Online Appendix for

Beliefs in Repeated Games

Masaki Aoyagi   Guillaume R. Fréchette   Sevgi Yuksel

CONTENTS:

A Related Literature

The experimental literature studying beliefs in one-shot games has focused on two important questions. The first investigates whether beliefs are correct and, more generally, what factors or features of the game impact beliefs. The second studies the extent to which behavior in a game best-responds to subjective beliefs. Nyarko and Schotter [2002] are among the first to study elicited beliefs in repeated games. Studying a finite repetition of a 2×2 game with a unique mixed Nash equilibrium (NE) played in fixed and random pairing, Nyarko and Schotter [2002] find the subjects' beliefs over actions are not empirical in the sense that they cannot be approximated by the weighted average of the opponent's past actions.[38] Following this, many papers that elicited beliefs have focussed on factors that determine beliefs. For instance, Hyndman et al. [2010] study beliefs when a stage game with a unique mixed NE (and two pure NE) is repeated 20 times, and find subjects' beliefs about the other's action in the present round do take into account the effect of their own action choice in the preceding rounds, and hence cannot be expressed by the weighted average of the other player's actions in the past. Hyndman et al. [2012b] advance this observation in an experiment in which subjects play a finite repetition of 3×3 and 4×4 normal form games with and without dominance-solvable NE. Hyndman et al. [2012b] note some players attempt to influence the beliefs of other players through their own actions, and thus help the process converge to an NE.[39]

Some of these, as well as other papers, in the experimental literature on beliefs examine the question of whether actions are best responses to beliefs with no definite answers. Nyarko and Schotter [2002] find the actions in each round mostly best respond to the stated beliefs, but also find fictitious-play beliefs better predict the opponents' action than the stated beliefs. Costa-Gomes and Weizsäcker [2008] use 14 3×3 games to investigate the relationship between subjects' elicited beliefs and their strategy choice. Regardless of whether belief elicitation precedes strategy choice, Costa-Gomes and Weizsäcker [2008] find the strategies are not best responses to the beliefs in a half of the games, and attribute this finding to the difference in the perception of the game in the two situations. Danz et al. [2012] use a dominance-solvable 3×3 game repeated 20 times to study beliefs under various combinations of feedback and matching conditions. Danz et al. [2012] find feedback of past actions helps advance the iterative elimination process both in terms of actions and beliefs.

[38] Nyarko and Schotter [2002] specifically consider a generalization of fictitious play called the γ-empirical average as proposed by Cheung and Friedman [1997].

[39] Hyndman et al. [2012a] have outside observers predict the actions of the subjects in Hyndman et al. [2012b], and find a large variance in their beliefs both in terms of accuracy and updating.

Table 4: Experiments Eliciting Beliefs in One-Shot Games

Article	Pairing	Repetitions	Games	Equilibria	Feedback	Best Response
Nyarko and Schotter (2002)	Fixed (exp. 1) Random (exp. 3)	60	One 2x2	Unique Mixed NE	Yes	75% (exp. 1) 79% (exp. 3)
Costa-Gomes and Weizsäcker (2008)	Random	None	14 3x3	Unique Pure NE	No	54%
Rey-Biel (2009)	Random	None	10 3x3	Unique Pure NE	No	67%
Hyndman et al. (2010)	Fixed	20	Four 2x2	Two Pure and one Mixed NE	Yes	74%
Danz et al. (2012)	Random (RM) Fixed (FP)	20	One 3x3	Unique Pure NE	Yes (RM+FM) No (NF)	63%
Hyndman et al. (2012)	Fixed	20+20	Two 3x3 Two 4x4	Unique Pure (+ mixed for some)	Yes	Periods 1-10: 60% and 49% Periods 11-20: 73% and 63%
Manski and Neri (2013)	Random	Four	2x2	Unique Mixed NE	Yes	89%
Hyndman et al. (2022)	Random	None	12 3x3	One or two Pure NE	No	60%

Using a series of 3×3 games each with a unique NE, Rey-Biel [2009] find more than two-thirds of subjects choose actions that best respond to their elicited beliefs.

Table 4 summarizes basic information about these papers (and a few more). In particular, even though it was not necessarily the focus of all of these papers, for each of them we can obtain the percentage of best response behavior. This reveals one interesting pattern: studies where the game is not played multiple times or that give no feedback [Costa-Gomes and Weizsäcker, 2008, Hyndman et al., 2022, Danz et al., 2012, Rey-Biel, 2009] have lower rates of best response.[40] Also in line with this observation is the fact that Hyndman et al. [2012a] reports increasing rates of best response behavior as experience increases. In that paper, for instance, a rate of 63% is much higher than random given that the games are four-by-fours. Overall, these suggests that subjects best-respond at fairly high rates when given experience and feedback.

The literature on voluntary-contribution games often finds conditional cooperation, which refers to the fact that subjects make higher contributions if they believe other members of their group make higher contributions. This relationship is observed, for example by Gächter and Renner [2010], Fischbacher and Gächter [2010] and Kocher et al. [2015].[41] Neugebauer et al. [2009] confirm this relationship in their

[40]The rate for Danz et al. [2012] mixes treatments with feedback and one without.

[41]Costa-Gomes et al. [2014] analyze the relationship in the trust game. Smith [2013, 2015] note the beliefs are endogenous, and hence that the effect on contribution, if interpreted as causal, is

experiment on a finitely repeated voluntary contribution game, and further observe that both contribution levels and beliefs about others' contribution levels decline toward the end. Chaudhuri et al. [2017] observe similar joint dynamics of contribution and beliefs, allowing for heterogeneity across subjects and classifying them into types according to their initial beliefs about others' contributions.

Among those papers, Neugebauer et al. [2009], Gächter and Renner [2010], Fischbacher and Gächter [2010] all play 10 periods in fixed pairing with feedback. As such, although these papers do not focus on supergame strategies, they do provide a point of comparison by formally creating one finitely repeated game. All three papers report beliefs that are higher than the actual contributions. In the case of Neugebauer et al. [2009] and Fischbacher and Gächter [2010], Figures suggest that this is directionally true in every period (at the treatment level).

On cooperation and strategies in finitely and infinitely repeated PD, Dal Bó and Fréchette [2018] and Embrey et al. [2018] find some key patterns by re-analyzing data from a collection of laboratory experiments.[42] First, in finitely repeated PD, the fraction of threshold strategies increases with experience.[43] By the end, threshold strategies always account for the majority of the data, and use of the threshold strategies with lower thresholds increases with experience. This contributes to a (sometimes very) slow aggregate movement toward earlier defection.[44] In the finitely repeated PD, if the parameters are conducive to cooperation, round-one cooperation increases with experience, whereas last-round cooperation decreases with it.[45] Otherwise, cooperation remains low in all rounds. In indefinitely repeated PD, on the other hand, experience leads cooperation (in the first or last round) to almost any level, depending on how conducive the parameters are to cooperation. Experience also amplifies the magnitude of the effects of the parameters, although it does not change the direction of those effects. In most experiments with perfect monitoring, a few simple strategies account for more than 50% of the strategies used. They are "always defect" (AD), "always cooperate" (AC), "grim trigger" (Grim), "tit-for-tat" (TFT), and "suspicious-tit-for-tat" (STFT).[46] AD, Grim, and TFT are generally the

overestimated.

[42]Experimental research on the subject goes as far back as Flood [1952].

[43]A threshold strategy (with threshold $k \geq 2$) starts with C and plays like grim-trigger before round k, but reverts to the unconditional play of D from round k on.

[44]Embrey et al. [2018] find that in the treatment most conducive to cooperation (replicated by the finite treatment of this study), the modal round at which cooperation breaks down moves earlier approximately by one round every 10 supergames.

[45]A longer horizon T, a higher discount factor δ, a lower temptation payoff $1 + g$, or a higher sucker payoff $-\ell$ all induce more cooperation.

[46]Grim cooperates until a defection is observed, at which point it defects forever; TFT starts

most popular, and Grim becomes more popular with experience and appears to be a counterpart to the threshold strategies in finite games. The implementation error term in Grim also decreases with experience.[47] Experience also increases *responsiveness*, which is measured as the difference between the probability of cooperative action after cooperation by the other player and that after defection by the other player. This is documented in Aoyagi et al. [2019] and confirmed by Dal Bó and Fréchette [2018] in their analysis of the meta-data and new experiments: according to a simple regression, experience has a significant positive impact on responsiveness in 11 paper-treatments, whereas it is insignificant in 20 paper-treatments.[48]

There are many papers on repeated games in the laboratory. Two that are more directly relevant are Kagel and McGee [2016] and Cooper and Kagel [2023] because they both study the same PD payoff matrix, one finitely repeated for 10 rounds, the other indefinitely repeated with a 10% random termination (hence 10 rounds in expectation); while the rest of the procedures are the same. Both papers' main focus is the comparisons of individual play (the typical implementation) versus team play (two players together in each of the row and column player's role). The results of the individual play treatments show (for experienced subjects): similar levels of round one cooperation for finite and indefinite. 2) Cooperation rates that drop over rounds of a supergame when it is finite, but not when it is indefinite. 3) Almost no cooperation in the last round of the finite game. Both papers show that teams initially cooperate less, but learn to cooperate more; and their behavior over rounds is more stable. The initially lower cooperation rates for teams are consistent with the discontinuity effect from psychology. However, the literature in psychology fails to identify that with experience the effect is reversed, i.e. teams cooperate more than individuals.

---

by cooperating and thereafter matches what the other did in the previous round; STFT starts by defecting and thereafter matches what the other did in the previous round.

[47] See Dal Bó and Fréchette [2019], Tables 8 and A10.

[48] This analysis eliminates all data in within-subjects designs after a change in treatment and only preserves the initial treatment. Most of the insignificant cases have a small number of observations. One treatment sees a negatively significant impact, perhaps because of relatively low round-one cooperation at 0.36.

B Additional Details and Analysis on Actions and Round Beliefs

Table 5: Session Summary

				No. of Game Rounds				Total no. of
		No. of	No. of	Actions	Actions and Beliefs			Obs.
Treatment	Session	Subjects	Supergames	Only	Early		Late	Rounds
	1	20	12			8, 8,		96
	2	20	12			8, 8,		96
	3	20	13			8, 8, 8,		104
Finite	4	20	11	8, 8, 8, 8	*8, 8, 8*	8,	8, 8, 8	88
	5	20	13			8, 8, 8,		104
	6	20	13			8, 8, 8,		104
	7	18	12			8, 8, 8,		96
	8	20	12			8, 8,		96
	1	20	10	9, 7, 13, 7	*1, 2, 23,*		4, 1, 19	112
	2	20	9	8, 15, 7, 32	*2, 10,*		5, 1, 8	105
	3	18	7	8, 2, 3, 14	*25,*		17, 10	90
Indefinite	4	16	8	9, 7, 10, 13	*32,*		7, 7, 6	96
	5	14	12	7, 22, 7, 3	*2, 5, 8,*	4, 14,	9, 3, 10	119
	6	18	6	1, 31, 4, 3	*24,*		15	94
	7	18	10	5, 6, 7, 14	*30, 8, 5,*		4, 9, 4	109
	8	20	9	11, 1, 4, 13	*9, 5,*		2, 4, 2	81

302 subjects in total.
Payment: \$8 + choices from two supergames (pre/post) + beliefs in one.
Earnings from \$22.00 to \$63.75 (with an average of \$35.30).
How to read: In the Finite treatment, session 1 had 20 subjects, they played a total of 12 supergames: 4 supergames of 8 rounds without belief elicitation, in the remaining 8 supergames that follow and where beliefs are also elicited, the first three (each with 8 rounds) are labelled "Early" supergames, three (each with 8) rounds are labelled "Late" supergames, and the two in between (supergames 8 and 9—each having 8 rounds) fall in neither Early nor Late category. In total subjects in that treatment played 96 rounds.

We aimed for three supergames for both early and late supergame categories when possible. When that was not possible, we aimed to have a division of total rounds that was as balanced as possible.

Tables 6 and 7 show no statistically significant differences in the probability of cooperation in round one $\bar{x}^1$ for supergames where beliefs are elicited. The other regressors are variables that have been considered in similar analysis.

Table 6: Correlated Random Effects Probit
Determinants of Cooperation in Round One

	Finite	Finite	Finite	Indefinite	Indefinite	Indefinite
Beliefs Are Elicited	0.109	0.0427	0.0654	0.891***	0.175	0.188
	(0.119)	(0.265)	(0.294)	(0.129)	(0.219)	(0.280)
Supergame		0.0106	0.0131		0.156***	0.187***
		(0.0323)	(0.0431)		(0.0475)	(0.0532)
Other Cooperated in Previous Supergame			0.250***			0.624***
			(0.0661)			(0.181)
Cooperated in Supergame 1			2.571***			2.830***
			(0.754)			(0.649)
Risk Measure			0.0189***			0.00534
			(0.00691)			(0.00663)
Length of Previous Supergame						-0.00119
						(0.00807)
Constant	2.280***	2.253***	-0.509	1.322***	0.955***	-1.461***
	(0.569)	(0.551)	(0.334)	(0.297)	(0.338)	(0.567)
Observations	1936	1936	1778	1270	1270	1126

Standard errors clustered (at the session level) in parentheses. ***1%, **5%, *10% significance.
All variables refer to behavior in Round 1.
Risk Measure is equal to the number of boxes collected in the bomb task.

Table 7: Correlated Random Effects Probit (Marginal Effects)
Determinants of Cooperation in Round One

	Finite	Finite	Finite	Indefinite	Indefinite	Indefinite
Beliefs Are Elicited	0.0115	0.00454	0.00660	0.108***	0.0208	0.0193
	(0.0140)	(0.0285)	(0.0295)	(0.0263)	(0.0256)	(0.0288)
Supergame		0.00113	0.00132		0.0186***	0.0193***
		(0.00337)	(0.00439)		(0.00717)	(0.00497)
Other Cooperated in Previous Supergame			0.0252***			0.0642***
			(0.00559)			(0.0188)
Cooperated in Supergame 1			0.259***			0.291***
			(0.0790)			(0.0482)
Risk Measure			0.00191***			0.000549
			(0.000615)			(0.000676)
Length of Previous Supergame						-0.000122
						(0.000827)
Observations	1936	1936	1778	1270	1270	1126

Standard errors clustered (at the session level) in parentheses. ***1%, **5%, *10% significance.
All variables refer to behavior in Round 1.
Risk Measure is equal to the number of boxes collected in the bomb task.

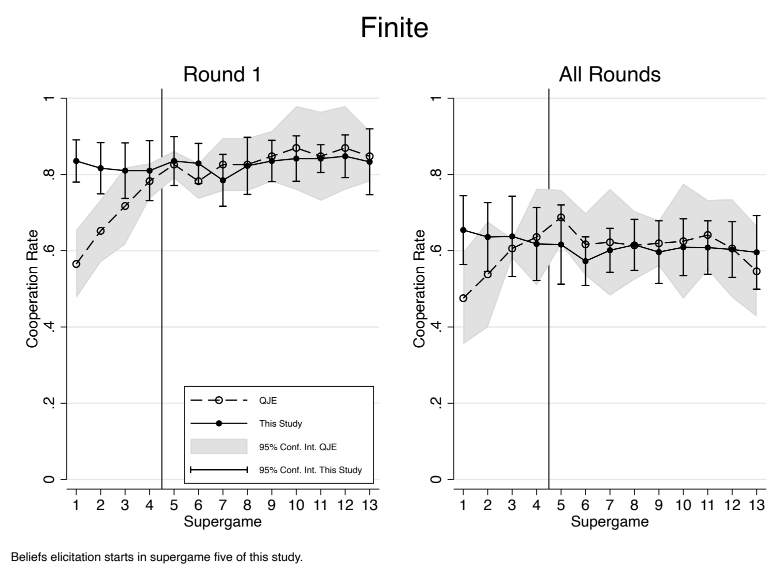


Figure 12: Comparison of Experiments With and Without Belief Elicitation

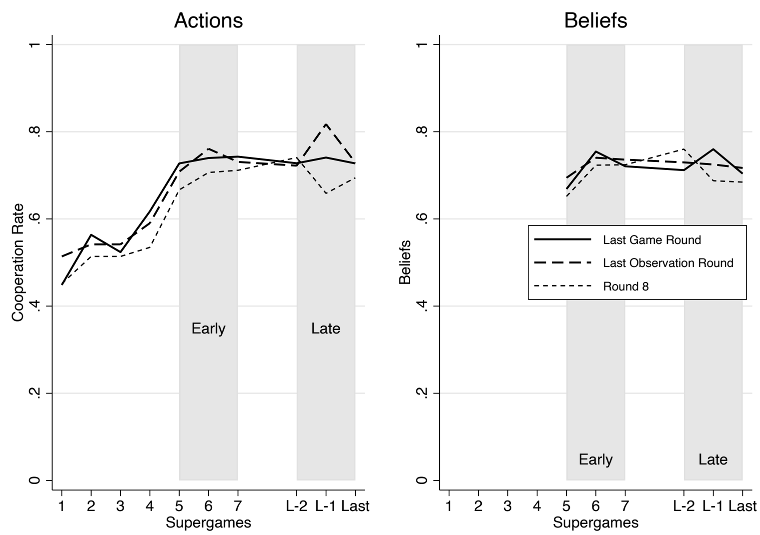


Figure 13: The Indefinite Game

In the Indefinite game, *observation rounds* refer to the rounds in which the subjects actually made action choices, and *game rounds* refer to those rounds that were part of the supergames. We denote by T the number of observation rounds in the Indefinite game so that $T = \max\{8, \text{“No. of game rounds”}\}$. For example, if an Indefinite game has five rounds, $T = 8$ because we observe the subject make eight choices even though only the first five mattered for payoffs, whereas if a supergame lasts 10 rounds, $T = 10$.

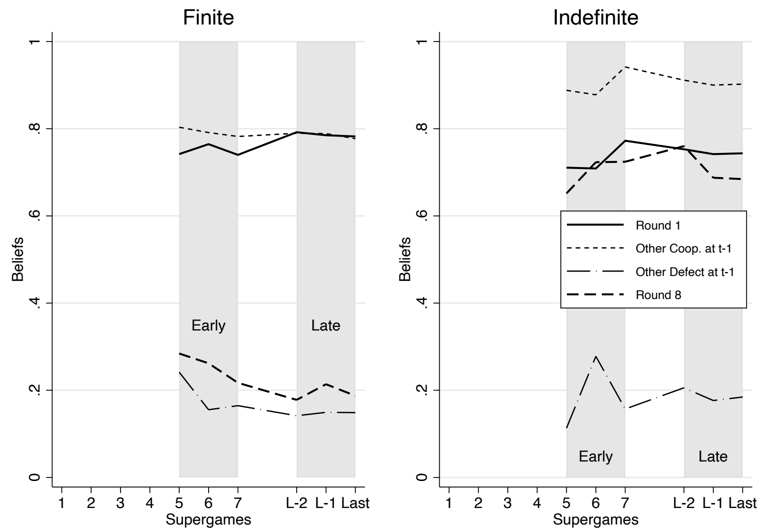


Figure 14: Beliefs Over Supergames

The evolution of beliefs depicted in Figure 14 mirrors the patterns observed for cooperation in Figure 1. $\bar{\mu}^1$ are high in both games. Beliefs are responsive in both games: $\bar{\mu}_i^t(*, a_j^{t-1} = C, *) - \bar{\mu}_i^t(*, a_j^{t-1} = D, *) > 0$. Beliefs $\bar{\mu}^T$ in the last round are low in the Finite game, but are high in the Indefinite game.

