant	0.320***	1.000***	2.175***	0.277***	1.011***	2.225***
	(0.032)	(0.072)	(0.287)	(0.032)	(0.073)	(0.324)
δ f.e.	NO	YES	YES	NO	YES	YES
Test Δ^{1st}_{above}=Δ^{2nd}_{above}	-	-	-	0.3178	0.2876	0.1933
Observations	34,319	34,319	34,319	18,536	18,536	18,536
R-squared	0.014	0.219	0.223	0.036	0.249	0.254

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.6: Reduced Threshold. Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by dummy Δ^{1st}_{above} indicating whether more than $\frac{1}{2}$ of matches in the 1st third of the experiment were longer than the theoretical median match length.

In Appendix Table D.8 we ask whether unusually short early matches or unusually long early matches have a larger effect. To this end we reproduce Table 1 using now a dummy Δ^{1st}_{below} which takes the value 1 if more $\frac{2}{3}$ of early matches are below the theoretical median length. For the sessions identified by this dummy cooperation rates are 32% lower in specification (1), indicating that the overall effect of match length realizations is broadly symmetric.

	Increased Threshold					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ^{1st}_{above}	0.192***	0.054	0.147**	0.149**	0.044	0.139**
	(0.067)	(0.038)	(0.059)	(0.070)	(0.038)	(0.059)
Δ^{2nd}_{above}				0.181***	0.079**	0.112*
				(0.062)	(0.038)	(0.063)
`SizeBAD`		-0.743***	-0.671***		-0.766***	-0.661***
		(0.067)	(0.077)		(0.068)	(0.090)
`SizeBAD` $\times$ Δ^{1st}_{above}			-0.215***			-0.221**
			(0.081)			(0.085)
`SizeBAD` $\times$ Δ^{2nd}_{above}						-0.060
						(0.092)
Constant	0.346***	0.951***	0.879***	0.297***	0.951***	0.846***
	(0.030)	(0.073)	(0.082)	(0.029)	(0.074)	(0.095)
δ f.e.	NO	YES	YES	NO	YES	YES
Test Δ^{1st}_{above}=Δ^{2nd}_{above}	-	-	-	0.3477	0.4220	0.050
Observations	34,319	34,319	34,319	18,536	18,536	18,536
R-squared	0.031	0.215	0.218	0.064	0.249	0.252

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.7: Reduced Threshold. Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by dummy Δ^{1st}_{above} indicating whether more than $\frac{3}{4}$ of matches in the 1st third of the experiment were longer than the theoretical median match length.

	Short matches					
	(1)	(2)	(3)	(4)	(5)	(6)
Δ^{1st}_{below}	-0.144***	-0.042	-0.145**	-0.063	-0.045	0.092
	(0.049)	(0.051)	(0.061)	(0.080)	(0.057)	(0.074)
Δ^{2nd}_{below}				-0.138	-0.041	0.040
				(0.083)	(0.060)	(0.102)
`SizeBAD`		-0.759***	-0.750***		-0.799***	-0.775***
		(0.066)	(0.067)		(0.068)	(0.071)
`SizeBAD` $\times$ Δ^{1st}_{below}			-1.093***			-0.768***
			(0.188)			(0.239)
`SizeBAD` $\times$ Δ^{2nd}_{below}						-0.228
						(0.152)
Constant	0.442***	1.010***	1.907***	0.468***	1.071***	1.824***
	(0.036)	(0.097)	(0.150)	(0.039)	(0.127)	(0.185)
δ f.e.	NO	NO	YES	NO	NO	YES
Test Δ^{1st}_{below}=Δ^{2nd}_{below}	-	-	-	0.6325	0.9650	0.6977
Observations	34,319	34,319	34,319	18,536	18,536	18,536
R-squared	0.018	0.213	0.215	0.033	0.242	0.244

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.8: Focus on Short Matches. Initial (first stage) cooperation rate in the 2nd and 3rd third of matches explained by dummy Δ^{1st}_{below} indicating whether more than $\frac{2}{3}$ of matches in the 1st third of the experiment were shorter than the theoretical median match length.