Contrasting aggregate bias in beliefs in early and late supergames

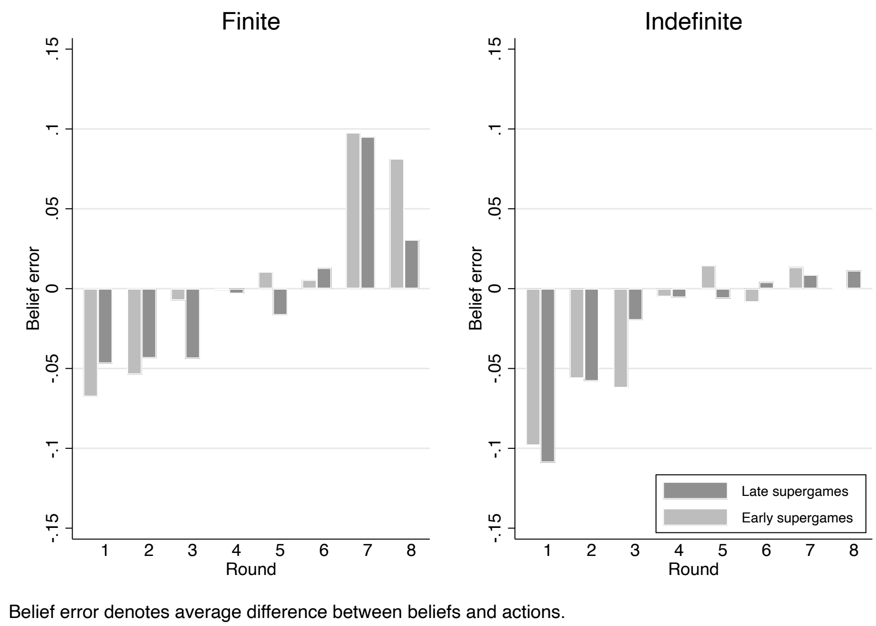


Figure 15: Belief Errors in Early vs. Late Supergames

One natural question is whether, with experience, subjects learn to correct their mispredictions. Figure 15 displays the error in key rounds for early versus late supergames. As the figure shows, in many cases where more substantial error occurs in early supergames, improvement is observed in late supergames, but not for round seven of the Finite game and round one of the Indefinite game. Even in these cases, however, subjects' beliefs do move in the right direction. As seen in Figure 16 which reports average cooperation rates and average beliefs for rounds one and seven over supergames, beliefs move in the correct direction with experience, but not fast enough to catch up with the changes in actions. We should note, however, that the changing behavior over the course of the session does not always imply beliefs are systematically off. For instance, in that same figure, one can see cooperation rates in round seven of the Indefinite game are changing with experience, but subjects correctly anticipate this change, as reflected in their beliefs.

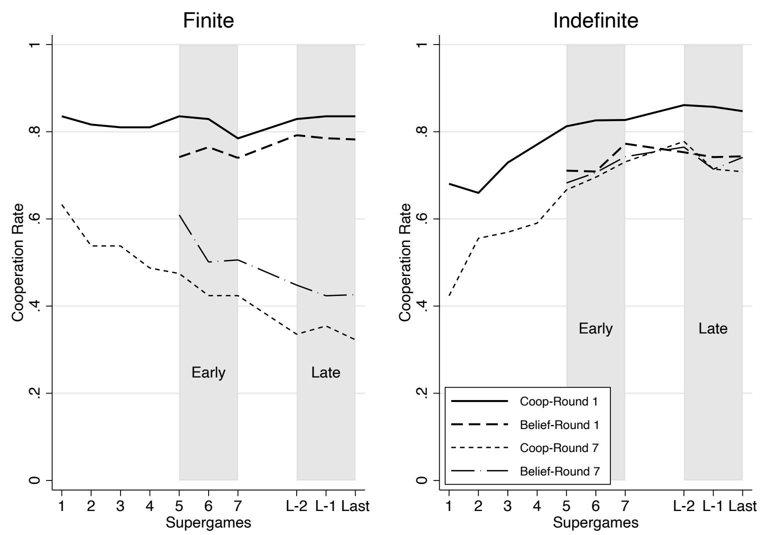


Figure 16: Average Cooperation and Belief

How are beliefs informed by experience?

Although determining exactly how beliefs are formed is not the goal of this study, understanding what allows subjects to predict actions relatively well is of clear interest. One conjecture is that subjects are simply reporting back their observations about others' behavior from previous supergames. Alternatively, subjects may form beliefs relying on introspection alone, or some combination of learning and introspection.[49] The data suggests that although experiences matter in shaping beliefs, they are not the sole determinant. Figure 17 shows the kernel density estimates of the differences between beliefs and the subject-specific experienced frequencies for the fifth (the first with belief elicitation) and last supergames of any given session. The figure reveals that subjects' beliefs differ substantially from the cooperation rates they have experienced. Consider round one where learning from past experiences is easiest (because there is no need to condition on history). In that round, beliefs differ from experienced frequencies by 17 and 16 percentage points, respectively in the first and last supergames (with belief elicitation) of the Finite game and 21 and 20 percentage points in the Indefinite game. This means that in many cases (58 per-

[49]The earlier observation about the Finite game—although behavior is changing in round seven, beliefs track action frequencies closely—already suggests subjects cannot be basing their beliefs only on empirical frequencies.

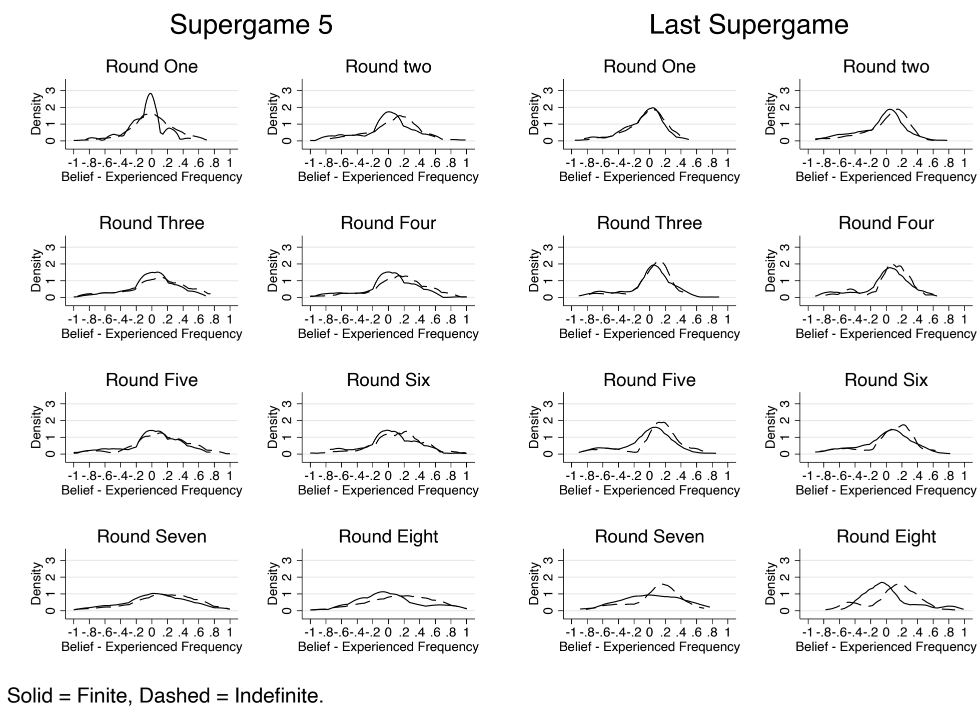


Figure 17: Difference Between Stated Beliefs and Experienced Frequencies of Cooperation by Subject

cent) beliefs are further than plus or minus 20 percentage points of the experienced cooperation rates.

Accuracy of beliefs on the subject-level

These results showing beliefs that are fairly accurate, both averaged over histories and along specific histories, do not speak directly to whether many or few subjects correctly anticipate actions at the individual level. One way to answer this question in a simple but structured way is to look at whether subjects are accurate in at least assessing whether cooperation by their opponent is a relatively likely or unlikely event. Specifically, we denote cooperation (by one's opponent) conditional on a history to be *unlikely* if the empirical frequency of cooperation is less than one third, *likely* if the empirical frequency is more than two thirds, and *uncertain* if the empirical frequency is between these values. Then, we identify the share of observations for which a subject's belief is accurate relative to this categorization; that is, we look at whether the belief lies in the same tercile (unlikely/likely/uncertain) as the observed average cooperation rate. We do so for rounds one and two.

Table 8: Accuracy (numbers are percentages)

Finite

		Early			**Late**		
		Correct	Within		Correct	Within	
		Tercile	10%	5%	Tercile	10%	5%
	Round 1	69	14	8	73	14	7
	Round 2						
	CC	87	60	7	91	60	9
Round 1	CD	63	16	8	67	16	9
Actions	DC	16	11	4	66	7	7
	DD	67	0	0	67	8	8
	Average	71	44	7	83	45	9

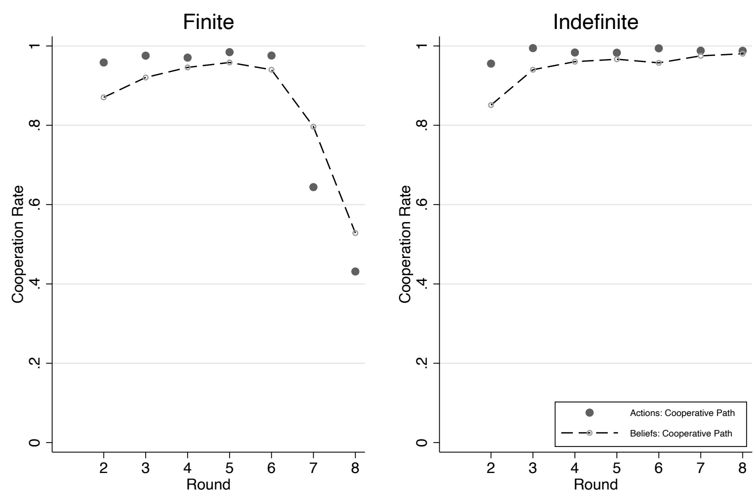
Indefinite

		Early			**Late**		
		Correct	Within		Correct	Within	
		Tercile	10%	5%	Tercile	10%	5%
	Round 1	65	13	7	67	10	5
	Round 2						
	CC	86	52	5	91	66	58
Round 1	CD	35	24	12	29	10	2
Actions	DC	65	6	6	56	17	12
	DD	11	0	0	79	0	0
	Average	73	40	6	80	52	45

Round 1 actions are listed own action first, other's action second: i.e. (a_i, a_j).
Average is weighted by the number of observations.
Note: the number of observations following DD is small, with 2% and 5%, for finite and indefinite respectively, of observations for late supergames.

Table 8 shows that accuracy of beliefs at the individual level, as defined above, is high both for round one (73% in the Finite game, 67% in the Indefinite game) and round two (83% in the Finite game, 80% in the Indefinite game). The accuracy rate is substantially above 33% (the benchmark if beliefs were random) and this is true even in early supergames (above 65% in rounds 1 and 2 for both treatments). However, after one history, accuracy is low: in round two of the Indefinite game along $h^1 = (C, D)$ (cooperation by oneself and defection by the other), beliefs fall in the correct tercile only 29% of the time. Interestingly, the opposite is not true: round-two beliefs along $h^1 = (D, C)$ (defection by oneself and cooperation by the other) fall in the correct tercile 79% of the time. Table 8 also considers more demanding tests of accuracy by reporting the fraction of times the empirical frequencies of cooperation are within ±5 and 10 percentage points of reported beliefs. Beliefs are fairly accurate along some histories (especially the more common ones, e.g., $h^1 = (C, C)$), but less so along other histories that are less common (particularly along $h^1 = (C, D)$ and (D, C) in the Indefinite game).

Beliefs on a cooperative path in early supergames



Cooperative path: joint cooperation up to that round.
Early supergames.

Figure 18: Cooperative Path (First Eight Rounds)

Are beliefs predictive of actions?

We use *round* to make the regressions succinct, but a specification with round indicator variables gives similar estimates.

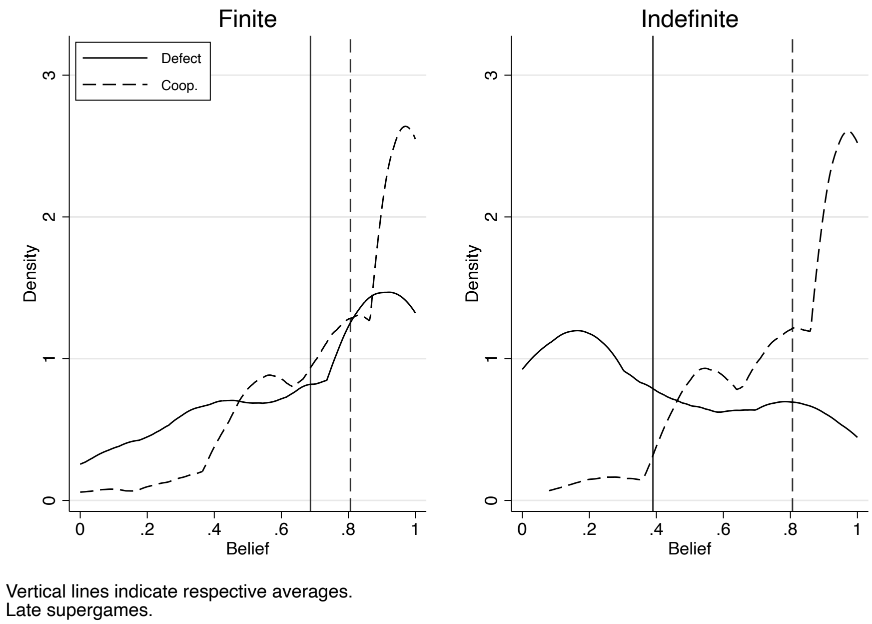
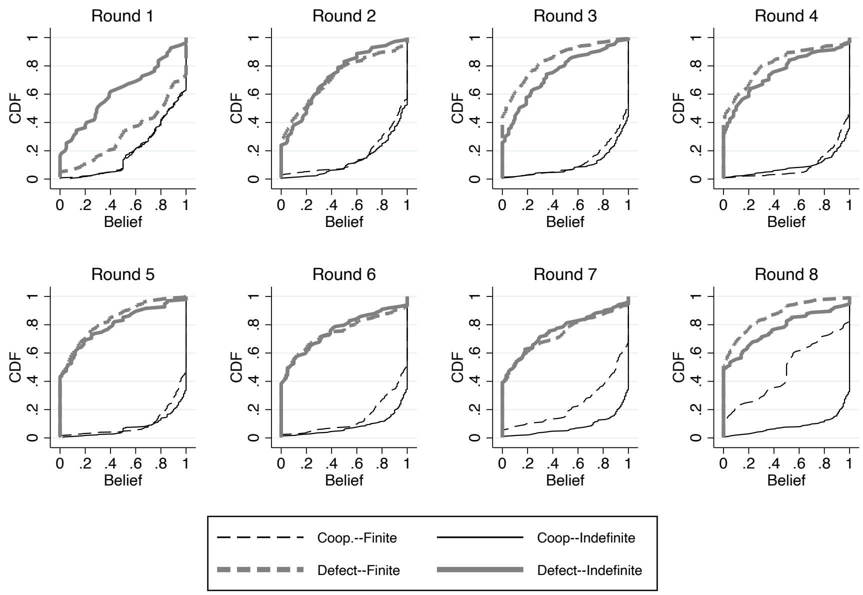


Figure 19: Beliefs of Defectors vs. Cooperators in Round One

Table 9: Correlated Random Effects Probit (Marginal Effects)
Dependent Variable: Cooperation in Round One

	Finite	Indefinite
Belief	0.0938***	0.258***
	(0.0272)	(0.0192)
Other Cooperated in Previous Supergame	-0.0382	0.0274
	(0.0379)	(0.0340)
Supergame	0.00143	0.00798
	(0.00925)	(0.00572)
Length of Previous Supergame		-0.00161
		(0.00121)
Cooperated in Supergame 1	0.413***	0.0493***
	(0.0810)	(0.0164)
Risk Measure	0.00163*	-0.000351
	(0.000848)	(0.000561)
Observations	474	378

Standard errors clustered (at the session level) in parentheses. ***1%, **5%, *10% significance.
All variables refer to behavior in Round 1.
Late supergames.
Risk Measure is equal to the number of boxes collected in the bomb task.



Late supergames.

Figure 20: Beliefs by Action and Treatment: Rounds One through Eight

Table 10: Correlated Random Effects Probit (Marginal Effects)
Dependent Variable: Cooperation

	Finite	Indefinite
Belief	0.462***	0.395***
	(0.0176)	(0.0146)
Round	-0.0336***	-0.00238
	(0.00339)	(0.00282)
Coop. in Round 1, Supergames 1-4	0.244***	0.0805***
	(0.0477)	(0.0244)
Coop. in Last Round, Supergames 1-4	0.126***	0.111***
	(0.0164)	(0.0321)
Risk Measure	-0.0000121	0.000105
	(0.000771)	(0.000633)
Observations	3792	3628

Standard errors clustered (at the session level) in parentheses. ***1%, **5%, *10% significance.
Late supergames.
Risk Measure is equal to the number of boxes collected in the bomb task.

C Additional Details and Analysis on Estimation of Strategies and Beliefs over Strategies

Table 11: Description of Strategies Estimated

Name of Strategy	Code	Description
Always Defect	AD	always play D.
Always Cooperate	AC	always play C.
Grim	GRIM	play C until either player plays D, then play D forever.
Tit-For-Tat	TFT	play C unless partner played D last round.
Suspicious Tit-For-Tat	STFT	play D in the first round, then TFT.
Threshold 8	T8	play Grim until round 8 (last round) then switch to AD.
Threshold 7	T7	play Grim until round 7 then switch to AD.
Threshold 6	T6	play Grim until round 6 then switch to AD.
Threshold 5	T5	play Grim until round 5 then switch to AD.
Threshold 4	T4	play Grim until round 4 then switch to AD.
Threshold 3	T3	play Grim until round 3 then switch to AD.
Threshold 2	T2	play C in round 1 then switch to AD.
Lenient Grim 2	GRIM2	play C until 2 consecutive rounds occur in which either player played D, then play D forever.
Tit-For-2 Tats	TF2T	play C unless partner played D in both of the last rounds.
2Tits-For-Tat	2TFT	play C unless partner played D in either of the last 2 rounds.
Lenient Grim 3	GRIM3	play C until 3 consecutive rounds occur in which either player played D, then play D forever.

Details on the two-step procedure to determine the set of strategies

We use a two-step procedure to determine the set of strategies in our analysis. First we rely on prior evidence to construct a consideration set of 16 strategies. The consideration set includes all strategies that Fudenberg et al. [2012] report have a statistically significant SFEM estimate in at least one indefinitely repeated game with perfect monitoring.[50] Motivated by the results of Embrey et al. [2018], who document the prevalent use of threshold strategies with experience in finitely repeated PD games, we also add to the consideration set all threshold strategies up to T8.[51] Appendix B provides a detailed description of each of these strategies. Results on

[50] Our aim was to be inclusive in the first step of the selection process. In particular, our selection criterion is such that we include all the strategies found to be important in a variety of different papers that have estimated strategies and covered in the meta-study of Dal Bó and Fréchette [2018]. It also means that we do not include strategies that are not observed in direct elicitation studies (Dal Bó and Fréchette [2019] and Romero and Rosokha [2023]).

[51] Thus, the consideration set is AD, AC, Grim, TFT, STFT, Grim2, Grim3, TF2T, 2TFT, and T2–T8. GrimX and TFXT are lenient versions of the corresponding strategy that punish after X consecutive defections by the opponent, 2TFT returns to cooperation only after two consecutive cooperate choices by the opponent.

this consideration set are reported in Online Appendix B. However, because our primary goal is to estimate beliefs over strategies, focusing on such a large set is more costly than is typical with SFEM: having more strategies can make identifying beliefs over different strategies difficult; it can also reduce the number of observations per type in the belief estimation. For these reasons, we use results from the larger consideration set to focus our analysis on the 10 strategies that are most important in terms of choices as well as beliefs. This set consists of AD, AC, Grim, TFT, STFT, Grim2, and TF2T, as well as threshold strategies T8, T7, and T6.[52]

Table 12: Strategy Prevalence and Typing

Finite			Indefinite		
	Share			Share	
Type	SFEM	Typing	Type	SFEM	Typing
T7	0.30	0.35	TFT	0.36	0.59
T8	0.22	0.20	Grim	0.18	0.09
AD	0.12	0.12	Grim2	0.11	0.11
TFT	0.09	0.12	AC	0.11	0.05
T6	0.08	0.08	TF2T	0.10	0.01
Grim	0.08	0.02	AD	0.09	0.10
TF2T	0.04	0.04	STFT	0.04	0.04
STFT	0.03	0.03	T8	0.01	0.01
AC	0.03	0.03	T7	0.00	0.00
Grim2	0.02	0.01	T6	0.00	0.00

Estimation using late supergames.
SFEM estimate for β are 0.94 for both.

[52]From the original set, we eliminate T2–T5, which our estimates indicate are not relevant in the Finite game, as well as 2TFT and Grim3, which are not popular enough in the Indefinite game to generate reliable belief estimates.

Table 13: Estimates for the Finite Game on Late Supergames

Type	Share		Estimated Beliefs - $\tilde{p}$																	
	SFEM	Typing	AD	AC	GRIM	TFT	STFT	T8	T7	T6	T5	T4	T3	T2	GRIM2	TF2T	2TFT	GRIM3	ν	$\tilde{\beta}$
T7	0.30	0.35	0.00	0.00	0.00	0.36	0.00	0.23	0.41	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	1.00
T8	0.22	0.20	0.05	0.00	0.00	0.00	0.00	0.52	0.00	0.00	0.00	0.00	0.00	0.00	0.15	0.15	0.00	0.13	0.04	1.00
AD	0.12	0.12	0.12	0.00	*[0.00]*	0.00	0.13	*[0.00]*	0.00	*[0.00]*	*[0.00]*	*[0.00]*	*[0.00]*	0.00	0.00	0.00	0.75	0.00	0.06	1.00
TFT	0.09	0.12	0.11	0.00	0.00	0.55	0.03	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	1.00
T6	0.08	0.08	0.00	0.00	*[0.00]*	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	1.00
GRIM	0.07	0.02	0.06	0.06	0.11	0.20	0.27	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	*[0.00]*	0.02	0.07	1.00
TF2T	0.03	0.04	0.00	0.17	0.83	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	1.00
GRIM3	0.03	0.03	0.00	0.00	*[0.00]*	0.78	0.00	0.00	0.00	0.00	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	1.00
STFT	0.02	0.02	0.00	0.00	*[0.00]*	0.81	0.00	*[0.00]*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.14	1.00
AC	0.02	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GRIM2	0.01	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T2	0.01	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T5	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T4	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T3	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2TFT	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ALL			0.04	0.01	0.03	0.22	0.04	0.24	0.21	0.00	0.01	0.00	0.00	0.00	0.03	0.04	0.09	0.03		

Estimation on late supergames. SFEM estimate for β is 0.94.
Estimates in *[square brackets]* are not estimated due to collinearity.

Table 14: Estimates for the Indefinite Game on Late Supergames

Type	Share		Estimated Beliefs - $\tilde{p}$																	
	SFEM	Typing	AD	AC	GRIM	TFT	STFT	T8	T7	T6	T5	T4	T3	T2	GRIM2	TF2T	2TFT	GRIM3	ν	$\tilde{\beta}$
TFT	0.34	0.58	0.08	0.12	0.08	0.28	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.26	0.14	0.00	0.01	1.00
GRIM	0.15	0.07	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06	1.00
AC	0.10	0.10	0.00	0.85	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	1.00
AD	0.09	0.10	0.90	0.01	0.07	0.01	0.00	*[0.00]*	*[0.00]*	*[0.00]*	*[0.00]*	*[0.00]*	*[0.00]*	*[0.00]*	0.01	0.00	0.00	0.01	0.04	1.00
TF2T	0.09	0.03	0.00	0.97	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	1.00
GRIM2	0.07	0.02	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.52	0.24	*[0.00]*	0.00	0.05	1.00
GRIM3	0.06	0.02	0.00	0.01	0.24	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.72	0.01	*[0.00]*	0.01	0.01	1.00
2TFT	0.05	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
STFT	0.04	0.04	0.48	0.00	0.15	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.21	0.14	0.00	0.00	0.00	0.07	1.00
T3	0.02	0.03	0.00	0.00	0.16	0.30	0.00	*[0.00]*	*[0.00]*	*[0.00]*	0.00	0.07	0.07	0.00	0.14	0.03	0.23	0.00	0.08	1.00
T8	0.01	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T7	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T6	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T5	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T4	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
T2	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ALL			0.13	0.24	0.22	0.12	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.09	0.11	0.06	0.06		

Estimation on late supergames. SFEM estimate for β is 0.94.
Estimates in *[square brackets]* are not estimated due to collinearity.

Complete Estimation Results for Baseline Treatments (Finite and Indefinite)

Table 15: Estimates for the Finite Game on Late Supergames

	Share		Estimated Beliefs - $\tilde{p}$											
	SFEM	TYPING	AD	AC	GRIM	TFT	STFT	T8	T7	T6	GRIM2	TF2T	ν	$\tilde{\beta}$
T7	0.30	0.35	0.00	0.00	0.18	0.00	0.00	0.39	0.43	0.00	0.00	0.00	0.04	1.00
			(0.02)	(0)	(0.13)	(0.09)	(0)	(0.15)	(0.14)	(0)	(0)	(0.01)		
T8	0.22	0.20	0.09	0.00	0.04	0.01	0.00	0.50	0.00	0.00	0.21	0.15	0.04	1.00
			(0.09)	(0.06)	(0.09)	(0.15)	(0.01)	(0.11)	(0.06)	(0)	(0.11)	(0.11)		
AD	0.12	0.12	0.07	0.00	*[0.00]*	0.00	0.18	*[0.00]*	0.75	*[0.00]*	0.00	0.00	0.06	1.00
			(0.09)	(0.01)		(0.1)	(0.09)		(0.16)		(0.03)	(0.06)		
TFT	0.09	0.12	0.11	0.00	0.00	0.53	0.03	0.33	0.00	0.00	0.00	0.00	0.05	1.00
			(0.08)	(0.05)	(0.07)	(0.22)	(0.04)	(0.12)	(0.03)	(0)	(0.06)	(0.12)		
T6	0.08	0.08	0.00	0.00	*[0.00]*	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.03	1.00
			(0.06)	(0)		(0.17)	(0.02)	(0.17)	(0.29)	(0.09)	(0)	(0.01)		
GRIM	0.08	0.02	0.34	0.10	0.17	0.16	0.00	0.22	0.00	0.00	0.00	0.01	0.07	1.00
			(0.21)	(0.06)	(0.35)	(0.08)	(0.12)	(0.13)	(0.08)	(0.01)	(0.03)	(0.03)		
TF2T	0.04	0.04	0.00	0.14	0.83	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.04	1.00
			(0.08)	(0.13)	(0.35)	(0.09)	(0.02)	(0.14)	(0.13)	(0.02)	(0.17)	(0.06)		
STFT	0.03	0.03	0.00	0.00	*[0.00]*	0.65	0.00	*[0.00]*	0.00	0.00	0.00	0.35	0.11	1.00
			(0.02)	(0.1)		(0.4)	(0.02)		(0.03)	(0.02)	(0.04)	(0.38)		
AC	0.03	0.03	0.04	0.00	0.16	0.30	0.03	0.46	0.00	0.00	0.00	0.00	0.07	1.00
			(0.04)	(0.04)	(0.13)	(0.17)	(0.04)	(0.2)	(0.02)	(0.03)	(0.04)	(0.11)		
GRIM2	0.02	0.01	-	-	-	-	-	-	-	-	-	-	-	-
ALL			0.07	0.01	0.12	0.09	0.03	0.29	0.30	0.00	0.05	0.04		

Estimation on late supergames. SFEM estimate for β is 0.94.
Estimates in *[square brackets]* are not estimated due to collinearity.
Estimates in *(brackets)* show bootstrapped standard deviation.

Table 16: Estimates for the Indefinite Game on Late Supergames

	Share		Estimated Beliefs - $\tilde{p}$											
	SFEM	TYPING	AD	AC	GRIM	TFT	STFT	T8	T7	T6	GRIM2	TF2T	ν	$\tilde{\beta}$
TFT	0.36	0.59	0.08	0.00	0.28	0.25	0.05	0.00	0.00	0.00	0.14	0.19	0.01	1.00
			(0.05)	(0.05)	(0.14)	(0.14)	(0.03)	(0)	(0)	(0)	(0.12)	(0.12)		
GRIM	0.18	0.09	0.00	0.00	0.80	0.13	0.00	0.00	0.00	0.00	0.05	0.02	0.06	1.00
			(0.07)	(0.07)	(0.23)	(0.17)	(0.06)	(0.02)	(0.02)	(0.01)	(0.06)	(0.11)		
GRIM2	0.11	0.11	0.00	0.23	0.22	0.00	0.00	0.00	0.00	0.00	0.31	0.23	0.02	1.00
			(0.03)	(0.13)	(0.16)	(0.06)	(0.02)	(0)	(0)	(0)	(0.2)	(0.14)		
AC	0.11	0.05	0.00	0.80	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.11	1.00
			(0.05)	(0.33)	(0.03)	(0.26)	(0.05)	(0.02)	(0.02)	(0.02)	(0.14)	(0.11)		
TF2T	0.10	0.01	0.00	0.27	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.01	1.00
			(0)	(0.13)	(0.2)	(0.05)	(0)	(0.01)	(0)	(0)	(0.05)	(0.18)		
AD	0.09	0.10	1.00	0.00	0.00	0.00	0.00	*[0.00]*	*[0.00]*	*[0.00]*	0.00	0.00	0.04	1.00
			(0.25)	(0.02)	(0.1)	(0.05)	(0.17)				(0.01)	(0.01)		
STFT	0.04	0.04	0.48	0.00	0.35	0.00	0.00	0.00	0.00	0.00	0.16	0.00	0.08	1.00
			(0.27)	(0.08)	(0.24)	(0.13)	(0.11)	(0.04)	(0.05)	(0.07)	(0.08)	(0.08)		
T8	0.01	0.01	-	-	-	-	-	-	-	-	-	-	-	-
T7	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
T6	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
ALL			0.14	0.14	0.32	0.14	0.02	0.00	0.00	0.00	0.10	0.14		

Estimation on late supergames. SFEM estimate for β is 0.94.
Estimates in *[square brackets]* are not estimated due to collinearity.
Estimates in *(brackets)* show bootstrapped standard deviation.

Table 17: Estimates for the Finite Game on Early Supergames

	Share		Estimated Beliefs - $\tilde{p}$											
	SFEM	TYPING	AD	AC	GRIM	TFT	STFT	T8	T7	T6	GRIM2	TF2T	ν	$\tilde{\beta}$
T8	0.30	0.36	0.01	0.00	0.40	0.00	0.00	0.58	0.00	0.00	0.00	0.00	0.05	1.00
			(0.06)	(0.01)	(0.14)	(0.07)	(0.00)	(0.09)	(0.04)	(0.00)	(0.05)	(0.04)		
T7	0.25	0.20	0.00	0.00	*[0.00]*	0.00	0.00	0.75	0.25	0.00	0.00	0.00	0.03	1.00
			(0.03)	(0.01)		(0.12)	(0.01)	(0.18)	(0.12)	(0.01)	(0.01)	(0.02)		
TFT	0.17	0.15	0.19	0.00	0.50	0.00	0.03	0.26	0.00	0.00	0.02	0.00	0.05	1.00
			(0.08)	(0.02)	(0.21)	(0.16)	(0.05)	(0.15)	(0.03)	(0.01)	(0.05)	(0.08)		
AD	0.12	0.13	0.25	0.00	*[0.00]*	0.21	0.00	0.55	*[0.00]*	*[0.00]*	0.00	0.00	0.11	1.00
			(0.12)	(0.04)		(0.17)	(0.09)	(0.22)			(0.06)	(0.09)		
TF2T	0.05	0.08	0.15	0.00	0.06	0.54	0.15	0.11	0.00	0.00	0.00	0.00	0.05	1.00
			(0.20)	(0.07)	(0.12)	(0.20)	(0.15)	(0.09)	(0.00)	(0.01)	(0.05)	(0.08)		
GRIM2	0.04	0.03	0.00	0.30	0.00	0.43	0.00	0.27	0.00	0.00	0.00	0.00	0.02	1.00
			(0.01)	(0.20)	(0.20)	(0.17)	(0.04)	(0.19)	(0.02)	(0.03)	(0.11)	(0.12)		
STFT	0.03	0.04	0.00	0.00	*[0.00]*	0.42	0.00	0.00	*[0.00]*	*[0.00]*	0.58	0.00	0.15	1.00
			(0.10)	(0.12)		(0.33)	(0.12)	(0.08)			(0.32)	(0.19)		
AC	0.03	0.02	0.03	0.78	*[0.00]*	0.16	0.00	0.03	*[0.00]*	*[0.00]*	0.00	0.00	0.16	0.92
			(0.27)	(0.36)		(0.09)	(0.09)	(0.08)			(0.08)	(0.09)		
GRIM	0.01	0.00	-	-	-	-	-	-	-	-	-	-	-	-
T6	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
ALL			0.07	0.04	0.21	0.09	0.01	0.49	0.06	0.00	0.02	0.00		

Estimation on early supergames. SFEM estimate for β is 0.92.
Estimates in *[square brackets]* are not estimated due to collinearity.
Estimates in *(brackets)* show bootstrapped standard deviation.

Table 18: Estimates for the Indefinite Game on Early Supergames

	Share		Estimated Beliefs - $\tilde{p}$											
	SFEM	TYPING	AD	AC	GRIM	TFT	STFT	T8	T7	T6	GRIM2	TF2T	ν	$\tilde{\beta}$
TFT	0.36	0.60	0.08	0.19	0.40	0.16	0.05	0.00	0.00	0.00	0.00	0.12	0.01	1.00
			(0.03)	(0.09)	(0.12)	(0.10)	(0.03)	(0)	(0)	(0)	(0.04)	(0.06)		
GRIM	0.21	0.09	0.11	0.11	0.45	0.19	0.14	0.00	0.00	0.00	0.00	0.00	0.10	1.00
			(0.11)	(0.2)	(0.21)	(0.13)	(0.09)	(0.04)	(0.09)	(0.07)	(0.12)	(0.15)		
TF2T	0.14	0.10	0.12	0.00	0.25	0.37	0.11	0.00	0.00	0.00	0.06	0.09	0.02	1.00
			(0.09)	(0.05)	(0.12)	(0.11)	(0.08)	(0.01)	(0.01)	(0.01)	(0.06)	(0.08)		
AD	0.13	0.13	0.59	0.03	0.20	0.00	0.14	*[0.00]*	*[0.00]*	*[0.00]*	0.04	0.00	0.05	1.00
			(0.23)	(0.03)	(0.11)	(0.06)	(0.14)				(0.04)	(0.05)		
GRIM2	0.10	0.05	0.42	0.00	0.31	0.00	0.00	0.00	0.00	0.00	0.26	0.00	0.06	1.00
			(0.22)	(0.14)	(0.20)	(0.05)	(0.07)	(0.06)	(0.02)	(0.02)	(0.23)	(0.05)		
AC	0.05	0.02	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	1.00
			(0.01)	(0.09)	(0.02)	(0.44)	(0.01)	(0.01)	(0.01)	(0.01)	(0.05)	(0.13)		
STFT	0.02	0.01	0.00	0.02	0.15	0.16	0.53	*[0.00]*	*[0.00]*	*[0.00]*	0.10	0.03	0.05	1.00
			(0.04)	(0.02)	(0.08)	(0.08)	(0.31)				(0.06)	(0.02)		
T8	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
T7	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
T6	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
ALL			0.19	0.09	0.33	0.20	0.09	0.00	0.00	0.00	0.04	0.06		

Estimation on early supergames. SFEM estimate for β is 0.94.
Estimates in *[square brackets]* are not estimated due to collinearity.
Estimates in *(brackets)* show bootstrapped standard deviation.

Table 19: Best Response Analysis

Finite						Indefinite					
	Share		Best Response				Share		Best Response		
Type	SFEM	Typing	BRS	R_s	R_o	Type	SFEM	Typing	BRS	R_s	R_o
T7	0.30	0.35	T7	1	0.97	TFT	0.34	0.58	TF2T/GRIM2	0.99	0.94
T8	0.22	0.20	T7	0.90	0.90	GRIM	0.15	0.07	AC/GRIM/TFT/TF2T/GRIM2	1	1
AD	0.12	0.12	T8	0.29	0.64	AC	0.10	0.10	AC	0.61	0.74
TFT	0.09	0.12	T8	0.86	0.80	AD	0.09	0.10	AD	1	0.77
T6	0.08	0.08	T6	1	1	TF2T	0.09	0.03	AC/GRIM/TFT/TF2T/GRIM2	1	0.98
GRIM	0.07	0.02	T7	0.85	0.82	GRIM2	0.07	0.02	AC/GRIM/TFT/TF2T/GRIM2	1	1
Other	0.12	0.11	T6			Other	0.16	0.10	TFT		
All			T7			All			TFT		

Estimation on late supergames out of 16 strategies: AD, AC, Grim, Grim2, Grim3, TFT, TF2T, 2TFT, STFT, T2-T8.