D.2 Additional Tables for Section 3.1

	Average Cooperation Rates			
	(1)	(2)	(3)	(4)
Replication	-14.46***	-14.72***		
	(4.722)	(4.955)		
Inverse	-3.142	-3.582		
	(3.555)	(3.548)		
Match Stochastic	-6.641	-6.956	-6.923	-7.210
	(5.154)	(5.059)	(5.093)	(4.993)
L.Duration		-0.022		0.013
		(0.045)		(0.054)
Constant	54.04***	55.31***	42.03***	42.43***
	(3.454)	(3.497)	(8.895)	(8.445)
Demographics	NO	NO	YES	YES
Observations	6,624	6,336	4,416	4,224
Number of id	288	288	192	192

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.9: Random Effects OLS regression of average cooperation rate on treatment dummies and covariates (age, gender, birthplace, nationality). Standard errors clustered at session level. Columns (3) and (4) only use data from match stochastic and deterministic sessions.

	Initial Cooperation Rates			
	(1)	(2)	(3)	(4)
Replication	-22.06***	-22.44***		
	(3.921)	(3.756)		
Inverse	-0.589	-0.538		
	(3.367)	(3.451)		
Match Stochastic	-12.77**	-13.11**	-12.94**	-13.23**
	(6.306)	(6.391)	(6.429)	(6.533)
L.Duration		0.061		0.119*
		(0.041)		(0.065)
Constant	73.96***	73.68***	57.61***	56.58***
	(3.144)	(3.369)	(12.20)	(11.97)
Demographics	NO	NO	YES	YES
Observations	6,624	6,336	4,416	4,224
Number of id	288	288	192	192

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.10: Random Effects OLS regression of initial cooperation rate on treatment dummies and covariates (age, gender, birthplace, nationality). Standard errors clustered at session level. Columns (3) and (4) only use data from match stochastic and deterministic sessions.

D.3 Additional Tables for Section 3.2

	First Round Cooperation Rate				
Decision Setting	Finite	Prob All	Prob S1	Prob S23	S1 vs S23
$N = 4$, MPCR$= 0.3$	24.9	28.8	16.56	34.95	
		$<$	$>^{*}$	$<^{**}$	$p = 0.002$
$N = 4$, MPCR$= 0.6$	44.0	47.7	38.88	51.00	
		$<$	$>$	$<$	$p = 0.224$
$N = 2$, MPCR$= 0.6$	52.4	44.3	42.56	45.46	
		$>$	$>$	$>$	$p = 0.678$
$N = 2$, MPCR$= 0.6$, Binary	76.1	57.8	54.28	59.23	
		$>^{***}$	$>^{*}$	$>^{***}$	$p = 0.613$

Table D.11: Initial cooperation rates across all rounds in the finite sessions (column (1)) and across all sessions with a probabilistic ending (column (2)) as in Table 3 in Lugovskyy et al. (2017). We further split the sessions with probabilistic ending in those with initially short matches (S1) and those with initially long matches (S23). Below each cooperation rate we show how the finite setting compares to the rate in question (as in Lugovskyy et al. (2017)). The last column shows the p-value when comparing initially short and long sequences. Following Lugovskyy et al. (2017) standard errors are clustered at the participant level in all regressions.

D.4 Additional Tables for Section 3.3

with RO	All groups	Short first matches	Other Groups	Short vs Other
Substitutes	16.5	9.0	20.1	
Complements	23.2	34.9	17.3	
Treatment Effect	- 6.7	-25.2**	2.8	$p = 0.1280$
without RO	All groups	Short first matches	Other Groups	Short vs Other
Substitutes	17.6	20.9	15.9	
Complements	10.9	16.2	8.1	
Treatment Effect	6.7*	4.7	7.8	$p = 0.8949$

Table D.12: Efficiency measure from Table 2 (all matches) in Embrey et al. (2019). Treatment Effect is the difference between substitutes and complements. The two rightmost columns split out the groups with short initial matches from the rest.