Rows represent top 6 played strategies; BRS: Best Response strategy given beliefs.

In Finite games the best response strategy to the actual distribution (SFEM) is T6; in Indefinite games it is GRIM2.

R_s: Expected payoff from strategy/Best response payoff given beliefs.

R_o: Expected payoff from strategy/Best response payoff given actual distribution (SFEM).

C.1 Contrasting Early and Late Supergames

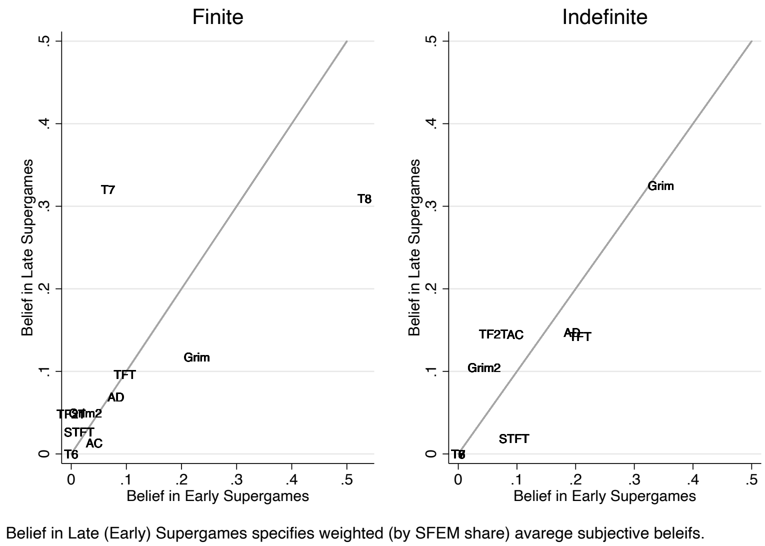
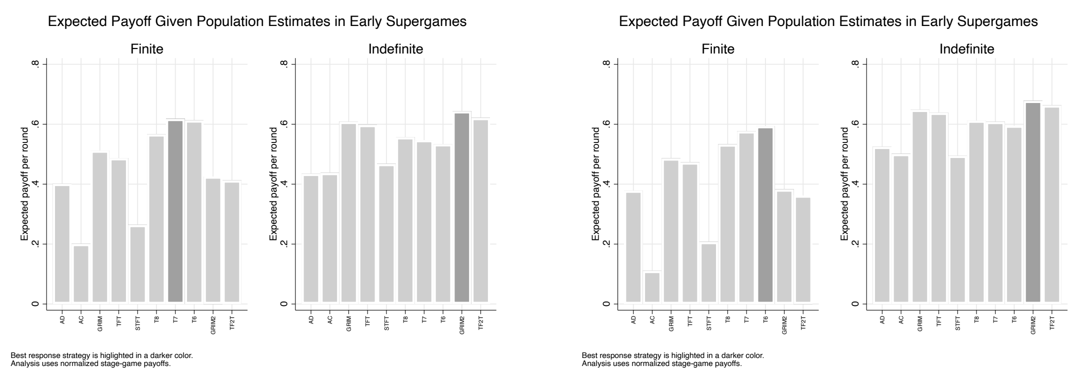


Figure 21: Change in Beliefs from Early to Late Supergames



(a) (b)

Figure 22: Normalized Expected Payoff by Type Given Strategy Distribution in Early and Late Supergames

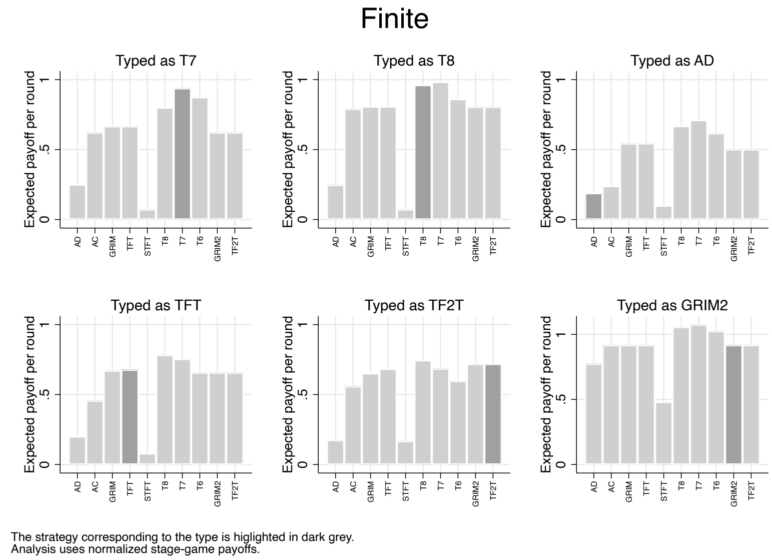


Figure 23: Best Response for Top 6 Types in the Finite Game in Early Supergames

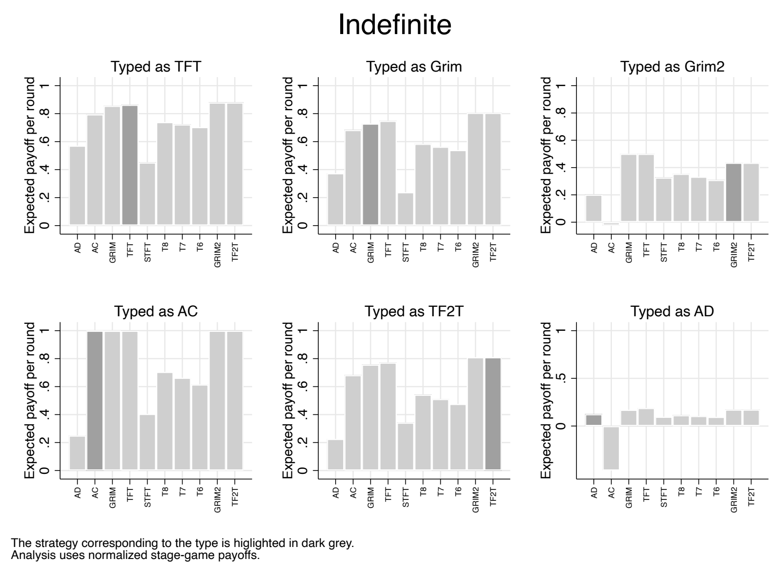


Figure 24: Best Response for Top 6 Types in the Indefinite Game in Early Supergames

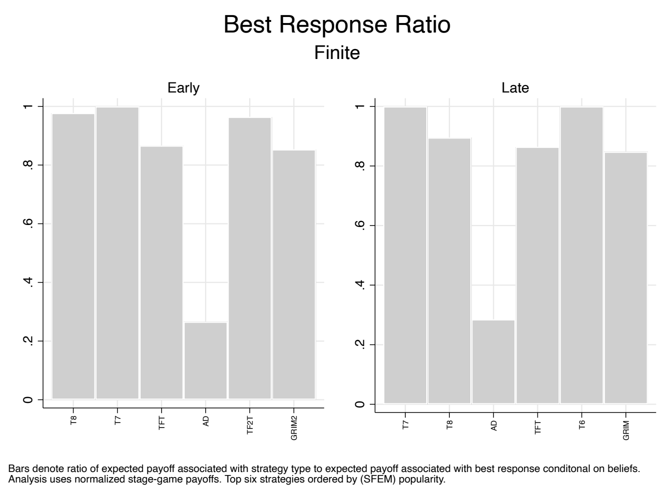


Figure 25: Best Response Ratio for Top 6 Types in the Finite Game

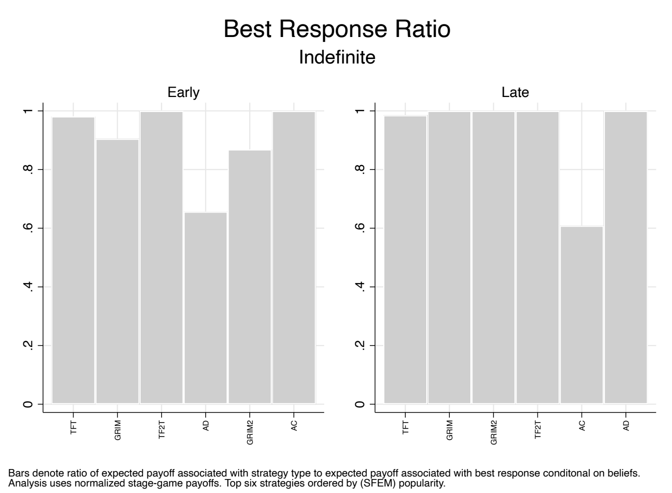
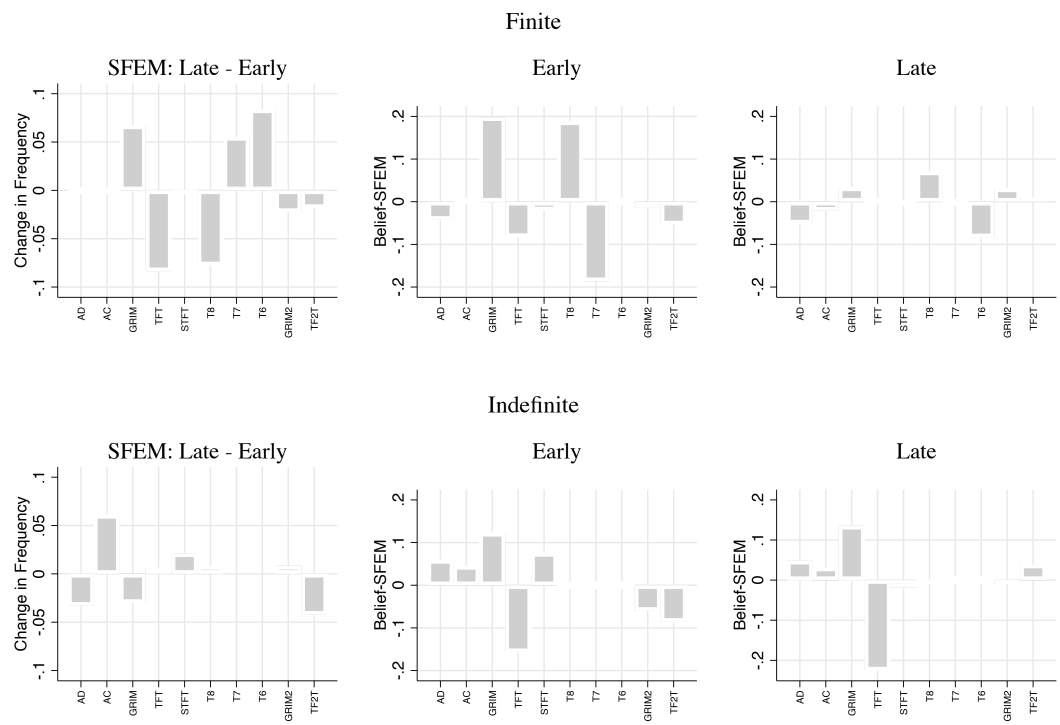


Figure 26: Best Response Ratio for Top 6 Types in the Indefinite Game



Y-axis for Early and Late panels denotes distance between weigted average subjective beliefs and SFEM values.

Figure 27: Strategy Changes and Belief Accuracy

The accuracy of beliefs over strategies can be studied more directly without relying on the cooperativeness order. In Figure 28 of Online Appendix B, we compute, for each type, the Euclidean distance between beliefs and the estimated frequency of strategies. To study whether beliefs become more accurate with experience, we also look at how this distance changes from early to late supergames. We find that, in aggregate, beliefs are becoming more accurate with experience in the Finite game, whereas accuracy changes little in the Indefinite game. In both cases, the most popular strategy types (T7 in Finite and TFT in Indefinite) have the most accurate beliefs in late supergames.[53]

[53]In the Finite game, early beliefs overestimate the likelihood of T8 and underestimate the likelihood of T7. Both of these errors are reduced (or eliminated) with experience. For the Indefinite game, early beliefs overestimate the likelihood of Grim and underestimate the likelihood of TFT; however, these errors (which are less costly than those observed in the Finite game) are not corrected with experience.

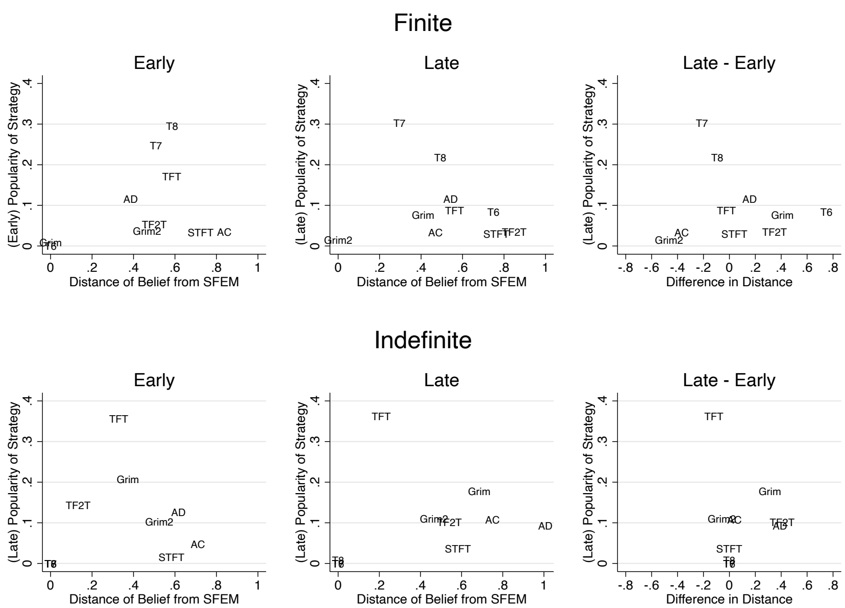


Figure 28: Change in Accuracy

Additionally, in Figures 21-27 we study learning effects more generally. We document in detail how the distribution of strategies, types, and beliefs for each type change from early to late supergames. We summarize the key observations from these results here. While behavior stabilizes quickly in the Indefinite game—with little change in distribution of strategies, types and beliefs observed from early to late supergames—there is clear evidence of learning in the Finite game. Most significantly, there is a shift towards less cooperative strategies: popularity of T8 declines while the popularity of T7 and T6 increase. The observed shift in strategies is anticipated by beliefs. These results suggest, in the Finite game, subjects to be updating their beliefs about the cooperativeness of their counterpart throughout the session and adjusting their strategy choices in response to these changing beliefs.

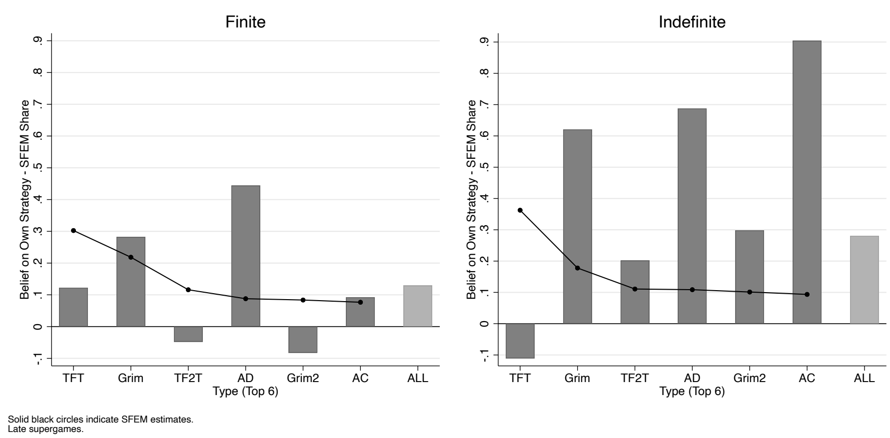


Figure 29: Overestimation in Beliefs of the Prevalence of One's Own Startegy

D Model of Heterogeneous Beliefs about the Sophistication of Others

This section formally describes the level-k model adapted to the supergame environment presented in Section 6.

Let σ_k and $\tilde{p}_k$ denote the supergame strategy and supergame belief, respectively, of a level-k player. Let also ζ_k be the proportion of level-k players in the population. For simplicity, we assume that there are three levels of sophistication so that $\zeta_0+\zeta_1+\zeta_2=1$.[54] Suppose that RD is a (stationary) supergame strategy that plays C with probability q and D with probability $1-q$ in each round after every history for some $q \in [0,1]$.[55] We assume that a level-0 player has a belief $\tilde{p}_0$ that places probability one on RD.[56] For $k \geq 1$, a level-k player has a belief $\tilde{p}_k$ which has support over strategies played by players whose levels are at or below k.[57]

Level-1 and level-2 players best respond to their beliefs: $\sigma_1 \in \text{BR}(\tilde{p}_1)$ and $\sigma_2 \in \text{BR}(\tilde{p}_2)$. On the other hand, we assume that the strategy σ_0 of level-0 players is such that for $\omega \in (0,1)$, σ_0 randomizes between Grim and RD as follows:[58]

$$\sigma_0 = \omega \cdot \text{Grim} + (1-\omega) \cdot \text{RD}.$$

Recall that $\delta = \frac{7}{8}$, $g = 1$ and $\ell = \frac{17}{12}$ in our parametrization.

D.1 Indefinite Games

We begin with the following observation:

[54] Increasing the number of sophistication levels leads essentially to the same conclusion in the Indefinite game but advances unraveling in the Finite game.

[55] RD with $q=0$ hence equals AD.

[56] Although the level-k theory does not usually specify the belief of level-0 players, it is needed here for the computation of the average belief in the population.

[57] It is standard in the level-k theory to assume that a level-k player believes that only those types below level k are present so that $\tilde{p}_k(\sigma_k)=0$. We allow the possibility that $\tilde{p}_k(\sigma_k) > 0$ for $k \geq 1$ to align the theory with the experimental finding that the subjects tend to place a positive belief weight on their own strategy.

[58] While the level-k theory usually assumes that the level-0 strategy is a random action choice, it is necessary to include a conditionally cooperative strategy such as Grim as a component of σ_0 since otherwise AD would become the unique best response to σ_0.

Observation 1 If $\sigma_1 = \sigma_2 = \text{Grim}$, then for $k = 1, 2$, $\sigma_k \in \text{BR}(\tilde{p}_k)$ if and only if

$$\frac{\tilde{p}_k(\text{Grim})}{\tilde{p}_k(\text{RD})}\Big[1-(1-\delta)(1+g)\Big] \geq u_i(\text{AD},\text{RD}) - u_i(\text{Grim},\text{RD}). \tag{1}$$

This condition holds if $\frac{\tilde{p}_1(\text{Grim})}{\tilde{p}_1(\text{RD})}$ is sufficiently large.

Consider a level-k player with belief $\tilde{p}_k$ for $k = 1, 2$. Since $\sigma_1 = \sigma_2 = \text{Grim}$ by assumption, $\tilde{p}_k$ places positive weight only on RD and Grim. Hence, after any history along which either player plays D, playing AD is optimal against $\tilde{p}_k$. In round 1, on the other hand, playing Grim against $\tilde{p}_k$ yields

$$u_i(\text{Grim}, \tilde{p}_k) = \tilde{p}_k(\text{Grim}) + \tilde{p}_k(\text{RD})\, u_i(\text{Grim},\text{RD}).$$

On the other hand, a one-step deviation to D in round 1 yields

$$\tilde{p}_k(\text{Grim})(1-\delta)(1+g) + \tilde{p}_k(\text{RD})\, u_i(\text{AD},\text{RD}).$$

It follows that Grim is a best response against $\tilde{p}_k$ if

$$\begin{aligned}&\tilde{p}_k(\text{Grim}) + \tilde{p}_k(\text{RD})\, u_i(\text{Grim},\text{RD})\\ &\quad\geq \tilde{p}_k(\text{Grim})(1-\delta)(1+g) + \tilde{p}_k(\text{RD})\, u_i(\text{AD},\text{RD}),\end{aligned}$$

which is equivalent to (1). Since $1-(1-\delta)(1+g) > 0$ holds when $\delta = \frac{7}{8} > \frac{1}{2} = \frac{g}{1+g}$, (1) holds when $\frac{\tilde{p}_k(\text{Grim})}{\tilde{p}_k(\text{RD})}$ is sufficiently large.

When (1) holds, hence, $\sigma_1 = \sigma_2 = \text{Grim}$ is consistent with subjective rationality. It follows that the proportion of strategies in the population is given by

$$\big(\omega\zeta_0 + \zeta_1 + \zeta_2\big)\cdot \text{Grim} + (1-\omega)\zeta_0 \cdot \text{RD}. \tag{2}$$

Denote by $\tilde{p}_k(h^t)$ level-k's continuation belief over strategies at history h^t, and let h^t_* be the t-length cooperative history that consists exclusively of (C, C)'s:

$$h^t_* = \big(\underbrace{(C,C),\ldots,(C,C)}_{t \text{ rounds}}\big).$$

A level-k player's continuation belief at h^{t-1}_* in round t is given by

$$\begin{aligned}\tilde{p}_k(h^{t-1}_*)(\text{Grim}) &= \frac{\tilde{p}_k(\text{Grim})}{\tilde{p}_k(\text{Grim}) + \tilde{p}_k(\text{RD})\, q^{t-1}},\\ \tilde{p}_k(h^{t-1}_*)(\text{RD}) &= \frac{\tilde{p}_k(\text{RD})\, q^{t-1}}{\tilde{p}_k(\text{Grim}) + \tilde{p}_k(\text{RD})\, q^{t-1}},\end{aligned}$$

and $\tilde{p}_k(h^{t-1}) = \text{RD}$ if $h^{t-1} \neq h_*^{t-1}$.[59] Suppose that two players from the population with the proportions of RD and Grim as in (2) are randomly matched. We can compute the ex ante mean of the round belief in round t (belief wight placed on the other player's choice of C in round t) as:

$$\begin{aligned}
\bar{\mu}^t = \zeta_0 q & \\
+ \zeta_1 \Big[& \big\{\zeta_0\omega + \zeta_0(1-\omega)q^{t-1} + \zeta_1 + \zeta_2\big\} \\
& \times \big\{\tilde{p}_1(h_*^{t-1})(\text{Grim}) + \tilde{p}_1(h_*^{t-1})(\text{RD})\, q\big\} \\
& \qquad + \zeta_0(1-\omega)(1-q^{t-1})q\Big] \\
+ \zeta_2 \Big[& \big\{\zeta_0\omega + \zeta_0(1-\omega)q^{t-1} + \zeta_1 + \zeta_2\big\} \\
& \times \big\{\tilde{p}_2(h_*^{t-1})(\text{Grim}) + \tilde{p}_2(h_*^{t-1})(\text{RD})\, q\big\} \\
& \qquad + \zeta_0(1-\omega)(1-q^{t-1})q\Big].
\end{aligned}$$

On the other hand, the ex ante mean of the cooperation rates in round t are given by

$$\bar{x}^t = \zeta_0(1-\omega)\, q + (\zeta_0\omega + \zeta_1 + \zeta_2)\big\{\zeta_0\omega + \zeta_0(1-\omega)q^{t-1} + \zeta_1 + \zeta_2\big\}.$$

D.2 Finite Games

We suppose that $\sigma_1 = \text{T8}$ and $\sigma_2 = \text{T7}$, and identify conditions which ensure that these strategies are indeed subjectively rational. By assumption, $\tilde{p}_1(\text{T7}) = 0$. Suppose first that $t \leq 7$. For $k = 1, 2$, the continuation belief of a level-k player at

[59] For any h^{t-1} that occurs only after one's own deviation, Bayes rule would imply a different specification of $\tilde{p}_k(h^{t-1})$. For example, after h^1 which involves the own choice of D and the other player's choice of C, the above specifies $\tilde{p}_k(h^1) = \text{RD}$. However, application of Bayes rule would suggest that $\tilde{p}_k(h^1) = \tilde{p}_k(h_*^1)$. This however is immaterial in the subsequent analysis.

history h_*^{t-1} in round t with prior belief $\tilde{p}_k$ is given by

$$\tilde{p}_k(h_*^{t-1})(\text{Grim}) = \frac{\tilde{p}_k(\text{Grim})}{\tilde{p}_k(\text{Grim}) + \tilde{p}_k(\text{T8}) + \tilde{p}_k(\text{T7}) + \tilde{p}_k(\text{RD})\, q^{t-1}},$$
$$\tilde{p}_k(h_*^{t-1})(\text{RD}) = \frac{\tilde{p}_k(\text{RD})\, q^{t-1}}{\tilde{p}_k(\text{Grim}) + \tilde{p}_k(\text{T8}) + \tilde{p}_k(\text{T7}) + \tilde{p}_k(\text{RD})\, q^{t-1}},$$
$$\tilde{p}_k(h_*^{t-1})(\text{T8}) = \frac{\tilde{p}_k(\text{T8})}{\tilde{p}_k(\text{Grim}) + \tilde{p}_k(\text{T8}) + \tilde{p}_k(\text{T7}) + \tilde{p}_k(\text{RD})\, q^{t-1}},$$
$$\tilde{p}_k(h_*^{t-1})(\text{T7}) = \frac{\tilde{p}_k(\text{T7})}{\tilde{p}_k(\text{Grim}) + \tilde{p}_k(\text{T8}) + \tilde{p}_k(\text{T7}) + \tilde{p}_k(\text{RD})\, q^{t-1}}.$$

On the other hand, the continuation belief of a level-1 player at history h_*^7 in the last round $t = 8$ is given by[60]

$$\tilde{p}_1(h_*^7)(\text{Grim}) = \frac{\tilde{p}_1(\text{Grim})}{\tilde{p}_1(\text{Grim}) + \tilde{p}_1(\text{T8}) + \tilde{p}_1(\text{RD})\, q^7},$$
$$\tilde{p}_1(h_*^7)(\text{RD}) = \frac{\tilde{p}_1(\text{RD})\, q^7}{\tilde{p}_1(\text{Grim}) + \tilde{p}_1(\text{T8}) + \tilde{p}_1(\text{RD})\, q^7},$$
$$\tilde{p}_1(h_*^7)(\text{T8}) = \frac{\tilde{p}_1(\text{T8})}{\tilde{p}_1(\text{Grim}) + \tilde{p}_1(\text{T8}) + \tilde{p}_1(\text{RD})\, q^7}.$$

For a level-2 player who plays T7, the history h_*^7 does not arise on the path of play. Instead, the relevant histories are given by $(h_*^6, (D, C))$ and $(h_*^6, (D, D))$: $(h_*^6, (D, C))$ is the history where (D, C) (own choice of D and the other's choice of C) in round 7 follows h_*^6, and $(h_*^6, (D, D))$ is the history where (D, D) in round 7 follows h_*^6. Note that at these histories, level-2 expects the other player to play C with positive probability in round 8 only when the other player plays RD. The continuation beliefs of a level-2 player at these histories in round 8 that the other player plays RD are given by

$$\tilde{p}_2(h_*^6, (D, C))(\text{RD}) = \frac{\tilde{p}_2(\text{RD})\, q^7}{\tilde{p}_2(\text{Grim}) + \tilde{p}_2(\text{T8}) + \tilde{p}_2(\text{RD})\, q^7},$$
$$\tilde{p}_2(h_*^6, (D, D))(\text{RD}) = \frac{\tilde{p}_2(\text{RD})\, q^6(1-q)}{\tilde{p}_2(\text{T7}) + \tilde{p}_2(\text{RD})\, q^6(1-q)}.$$

[60]See Footnote 59.

The ex ante mean of the round belief in round t for $t = 1, \ldots, 6$ is given by

$$\begin{aligned}
\bar{\mu}^t = {} & \zeta_0 q \\
& + \zeta_1 \Big[\big\{ \zeta_0 \omega + \zeta_0 (1-\omega) q^{t-1} + \zeta_1 + \zeta_2 \big\} \\
& \quad \times \big\{ \tilde{p}_1(h_*^{t-1})(\mathrm{Grim}) + \tilde{p}_1(h_*^{t-1})(\mathrm{T8}) + \tilde{p}_1(h_*^{t-1})(\mathrm{RD})\, q \big\} \\
& \qquad + \zeta_0 (1-\omega)(1-q^{t-1})\, q \Big] \\
& + \zeta_2 \Big[\big\{ \zeta_0 \omega + \zeta_0 (1-\omega) q^{t-1} + \zeta_1 + \zeta_2 \big\} \\
& \quad \times \big\{ \tilde{p}_2(h_*^{t-1})(\mathrm{Grim}) + \tilde{p}_2(h_*^{t-1})(\mathrm{T8}) + \tilde{p}_2(h_*^{t-1})(\mathrm{T7}) + \tilde{p}_2(h_*^{t-1})(\mathrm{RD})\, q \big\} \\
& \qquad + \zeta_0 (1-\omega)(1-q^{t-1})\, q \Big].
\end{aligned}$$

Likewise, the ex ante mean of the round belief in round 7 is given by

$$\begin{aligned}
\bar{\mu}^7 = {} & \zeta_0 q \\
& + \zeta_1 \Big[\big\{ \zeta_0 \omega + \zeta_0 (1-\omega) q^{6} + \zeta_1 + \zeta_2 \big\} \\
& \quad \times \big\{ \tilde{p}_1(h_*^{6})(\mathrm{Grim}) + \tilde{p}_1(h_*^{6})(\mathrm{T8}) + \tilde{p}_1(h_*^{6})(\mathrm{RD})\, q \big\} \\
& \qquad + \zeta_0 (1-\omega)(1-q^{6})\, q \Big] \\
& + \zeta_2 \Big[\big\{ \zeta_0 \omega + \zeta_0 (1-\omega) q^{6} + \zeta_1 + \zeta_2 \big\} \\
& \quad \times \big\{ \tilde{p}_2(h_*^{6})(\mathrm{Grim}) + \tilde{p}_2(h_*^{6})(\mathrm{T8}) + \tilde{p}_2(h_*^{6})(\mathrm{RD})\, q \big\} \\
& \qquad + \zeta_0 (1-\omega)(1-q^{6})\, q \Big],
\end{aligned}$$

and that in round 8 is given by

$$\begin{aligned}
\bar{\mu}^8 = {} & \zeta_0 q \\
& + \zeta_1 \Big[\big\{ \zeta_0 \omega + \zeta_0 (1-\omega) q^{7} + \zeta_1 \big\} \big\{ \tilde{p}_1(h_*^{7})(\mathrm{Grim}) + \tilde{p}_1(h_*^{7})(\mathrm{RD})\, q \big\} \\
& \quad + \big\{ \zeta_0 (1-\omega)(1-q^{7}) + \zeta_2 \big\}\, q \Big] \\
& + \zeta_2 \Big[\big\{ \zeta_0 \omega + \zeta_0 (1-\omega) q^{7} + \zeta_1 \big\}\, \tilde{p}_2(h_*^{6}, (D, C))(\mathrm{RD}) \\
& \quad + \big\{ \zeta_0 (1-\omega) q^{6} (1-q) + \zeta_2 \big\} \tilde{p}_2(h_*^{6}, (D, D))(\mathrm{RD}) \\
& \quad + \zeta_0 (1-\omega)(1-q^{6}) \Big]\, q.
\end{aligned}$$

Observation 2 For a level-1 player, T8 $\in$ BR$(\tilde{p}_1)$ if for $t = 1, \dots, 7$,

$$\begin{aligned} u_i(\text{T8}, \tilde{p}_1 \mid h_*^{t-1}) \geq &\Big[\tilde{p}_1(h_*^{t-1})(\text{Grim}) + \tilde{p}_1(h_*^{t-1})(\text{T8})\Big]\,(1+g) \\ &+ \tilde{p}_1(h_*^{t-1})(\text{RD}) \cdot (9-t)q(1+g). \end{aligned} \tag{3}$$

These conditions hold if $\frac{\tilde{p}_1(\text{T8})+\tilde{p}_1(\text{RD})}{\tilde{p}_1(\text{Grim})}$ is sufficiently small.

It is clear that playing D as specified by T8 is a best response against $\tilde{p}_1$ in round 8 after h_*^7. In round $t \leq 7$ after h_*^{t-1}, a one-step deviation to D yields

$$(1+g)\,\Big[\tilde{p}_1(h_*^{t-1})(\text{Grim}) + \tilde{p}_1(h_*^{t-1})(\text{T8})\Big] + (9-t)q(1+g)\,\tilde{p}_1(h_*^{t-1})(\text{RD}).$$

Hence, no such deviation is profitable if (3) holds. On the other hand, playing T8 against $\tilde{p}_1$ yields

$$\begin{aligned} &u_i(\text{T8}, \tilde{p}_1 \mid h_*^{t-1}) \\ &= (9-t+g)\,\tilde{p}_1(h_*^{t-1})(\text{Grim}) + (8-t)\,\tilde{p}_1(h_*^{t-1})(\text{T8}) \\ &+ \Big[q\Big\{1 + u_i(\text{T8}, \text{RD} \mid h_*^{t})\Big\} + (1-q)\Big\{-\ell + (8-t)q(1+g)\Big\}\Big]\,\tilde{p}_1(h_*^{t-1})(\text{RD}). \end{aligned}$$

We can show by induction that $u_i(\text{T8}, \text{RD} \mid h_*^{t}) \geq -\ell$. It hence follows that

$$\begin{aligned} &u_i(\text{T8}, \tilde{p}_1 \mid h_*^{t-1}) \\ &\geq (9-t+g)\,\tilde{p}_1(h_*^{t-1})(\text{Grim}) + (8-t)\,\tilde{p}_1(h_*^{t-1})(\text{T8}) \\ &+ \Big[q - \ell + (1-q)(8-t)q(1+g)\Big]\,\tilde{p}_1(h_*^{t-1})(\text{RD}). \end{aligned}$$

Hence, (3) is implied if

$$\begin{aligned} &(8-t)\,\tilde{p}_1(h_*^{t-1})(\text{Grim}) + (7-t-g)\,\tilde{p}_1(h_*^{t-1})(\text{T8}) \\ &+ \Big[q - \ell + (1-q)(8-t)q(1+g) - (9-t)q(1+g)\Big]\,\tilde{p}_1(h_*^{t-1})(\text{RD}) \\ &= (8-t)\,\tilde{p}_1(h_*^{t-1})(\text{Grim}) + (7-t-g)\,\tilde{p}_1(h_*^{t-1})(\text{T8}) \\ &+ \Big[-\ell - q^2(8-t)(1+g) - qg\Big]\,\tilde{p}_1(h_*^{t-1})(\text{RD}) \geq 0. \end{aligned}$$

Since $\frac{\tilde{p}_1(h_*^{t-1})(\text{T8})}{\tilde{p}_1(h_*^{t-1})(\text{Grim})} = \frac{\tilde{p}_1(\text{T8})}{\tilde{p}_1(\text{Grim})}$ and $\frac{\tilde{p}_1(h_*^{t-1})(\text{RD})}{\tilde{p}_1(h_*^{t-1})(\text{Grim})} \leq \frac{\tilde{p}_1(\text{RD})}{\tilde{p}_1(\text{Grim})}$, this inequality holds if $\frac{\tilde{p}_1(\text{T8})+\tilde{p}_1(\text{RD})}{\tilde{p}_1(\text{Grim})}$ is sufficiently small.

Observation 3 For a level-2 player, $\mathrm{T7} \in \mathrm{BR}(\tilde{p}_2)$ if

$$\begin{aligned} u_i(\mathrm{T7}, \tilde{p}_2 \mid h_*^{t-1}) \geq (1+g) \Big[& \tilde{p}_2(h_*^{t-1})(\mathrm{Grim}) + \tilde{p}_2(h_*^{t-1})(\mathrm{T8}) + \tilde{p}_2(h_*^{t-1})(\mathrm{T7}) \Big] \\ & + (9-t) q (1+g) \, \tilde{p}_2(h_*^{t-1})(\mathrm{RD}) \\ & \text{for } t = 1, \ldots, 6, \end{aligned} \tag{4}$$

and

$$\begin{aligned} u_i(\mathrm{T7}, \tilde{p}_2 \mid h_*^6) \geq (2+g) \, & \tilde{p}_2(h_*^6)(\mathrm{Grim}) + \tilde{p}_2(h_*^6)(\mathrm{T8}) - \ell \, \tilde{p}_2(h_*^6)(\mathrm{T7}) \\ & + \Big[q + (1-q)(-\ell) + q(1+g) \Big] \, \tilde{p}_2(h_*^6)(\mathrm{RD}). \end{aligned} \tag{5}$$

These conditions hold when $\frac{\tilde{p}_2(\mathrm{Grim})}{\tilde{p}_2(\mathrm{T7}) + \tilde{p}_2(\mathrm{T8})}$ and $\frac{\tilde{p}_2(\mathrm{T7}) + \tilde{p}_2(\mathrm{RD})}{\tilde{p}_2(\mathrm{Grim}) + \tilde{p}_2(\mathrm{T8})}$ are sufficiently small.