D.5 Additional Tables for Section 4

	$\delta = 0.5$				$\delta = 0.75$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
mean duration	0.021		-0.001	0.022	0.015*		-0.004	-0.002
	(0.070)		(0.077)	(0.075)	(0.008)		(0.016)	(0.012)
median duration		0.104*	0.105	0.087		0.020**	0.024*	0.035**
		(0.063)	(0.082)	(0.105)		(0.008)	(0.016)	(0.016)
SizeBAD	-0.662***	-0.704***	-0.705***	-0.642***	-0.825***	-0.847***	-0.848***	-0.965***
	(0.113)	(0.113)	(0.103)	(0.089)	(0.071)	(0.065)	(0.064)	(0.077)
Constant	0.665***	0.519***	0.521***	0.527**	0.776***	0.753***	0.759***	0.575***
	(0.150)	(0.174)	(0.151)	(0.201)	(0.061)	(0.062)	(0.061)	(0.098)
paper fixed effects	NO	NO	NO	YES	NO	NO	NO	YES
Observations	10,786	10,786	10,786	10,786	12,380	12,380	12,380	12,380
R-squared	0.151	0.160	0.160	0.173	0.185	0.189	0.189	0.225

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table D.13: First round cooperation in the 2nd and 3rd third of matches regressed on mean and median match length in the 1st third. We selected the two discounted factors with most papers in the meta-study (8 papers for $\delta = 0.75$ and 4 papers for $\delta = 0.5$).

E Additional Figures

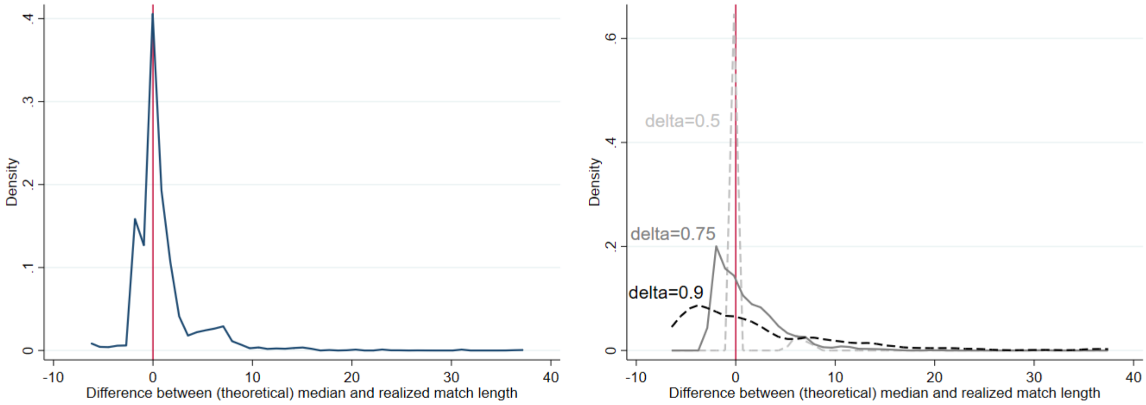

Figure E.1: Meta Data used in Section 2.2: Kernel density estimates of the difference between theoretical median match length and realized match length overall (left panel) and separately for $\delta = 0.5, 0.75$ and $\delta = 0.9$ (right panel).