In round 7, if the history up to round 6 equals h_*^6, a one-step deviation to C at h_*^6 yields

$$\begin{aligned} & \{1 + (1+g)\} \, \tilde{p}_2(h_*^6)(\mathrm{Grim}) + \tilde{p}_2(h_*^6)(\mathrm{T8}) + (-\ell) \, \tilde{p}_2(h_*^6)(\mathrm{T7}) \\ & + \Big[q + (1-q)(-\ell) + q(1+g) \Big] \, \tilde{p}_2(h_*^6)(\mathrm{RD}). \end{aligned}$$

Hence, no such deviation is profitable if (5) holds. On the other hand, playing T7 at h_*^6 yields

$$\begin{aligned} u_i(\mathrm{T7}, \tilde{p}_2 \mid h_*^6) = (1+g) \, & \Big[\tilde{p}_2(h_*^6)(\mathrm{Grim}) + \tilde{p}_2(h_*^6)(\mathrm{T8}) \Big] \\ & + 2q(1+g) \, \tilde{p}_2(h_*^6)(\mathrm{RD}). \end{aligned}$$

After simplification, we see that (5) holds if and only if

$$\begin{aligned} & - \tilde{p}_2(h_*^6)(\mathrm{Grim}) + g \, \tilde{p}_2(h_*^6)(\mathrm{T8}) + \ell \, \tilde{p}_2(h_*^6)(\mathrm{T7}) \\ & + \Big[qg + (1-q)\ell \Big] \, \tilde{p}_2(h_*^6)(\mathrm{RD}) \geq 0. \end{aligned} \tag{6}$$

Since $\frac{\tilde{p}_2(h_*^6)(\mathrm{Grim})}{\tilde{p}_2(h_*^6)(\mathrm{T7}) + \tilde{p}_2(h_*^6)(\mathrm{T8}) + \tilde{p}_2(h_*^6)(\mathrm{RD})} < \frac{\tilde{p}_2(\mathrm{Grim})}{\tilde{p}_2(\mathrm{T7}) + \tilde{p}_2(\mathrm{T8})}$, it follows that (6) holds if $\frac{\tilde{p}_2(\mathrm{Grim})}{\tilde{p}_2(\mathrm{T7}) + \tilde{p}_2(\mathrm{T8})}$ is sufficiently small.

In round $t \leq 6$, if the history up to round $t-1$ is h_*^{t-1}, then a one-step deviation to D at h_*^{t-1} yields

$$\begin{aligned} & (1+g) \Big[\tilde{p}_2(h_*^{t-1})(\mathrm{Grim}) + \tilde{p}_2(h_*^{t-1})(\mathrm{T8}) + \tilde{p}_2(h_*^{t-1})(\mathrm{T7}) \Big] \\ & + (9-t) q (1+g) \, \tilde{p}_2(h_*^{t-1})(\mathrm{RD}). \end{aligned}$$

It follows that no such deviation at h_*^{t-1} ($t \le 6$) is profitable if (4) holds. On the other hand, playing T7 against $\tilde{p}_2$ at h_*^{t-1} yields

$$
\begin{aligned}
& u_i(T7, \tilde{p}_2 \mid h_*^{t-1}) \\
& = (8 - t + g)\left[\tilde{p}_2(h_*^{t-1})(\text{Grim}) + \tilde{p}_2(h_*^{t-1})(\text{T8})\right] + (7 - t)\,\tilde{p}_2(h_*^{t-1})(\text{T8}) \\
& + \left[q\left\{1 + u_i(T7, R \mid h_*^t)\right\} + (1 - q)\left\{-\ell + (8 - t)q(1 + g)\right\}\right]\tilde{p}_2(h_*^{t-1})(\text{RD}).
\end{aligned}
$$

It follows that (4) holds if and only if

$$
\begin{aligned}
& (7 - t)\left[\tilde{p}_2(h_*^{t-1})(\text{Grim}) + \tilde{p}_2(h_*^{t-1})(\text{T8})\right] + (6 - t - g)\,\tilde{p}_2(h_*^{t-1})(\text{T7}) \\
& + \Big[q\left\{1 + u_i(T7, R \mid h_*^t)\right\} + (1 - q)\left\{-\ell + (8 - t)q(1 + g)\right\} \\
& \qquad - (9 - t)q(1 + g)\Big]\,\tilde{p}_2(h_*^{t-1})(\text{RD}) \ge 0.
\end{aligned}
\tag{7}
$$

Since $\frac{\tilde{p}_2(h_*^{t-1})(\text{T7})}{\tilde{p}_2(h_*^{t-1})(\text{Grim}) + \tilde{p}_2(h_*^{t-1})(\text{T8})} = \frac{\tilde{p}_2(\text{T7})}{\tilde{p}_2(\text{Grim}) + \tilde{p}_2(\text{T8})}$ and $\frac{\tilde{p}_2(h_*^{t-1})(\text{RD})}{\tilde{p}_2(h_*^{t-1})(\text{Grim}) + \tilde{p}_2(h_*^{t-1})(\text{T8})} = \frac{\tilde{p}_2(\text{RD})}{\tilde{p}_2(\text{Grim}) + \tilde{p}_2(\text{T8})}$, (7) holds if $\frac{\tilde{p}_2(\text{T7}) + \tilde{p}_2(\text{RD})}{\tilde{p}_2(\text{Grim}) + \tilde{p}_2(\text{T8})}$ is sufficiently small.

Under the conditions of Observations 2 and 3, hence, $\sigma_1 = \text{T8}$ and $\sigma_2 = \text{T7}$ are consistent with subjective rationality. The distribution of strategies in the population is hence given by

$$
\zeta_0(1 - \omega) \cdot \text{RD} + \zeta_0\omega \cdot \text{Grim} + \zeta_1 \cdot \text{T8} + \zeta_2 \cdot \text{T7}. \tag{8}
$$

Under (8), the ex ante mean of the cooperation rates in round $t \le 6$ is given by:

$$
\bar{x}^t = \zeta_0(1 - \omega)q + (\zeta_0\omega + \zeta_1 + \zeta_2)\left\{\zeta_0(1 - \omega)q^{t-1} + \zeta_0\omega + \zeta_1 + \zeta_2\right\}.
$$

Likewise, the ex ante means of the cooperation rates in round 7 and 8 are given by

$$
\bar{x}^7 = \zeta_0(1 - \omega)q + (\zeta_0\omega + \zeta_1)\left\{\zeta_0(1 - \omega)q^6 + \zeta_0\omega + \zeta_1 + \zeta_2\right\},
$$

and

$$
\bar{x}^8 = \zeta_0(1 - \omega)q + \zeta_0\omega\left\{\zeta_0(1 - \omega)q^7 + \zeta_0\omega + \zeta_1\right\}.
$$

D.3 Numerical Illustration

We use numerical computation to illustrate the transitions of $\bar{x}^t$ and $\bar{\mu}^t$ derived above based on two different specifications of prior beliefs. In the first specification, the belief $\tilde{p}_k$ of a level-k player ($k = 1, 2$) places positive probabilities only on those strategies played by levels below k. The level-k belief is further assumed to be proportional to the actual proportions of players at levels below k:

$$\tilde{p}_1 = \sigma_0, \quad \text{and} \quad \tilde{p}_2 = \frac{\zeta_0}{\zeta_0 + \zeta_1} \cdot \sigma_0 + \frac{\zeta_1}{\zeta_0 + \zeta_1} \cdot \sigma_1. \tag{9}$$

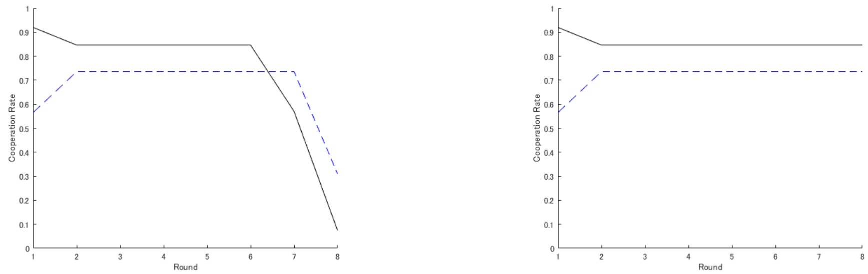
In the second specification, the belief $\tilde{p}_k$ of a level-k player ($k = 1, 2$) places positive probability also on the strategy σ_k played by level-k players. The level-k belief is assumed to be proportional to the actual proportions of players at levels k and lower:

$$\tilde{p}_1 = \frac{\zeta_0}{\zeta_0 + \zeta_1} \cdot \sigma_0 + \frac{\zeta_1}{\zeta_0 + \zeta_1} \cdot \sigma_1, \quad \text{and} \quad \tilde{p}_2 = \zeta_0 \cdot \sigma_0 + \zeta_1 \cdot \sigma_1 + \zeta_2 \cdot \sigma_2. \tag{10}$$

Figure 8 in the text as well as Figures 30 and 31 below depict $\bar{x}^t$ (solid line) and $\bar{\mu}^t$ (dashed line) for two different values of q, the probability with which RD plays C in each round. Figures 8 and 30 use the specification of prior beliefs in (9), and Figure 31 uses the specification in (10).[61]

These transition patterns can be interpreted as follows: First, in the Indefinite game, the cooperation rates $\bar{x}^t$ gradually decline over time since whenever RD plays D, Grim switches to AD and will never return to C. As time passes by, the average cooperation rates approach the probability that both players play Grim. On the other hand, there are two key forces behind the movement of the round beliefs $\bar{\mu}^t$. First, along the cooperative path h_*^t, the round beliefs monotonically increase (to 1) since that indicates that the strategy played by the other player is less likely to be RD. When $q = 0$, RD is immediately excluded after the other player plays C. On the other hand, the probability of the cooperative path h_*^t decreases with t as noted above and once $h^{t-1} \neq h_*^{t-1}$ is observed, the round t beliefs of levels 1 and 2 drop to q and stay there. The increasing pattern of $\bar{\mu}^t$ indicates that the first positive effect is stronger than the second negative effect. To see why the beliefs are more pessimistic initially (i.e., $\bar{x}^t - \bar{\mu}^t$ is positive but decreases with t), consider the second specification (9) and suppose that RD never plays C ($q = 0$). In this case, the round 1 belief just equals the proportion of Grim in the population as perceived by level-1

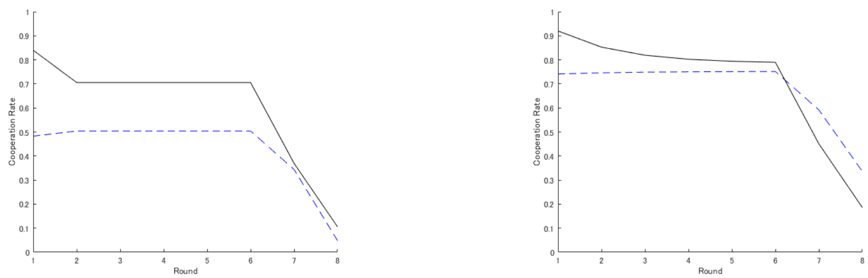
[61]The relevant conditions in Observations 1, 2 and 3 hold in all cases so that the level-k strategy σ_k is a best response to the level-k belief $\tilde{p}_k$ for $k = 1, 2$.



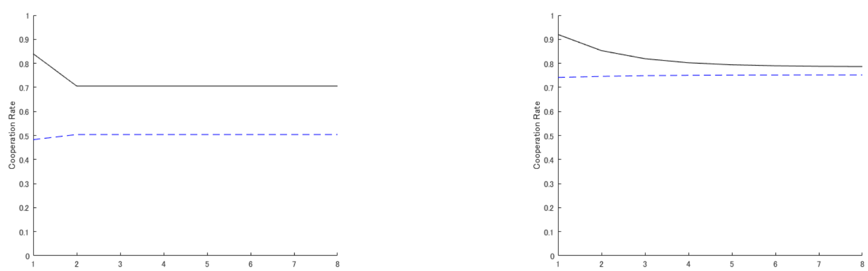
(a) Finite (b) Indefinite

Figure 30: Cooperation rates

Notes: $\bar{x}^t$ *(solid line) and round beliefs* $\bar{\mu}^t$ *(dashed line) when priors are given by (9).* $(\zeta_0, \zeta_1, \zeta_2) = (0.2, 0.5, 0.3)$, $\omega = 0.6$ *and* $q = 0$.



(a) Finite ($q = 0$) (b) Finite ($q = 0.5$)



(c) Indefinite ($q = 0$) (d) Indefinite ($q = 0.5$)

Figure 31: Cooperation rates

Notes: $\bar{x}^t$ *(solid line) and round beliefs* $\bar{\mu}^t$ *(dashed line) when priors are given by (10).* $(\zeta_0, \zeta_1, \zeta_2) = (0.4, 0.2, 0.4)$ *and* $\omega = 0.6$.

and level-2. On the other hand, the cooperation rates equal the actual proportion of level-1 and level-2 in the population. Since level-1 is not aware of the presence of level-2, the actual cooperation rates are higher in round 1. Conditional on the play of C by the other player in round 1, however, the level-1 correctly updates his belief and thinks that the other player also plays Grim. This correction helps reduce the gap between $\bar{\mu}^t$ and $\bar{x}^t$ from round 2 on.

Second, in the Finite game, unraveling is incomplete and the transitions of $\bar{x}^t$ and $\bar{\mu}^t$ are exactly the same as those in the Indefinite game up to round 6.[62] The decline of $\bar{x}^t$ in round 8 is caused by both level-1 and level-2, whereas its decline in round 7 is caused by level-2. Note that $\sigma_2 =$ T7 played by level-2 contributes to further reduction in cooperation in round 8 since it triggers D by Grim and T8 in round 8 by playing D in round 7. The round 7 belief $\bar{\mu}^7$ is different between the two specifications of prior beliefs. Under (9), there is no unraveling yet in round 7 since even level-2 does not expect any defection by T7. Under (10), on the other hand, unraveling begins in round 7 because level-2 correctly anticipates D by T7. The round 8 belief $\bar{\mu}^8$ is further lowered by two forces: First, level-2 (and level-1 in the case of (10)) expects $\sigma_1 =$ T8 to switch to D even along the cooperative path. Second, since level-2, who has played D in round 7, expects that T8 and Grim will revert to D.

D.4 Individual versus Team Play

Suppose that two individuals are randomly matched to form a team. A unit mass of these two-player teams are then randomly matched to play the repeated PD games against another team. Under the "Truth Wins norm," the sophistication level of a team equals the higher of the two sophistication levels of its members. For example, if an individual with level $k = 0$ is paired with an individual with level $k = 1$, the sophistication level of the resulting team equals $k = 1$. When the proportion of level-k individuals in the population equals ζ_k ($k = 0$, 1, 2), the proportion ξ_k of the level-k team under the truth wins norm equals

$$\xi_0 = \zeta_0^2, \quad \xi_1 = \zeta_0\zeta_1 + \zeta_1^2, \quad \xi_2 = 1 - (1 - \zeta_2)^2.$$

As for the prior belief of a level-k team over the strategy distribution, we assume that it is the ξ-adjusted belief of its member with the higher level of sophistication.

[62]This is because the only threshold strategies included in the analysis here are T7 and T8. If T6 is included as level-3, for example, the coincidence between the Finite and Indefinite games holds only in rounds 1-5.

For example, suppose that the two members of a team are level-1 and level-2, and assume that they place zero belief weight on the own level. The prior belief of the team is then level-2 based on the team strategy distribution above and given by

$$\tilde{p}_2 = \frac{\xi_0}{\xi_0 + \xi_1}\,\sigma_0 + \frac{\xi_1}{\xi_0 + \xi_1}\,\sigma_1,$$

where σ_k is the level-k strategy.

Figures 32 and 33 show the mean cooperation rates $\bar{x}^t$ (solid line) and mean beliefs $\bar{\mu}^t$ (dashed line) under individual and team play when the level-k belief places zero weight on the level-k strategy. Likewise, Figures 34 and 35 show the mean cooperation rates $\bar{x}^t$ (solid line) and mean beliefs $\bar{\mu}^t$ (dashed line) under individual and team play when the level-k belief places positive weight on the level-k strategy.

Whether the belief weight on the own strategy is zero or positive, the mean cooperation rates are higher under team play than under individual play in the Indefinite games. In the Finite games, on the other hand, the mean cooperation rates under team play are higher in earlier rounds, but drop more sharply toward the end. The mean cooperation rates are indeed lower in rounds 7 and 8 under team play than under individual play. These are consistent with the experimental evidence found by Kagel and McGee [2016] and Cooper and Kagel [2023] in the finitely and indefinitely repeated PD games, respectively.

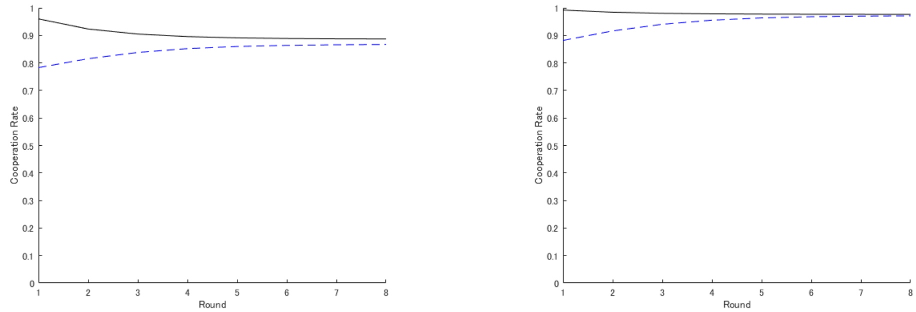


Figure 32: Indefinite Games: Individuals (left) and Teams (right)
Level-k belief places zero weight on the level-k strategy.
$(\zeta_0, \zeta_1, \zeta_2) = (0.5, 0.3, 0.2)$, $q = 0.5$.

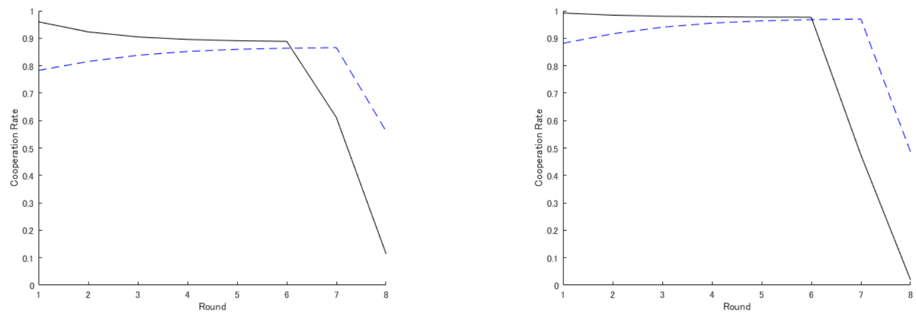


Figure 33: Finite Games: Individuals (left) and Teams (right)
Level-k belief places zero weight on the level-k strategy.
$(\zeta_0, \zeta_1, \zeta_2) = (0.5, 0.3, 0.2)$, $q = 0.5$.

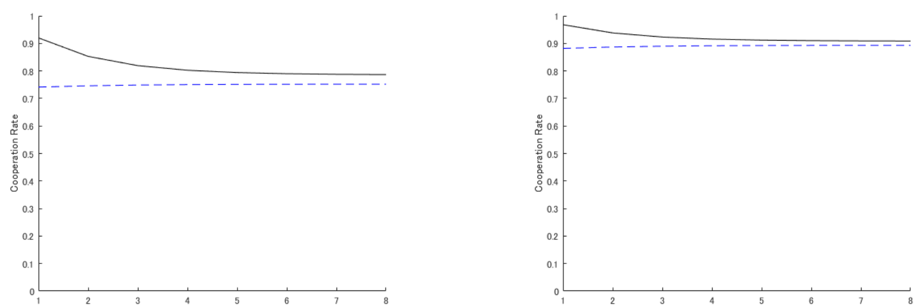


Figure 34: Indefinite Games: Individuals (left) and Teams (right)
Level-k belief places positive weight on the level-k strategy.
$(\zeta_0, \zeta_1, \zeta_2) = (0.4, 0.2, 0.4)$, $q = 0.5$.

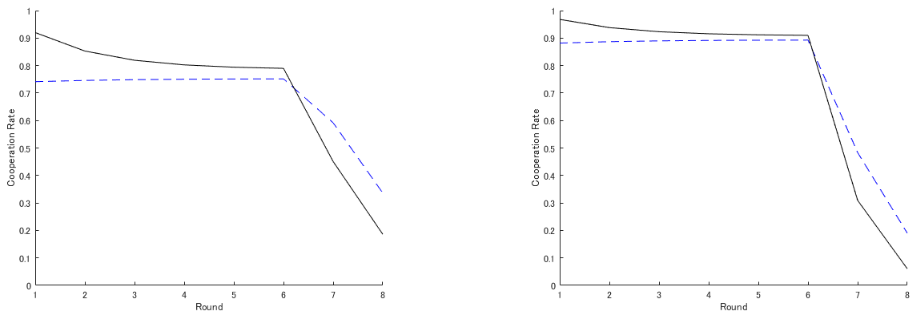


Figure 35: Finite Games: Individuals (left) and Teams (right)
Level-k belief places positive weight on the level-k strategy.
$(\zeta_0, \zeta_1, \zeta_2) = (0.4, 0.2, 0.4)$, $q = 0.5$.

E Additional Analysis on Robustness

E.1 Belief Recovery Method Simulations

In the next few sections, in addition to other robustness exercises, we document the behavior of our belief-estimation method using simulations. Here we organize these results to facilitate reading. We first present simulations pertaining to the main estimation that assumes Bayes updating. It is followed by simulations for the Grether updating specifications.

First, we show that the method recovers the correct beliefs in a simple stylized example of a population that consists of only AD (25%) , Grim (40%) and TFT types (35%). We simulate data—including both actions and round-by-round beliefs—based on the model of belief formation described in the paper assuming the following supergame beliefs for the different types. AD types believe others are playing AD with 40% probability, Grim with 10% probability, and TFT with 50% probability. Grim types believe others are playing AD with 10% probability, Grim with 30% probability, and TFT with 60% probability. TFT believe others are playing AD with 20% probability, Grim with 50% probability, and TFT with 30% probability.For this simulation, and all other simulations with this "three types" setup, with regards to simulating action choices, we set $\beta = 0.9365$, the average estimated value for this parameter in the experiment (using values from the Finite and Indefinite games). Also for all simulations of this type, with regards to simulating belief reports, we set $\tilde{\beta} = 1.00$ and $\nu = 0.05$, which are the median estimated values for these parameters from the experiment (including all types in the Finite and Indefinite games). The simulations are performed on supergames of eight rounds.[63] Table 20 summarizes the parameters of these simulations, as well as others presented in this appendix.

Figure 36 plots how well the belief recovery method estimates the beliefs of each simulated type. Note that this involves all three steps of our method: 1.(a) Estimating SFEM on the simulated data. 1.(b) Typing each simulated subject using the population level SFEM estimates as a prior and the subjects specific choices to determine the posterior. 2. Finally, for each strategy type, estimating beliefs over strategies given the simulated round-by-round beliefs. As such, it allows for errors at each of these steps, including incorrectly typing subjects. The figure highlights the impact of sample size by displaying results for simulations using two sessions, four sessions, and eight sessions. As can be seen, median parameter estimates are close

[63] Finite versus indefinite does not matter for the recovery technique except insofar as it affects the number of rounds. Eight rounds is the minimum we have, and thus a lower bound on performance.

Table 20: Simulations

Figure	Panel	Sessions	Termination	Types	DGP Updating	DGP ν	Grether c	Grether d	Estimator Updating
	Top	2	Finite	3	Bayes	logistc			Bayes
36	Middle	4	Finite	3	Bayes	logistc			Bayes
	Top	8	Finite	3	Bayes	logistc			Bayes
37		8	Finite	10	Bayes	logistc			Bayes
38		8	Indefinite	10	Bayes	logistc			Bayes
39		8	Finite	3	Bayes	normal			Bayes
44		8	Finite	3	Grether	logistic	0.75		Bayes
45	Top	8	Finite	3	Bayes	logistic			Grether
45	Bottom	8	Finite	3	Grether	logistic		0.75	Grether

100 Experiments per simulations (except in Figures 37 and 38).
18 subjects per session, each with 3 supergames per subject (except in Figures 37 and 38).
In all cases ν is truncated version.

to the true value in all cases. Furthermore, in a relatively simple setting such as this one, even with only two sessions, estimates are typically close to the true value.

Next we consider a similar exercise, but for conditions similar to the ones in our data set. Namely, the data generating process is assumed to correspond to the one we report in Tables 15 and 16. The sample size is assumed to be the same as the one we have collected in the experiment. 150 simulated data sets are produced for each treatment.[64] Figures 37 and 38 show that the input parameters are recovered quite well for the most common types. One notable exception is the supergame beliefs of the AC in the Indefinite game, which are not recovered as well as other types (AC as a SFEM estimate of 11% of the population). However, it is useful to note the nature of the discrepancy in this case: input values are such that the AC type puts 80% probability on others playing AC; the recovered values are such that some of this weight is shifted to TF2T. Thus, the discrepancy between the input and output values are among the most cooperative two strategies.

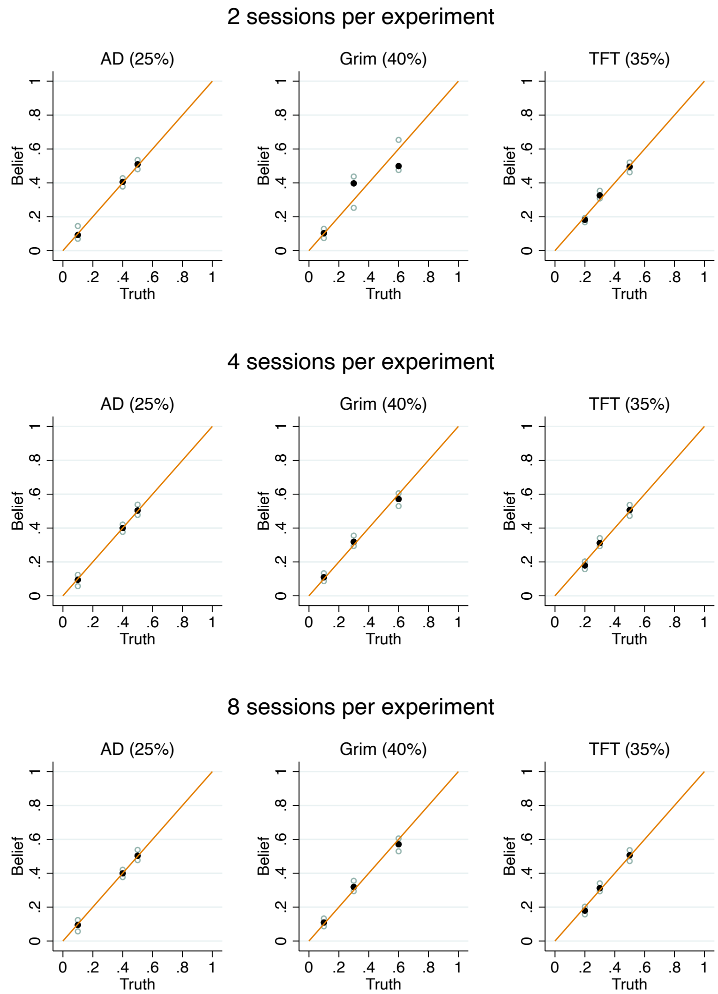
Figure 39 reports results from estimates that would result if the error in belief reporting ν is incorrectly specified in our estimation. Specifically, the data generating process assumes that reporting errors are distributed as a truncated normal, although our estimation assumed a truncated logistic. Other parameters of the simulations

[64]In the data from the experiment, in a few cases, beliefs over two strategies of a given type cannot be distinguished because no history is observed that would allow identification. When a simulated sample allow identification that is not in our original sample, we drop that sample.

are set as in the eight session simulation of Figure 36. Our estimates of beliefs are still very good in this case.

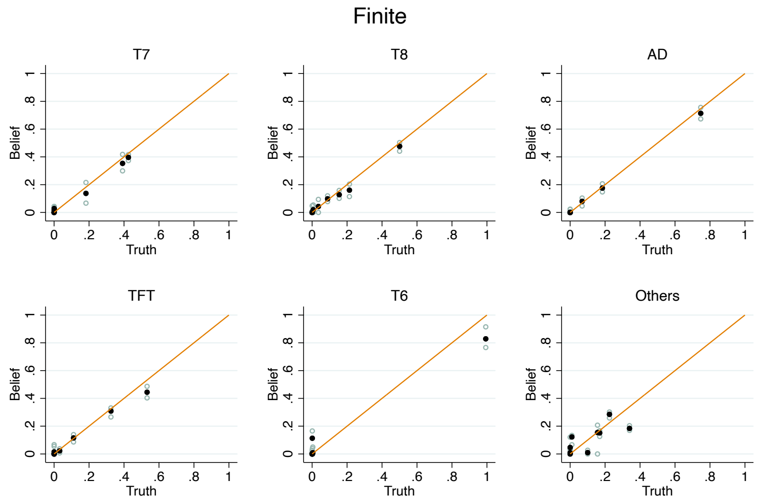
In the Robustness Section of the paper, we argue that given the $\tilde{\beta}$ estimated in our experiments, our results cannot be meaningfully affected by non-Bayesian updating that distorts signals in the form of $c \neq 1$ in the Grether updating formula. Figure 44 provides evidence of this by simulating data where agents are non-Bayesian, and in particular they have parameters $c = 0.75$ and $d = 1$ in the Grether formulation. However, our estimation assumes they update according to Bayes. Other parameters of the simulations are set as in the eight session simulation of Figure 36. These results align with the intuition provided in the text, namely that given the $\tilde{\beta}$ we observe, our results are robust to such non-Bayesian updating.

Figure 45 presents estimation results for the Grether style non-Bayesian belief recovery. In one case parameter d is assumed to be one, i.e. the simulated subjects are actually Bayesians. In the other case, $d = 0.75$, and the simulated subjects suffer from base-rate-neglect. Other parameters of the simulations are set as in the eight session simulation of Figure 36. As can be seen, the estimate of d move in the correct direction between the two simulations and the median estimates are close to the true value. The belief estimates are overall reasonable, although they become less precise.



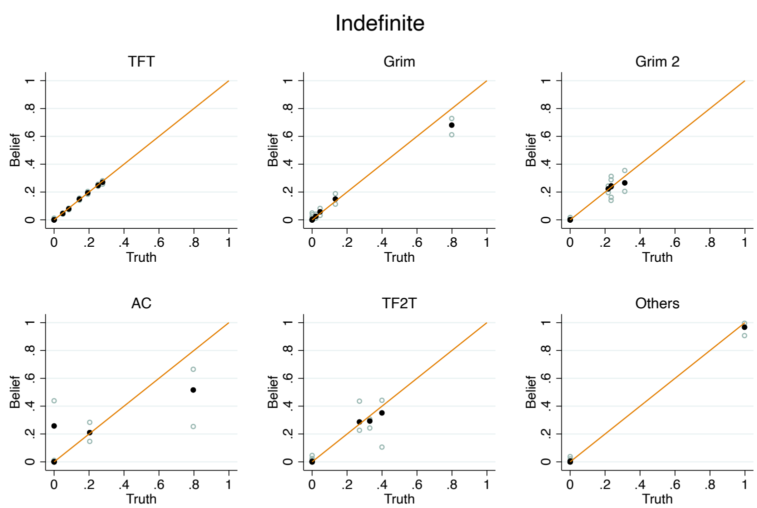
Estimation results from 100 simulated experiments with 18 subjects in each session.
Truth refers to input values.
Solid dots represent median estimate, hollow bubbles represent 25th and 75th percentile estimates.
Input values for all other parameters (not depicted in graphs) correspond to median values from belief estimates reported in the paper.

Figure 36: Estimation results using simulations



Truth refers to input values.
Solid dots represent median estimate, hollow bubbles represent 25th and 75th percentile estimates from 150 simulations.
We report simulation results for which identification matches original experiment (96% of data).

Figure 37: Estimation results using simulations



Truth refers to input values.
Solid dots represent median estimate, hollow bubbles represent 25th and 75th percentile estimates from 150 simulations.
We report simulation results for which identification matches original experiment (99% of data).

Figure 38: Estimation results using simulations

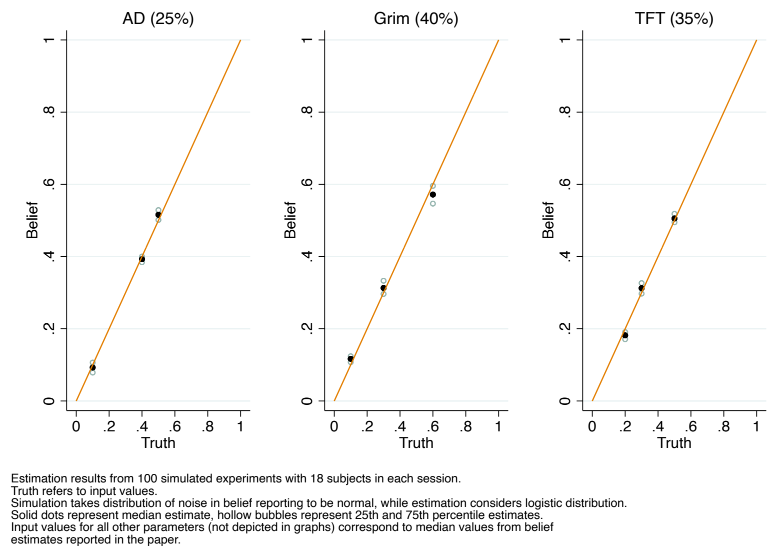
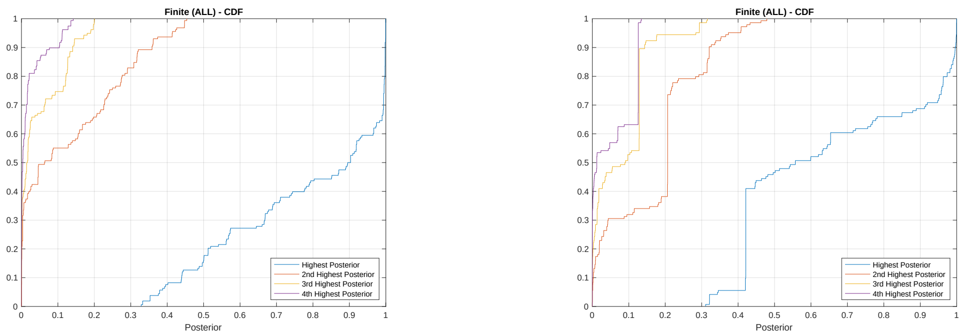


Figure 39: Estimation results using simulations with incorrect noise specification

E.2 Robustness with Respect to Typing



(a) (b)

Figure 40: Distribution of Posteriors in the Finite and Indefinite Game

Belief Estimates Under Alternative Simplified Typing

This section reports the belief-estimation results under a simplified alternative approach as described in Section 7.1. The method only considers (i) the consistency of actions with each strategy. To focus on subjects who are are clearly playing different strategies only a small set of strategies is considered. Namely, the most popular defective strategy (AD) and the most popular cooperative strategies (T7 for the Finite game and TFT for the Indefinite game). A subject is classified as playing one of these strategies if the consistency of their actions with that strategy is 90% or more and consistency of their actions with the other strategies is less than 90%. This classification labels 27% as T7, 18% as AD, and 9% as TFT for the Finite game. The numbers are 5%, and 60% for AD and TFT respectively for the Indefinite game (we do not include T7 since it only accounts for 1% of subjects).

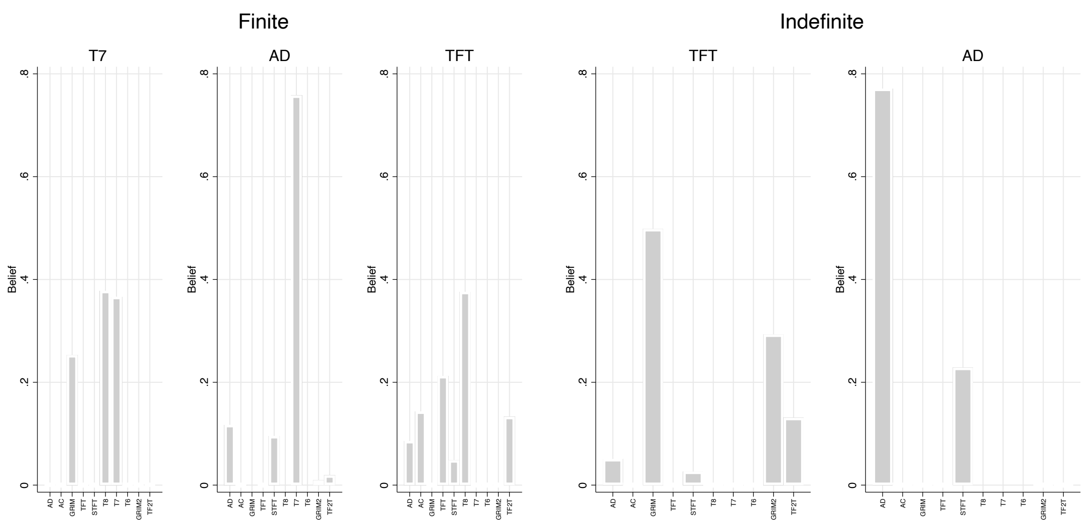
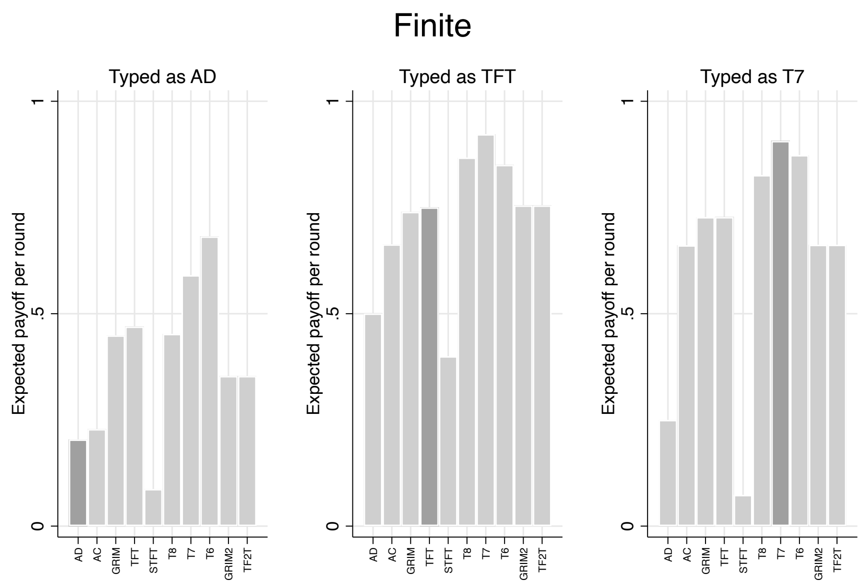
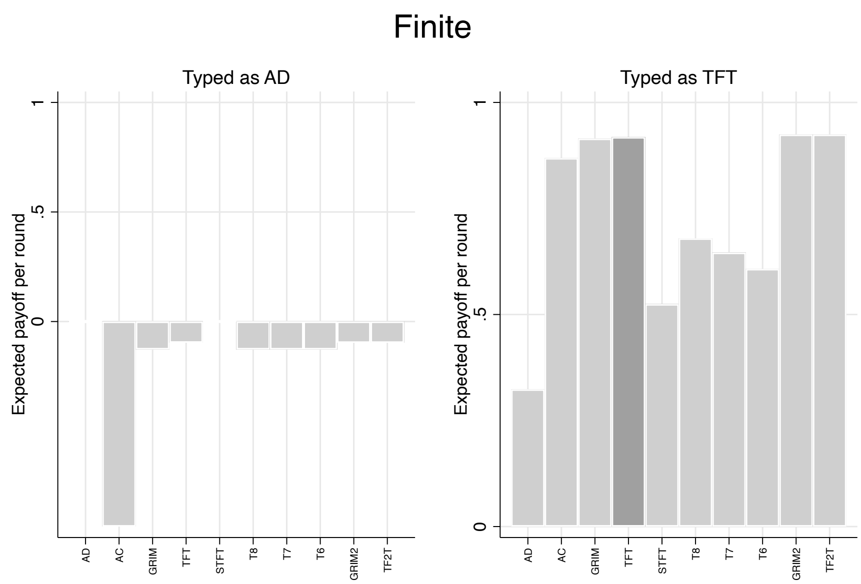


Figure 41: Estimated Beliefs Based on Simplified Typing for AD, T7 and TFT



The strategy corresponding to the type is higlighted in dark grey.
Analysis uses normalized stage-game payoffs.