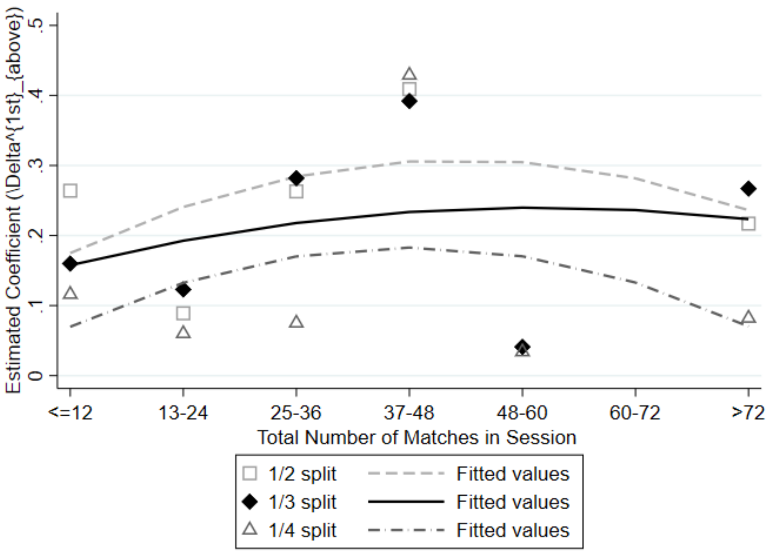

Figure E.2: Section 2.2 The Effect of Experience (Number of Matches Played in a Session) on Estimated Coefficient Δ^{1st}_{above}. Black diamonds show the estimated coefficient Δ^{1st}_{above} on cooperation in the last third of matches depending on the total number of matches played in the session. The black line shows fitted values using a square polynomial. Squares and light gray dashed line show the effect of the first half of matches on the second half and triangles and dark gray dashed line the effect of the first quarter on the last quarter of matches.

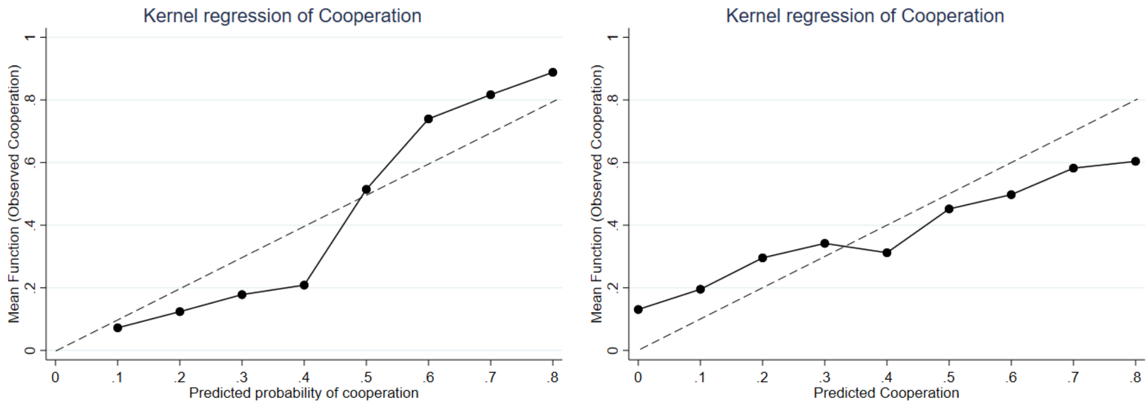

Figure E.3: Kernel density estimates of observed cooperation in stage 1 of a match depending on the predicted probability of cooperation in stage 1, i.e. the predicted probability of Grim Trigger ($p^{i,m}$). Left panel shows the basic model and right panel the extended model. The dashed line is the 45 degree line showing zero prediction error. (Note that the x-axis range differs from the y-axis range).

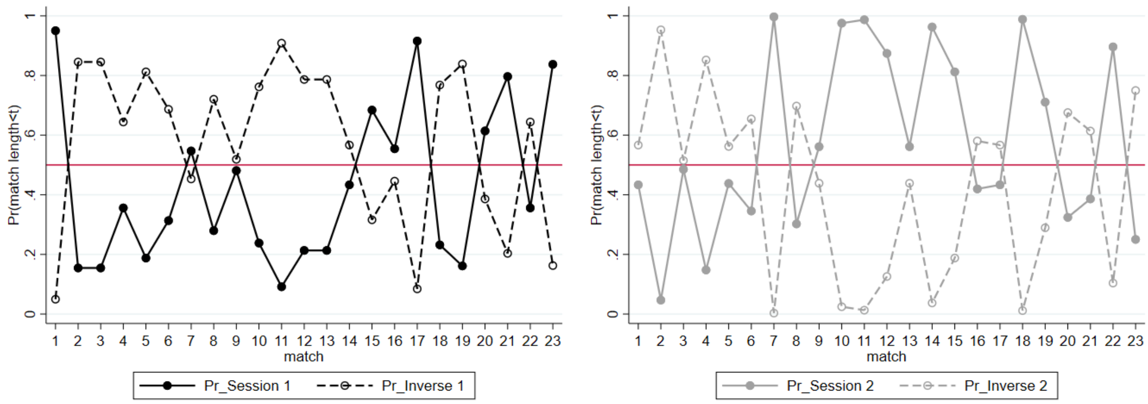

Figure E.4: Section 3.1: Illustration of how sequence of “inverse” match lengths is generated for the two sessions.

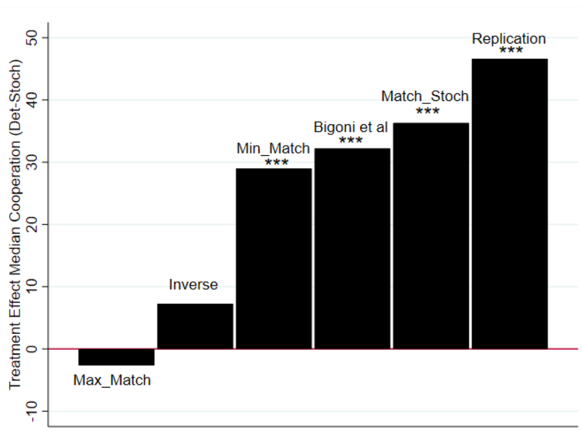

Figure E.5: Different Effect sizes Median Cooperation Frequency Det-Stoch.

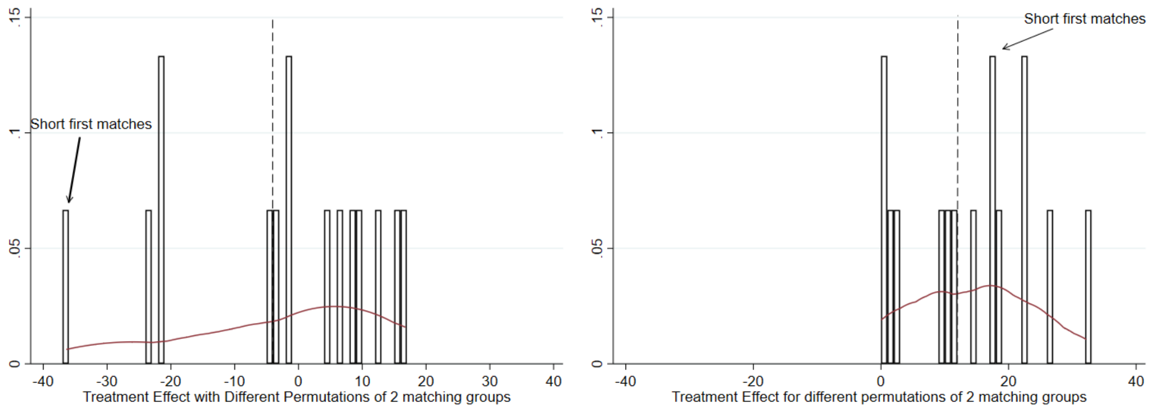

Figure E.6: Section 3.3: Illustration of treatment effect for different permutations of two matching groups. Histogram and estimated kernel density. Left Panel: treatments with revision opportunities. Right panel: treatments without revision opportunities.