The strategy corresponding to the type is higlighted in dark grey.
Analysis uses normalized stage-game payoffs.

Figure 42: Normalized Expected Payoff by Type Given Estimated Beliefs Late Supergames

E.3 Robustness with Respect to Non-Bayesian Updating

The goal of if this section is to study the extent to which the Bayesian assumption impacts our main results, we re-estimate beliefs allowing for non-Bayesian updating.

There are many ways in which our estimation method can be enriched to allow for deviations from the Bayesian benchmark. Grether [1980] provides a conceptual framework which differentiates between two types of non-Bayesian behavior: The first, denoted with parameter c, captures responsiveness to signals; and the second, denoted with parameter d, captures responsiveness to the prior.[65] In our setting, the prior corresponds to a subject's beliefs in round one about their opponent's strategy and the signals correspond to the actions taken by their opponent, which impacts the subject's updated beliefs in subsequent rounds. Our belief recovery procedure (as implemented in Section 5) already allows for errors. Indeed, $\tilde{\beta}$ captures (potentially incorrect) beliefs about how *noisy* actions are given strategy choice and therefore impacts how responsive updated beliefs are to observed actions, compressing belief reports toward 0.5; while the reporting error ν moves round beliefs up and down around the true value. However, these variables cannot be directly mapped into Grether's c and d parameters.[66] In our specific application, unlike in the typical bookbag-and-poker-chip inference experiment, the signals are perceived as very informative, i.e. $\tilde{\beta}$ is very close to one. An implication is that the Grether parameter c has little effect on updated beliefs.[67] For that reason, in what is presented below, we focus on a special case of the Grether framework with only one free parameter (d).[68]

[65]The Grether framework has become the standard approach to study non-Bayesian updating in empirical work (see Benjamin [2019]). Formally, given two states A and B and a signal S, the posterior π is given by:

$$\pi(A|S) = \frac{p(S|A)^c p(A)^d}{p(S|A)^c p(A)^d + p(S|B)^c p(B)^d}. \tag{11}$$

Hence, $c = d = 1$ corresponds to Bayesian updating, whereas $c < 1$ corresponds to underinference (sometimes also referred to as conservatism) while $d < 1$ to base rate neglect.

[66]To see this note that c and d directly capture deviations in updating and thus, by definition, can only impact beliefs after round one. By contrast, $\tilde{\beta}$ and ν have implications also for round one.

[67]This is because $p(S|A)$ in Equation 11 is either 0 or 1 (or very close to that). Figure 44 of the Online Appendix E.3 repeats the simulation of Figure 36 (with eight sessions) and shows that if the data generating process is actually one with $c = 0.8$ and $d = 1$, then estimates are almost identical to when $c = 1$.

[68]Base-rate neglect (captured by $d < 1$) is one of the most frequently documented biases in updating (going back to Kahneman and Tversky [1973]). See Benjamin, Bodoh-Creed, and Rabin [2019] and Esponda, Vespa, and Yuksel [2024] for recent perspectives on this bias. More broadly, an active literature in experimental and behavioral economics investigates the factors (parameters, context,

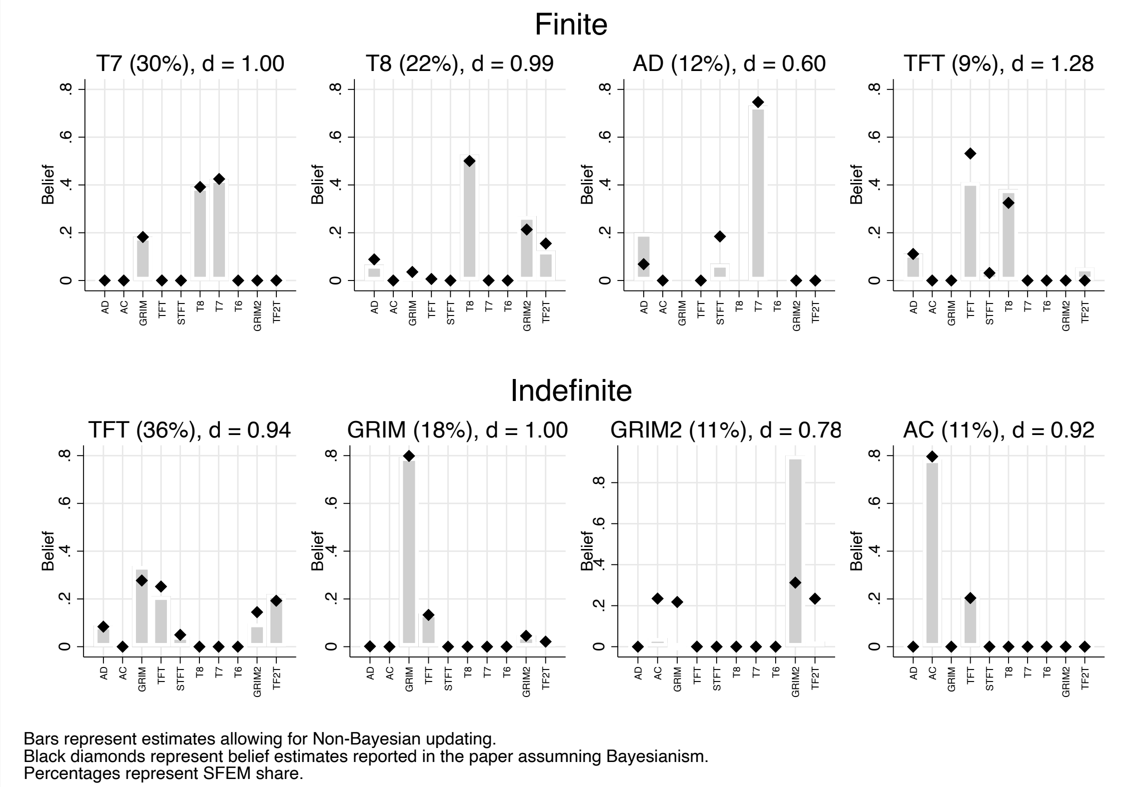


Figure 43: Beliefs over Strategies with Non-Bayesian Updating

These results are summarized for the four most common types in Figure 43, which also reports results from our original belief estimation for comparison.[69] At a qualitative level, allowing for non-Bayesian updating doesn't change our main results. When there are differences, beliefs move between similarly cooperative strategies.[70]

complexity) that predict the types of non-Bayesian behavior observed. Benjamin [2019] reviews the literature, in particular bookbag-and-poker-chip experiments, and finds varying results, but d is on average below 1. In recent papers, Augenblick et al. [2023] and Ba et al. [2023] identify that results from standard bookbag-and-poker-chip experiments can be reversed by changing elements of the paradigm (such as the number of states).

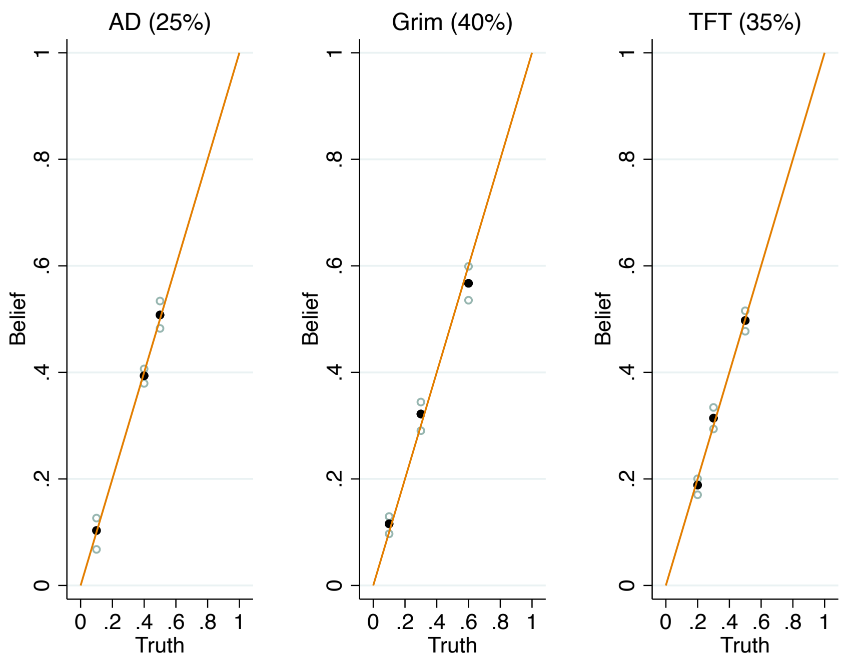
[69] Tables 21 and 22 present the complete results.

[70] Consider, for example, the AD type in the Finite game. The Grether parameter d is estimated to be low at 0.6 indicating base-rate neglect. Allowing for non-Bayesian updating mostly shifts beliefs from STFT to AD for this type. Similarly, in the Indefinite game, the d parameter is fairly low for type GRIM2 at 0.78. Allowing for non-Bayesian updating mostly shifts beliefs from AC, TF2T and Grim to Grim2 for this type. Nonetheless, overall, the estimates are fairly similar.

Importantly, the changes do not affect the interpretation of the results.[71]

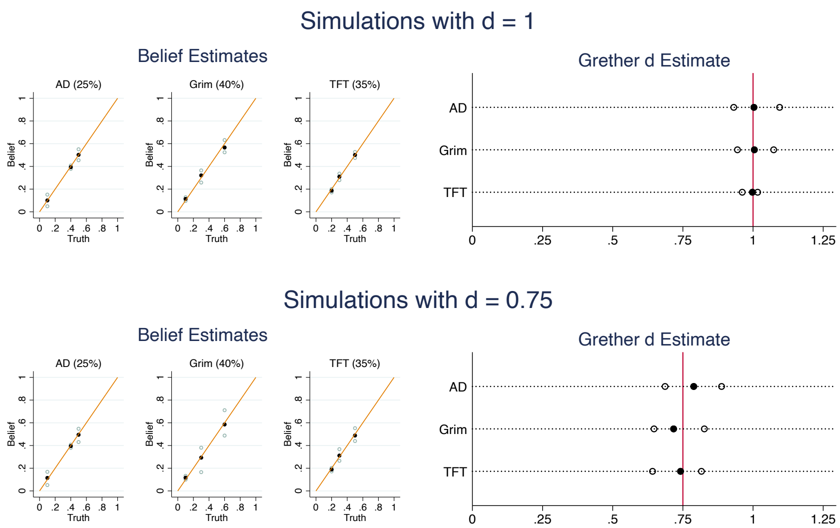
We note that our study, which focuses on beliefs in repeated games with perfect monitoring, does not provide the best setting to study deviations from Bayesian updating. But, in general, the belief recovery method can be generalized as demonstrated above to allow for such behavior. An environment with imperfect monitoring, for instance, where observed actions only carry limited information about the underlying strategies would be a richer setting to study non-Bayesian updating of beliefs in repeated games.

[71] It is still the case that beliefs over strategies capture the main differences between treatments: subjects mostly expect threshold strategies in the Finite game and conditionally cooperative strategies in the Indefinite game. In addition, the small changes in belief estimates do not change the finding that behavior is subjectively rational for most of the subjects. This can be seen in Figures 46 and 47 in that reproduce Figures 6 and 7 using the new estimates that allow for non-Bayesian updating.



Data is generated assuming Grether c = 0.75 and d = 1; but estimation assumes Bayesian updating.
Truth refers to input values.
Solid dots represent median estimate, hollow bubbles represent 25th and 75th percentile estimates.
Input values for all other parameters (not depicted in graphs) correspond to median values from belief estimates reported in the paper.

Figure 44: Simulation-Estimation Results with Grether Parameter $c < 1$



Estimation results from 100 simulated experiments with 18 subjects in each session.
Truth refers to input values (vertical line in panels to the right).
Solid dots represent median estimate, hollow bubbles represent 25th and 75th percentile estimates.
Input values for all other parameters (not depicted in graphs) correspond to median values from belief estimates reported in the paper.

Figure 45: Simulation-Estimation Results with Grether Parameter $d = 1$ and $d = 0.75$

Additional Grether parameter is represented as d in estimation results. See discussion above (Online Appendix E.3) for description of the parameter.

Table 21: Estimates for the Finite Game on Late Supergames

	Share		Estimated Beliefs - $\tilde{p}$												
	SFEM	TYPING	AD	AC	GRIM	TFT	STFT	T8	T7	T6	GRIM2	TF2T	ν	$\tilde{\beta}$	d
AD	0.12	0.12	0.20	0.00	*[0.00]*	0.00	0.07	*[0.00]*	0.73	*[0.00]*	0.00	0.00	0.06	1.00	0.60
			(0.13)	(0.02)		(0.16)	(0.14)		(0.23)		(0.05)	(0.04)			
AC	0.03	0.03	0.06	0.00	0.19	0.30	0.02	0.44	0.00	0.00	0.00	0.00	0.06	1.00	1.02
			(0.06)	(0.1)	(0.14)	(0.21)	(0.08)	(0.19)	(0.04)	(0.06)	(0.05)	(0.09)			
GRIM	0.08	0.02	0.33	0.10	0.02	0.15	0.00	0.40	0.00	0.00	0.00	0.00	0.07	1.00	0.85
			(0.27)	(0.11)	(0.33)	(0.07)	(0.08)	(0.12)	(0.09)	(0.01)	(0.03)	(0.13)			
TFT	0.09	0.12	0.11	0.00	0.00	0.41	0.04	0.38	0.00	0.00	0.00	0.05	0.05	1.00	1.28
			(0.1)	(0.09)	(0.16)	(0.29)	(0.04)	(0.13)	(0.06)	(0.01)	(0.14)	(0.14)			
STFT	0.03	0.03	0.00	0.00	*[0.00]*	0.82	0.00	*[0.00]*	0.00	0.00	0.00	0.18	0.07	1.00	0.39
			(0.02)	(0.16)		(0.42)	(0.02)		(0.07)	(0.02)	(0.09)	(0.36)			
T8	0.22	0.20	0.07	0.00	0.01	0.00	0.00	0.53	0.00	0.00	0.27	0.12	0.04	1.00	0.99
			(0.12)	(0.2)	(0.09)	(0.16)	(0.04)	(0.19)	(0.07)	(0.03)	(0.08)	(0.13)			
T7	0.30	0.35	0.00	0.00	0.18	0.00	0.00	0.39	0.42	0.00	0.00	0.00	0.04	1.00	1.00
			(0.02)	(0)	(0.11)	(0.17)	(0.01)	(0.15)	(0.14)	(0)	(0.01)	(0)			
T6	0.08	0.08	0.00	0.00	*[0.00]*	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.03	1.00	1.19
			(0.1)	(0.03)		(0.27)	(0.08)	(0.22)	(0.33)	(0.13)	(0.04)	(0.07)			
GRIM2	0.02	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-
TF2T	0.04	0.04	0.00	0.14	0.83	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.04	1.00	1.00
			(0.1)	(0.18)	(0.25)	(0.21)	(0.05)	(0.15)	(0.16)	(0.07)	(0.17)	(0.07)			
ALL			0.08	0.01	0.10	0.08	0.01	0.32	0.30	0.00	0.06	0.04			

Estimation on late supergames. SFEM estimate for β is 0.94.
Estimates in *[square brackets]* are not estimated due to collinearity.
Estimates in *(brackets)* show bootstrapped standard deviation.

Table 22: Estimates for the Indefinite Game on Late Supergames

	Share		Estimated Beliefs - $\tilde{p}$												
	SFEM	TYPING	AD	AC	GRIM	TFT	STFT	T8	T7	T6	GRIM2	TF2T	ν	$\tilde{\beta}$	d
AD	0.09	0.10	1.00	0.00	0.00	0.00	0.00	*[0.00]*	*[0.00]*	*[0.00]*	0.00	0.00	0.04	1.00	0.94
			(0.24)	(0.04)	(0.14)	(0.06)	(0.19)				(0.01)	(0.02)			
AC	0.11	0.05	0.00	0.78	0.00	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.10	1.00	0.92
			(0.42)	(0.17)	(0.31)	(0.16)	(0.06)	(0.05)	(0.05)	(0.04)	(0.08)	(0.55)			
GRIM	0.18	0.09	0.00	0.00	0.79	0.14	0.00	0.00	0.00	0.00	0.05	0.02	0.06	1.00	1.00
			(0.15)	(0.06)	(0.36)	(0.39)	(0.04)	(0.06)	(0.05)	(0.03)	(0.1)	(0.15)			
TFT	0.36	0.59	0.09	0.00	0.34	0.21	0.04	0.00	0.00	0.00	0.10	0.22	0.01	1.00	0.94
			(0.08)	(0.19)	(0.14)	(0.24)	(0.04)	(0)	(0)	(0)	(0.15)	(0.14)			
STFT	0.04	0.04	0.48	0.00	0.37	0.00	0.00	0.00	0.00	0.00	0.15	0.00	0.07	1.00	0.70
			(0.24)	(0.14)	(0.05)	(0.17)	(0.09)	(0.04)	(0.04)	(0.04)	(0.04)	(0.16)			
T8	0.01	0.01	-	-	-	-	-	-	-	-	-	-	-	-	-
T7	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-
T6	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-	-
GRIM2	0.11	0.11	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.93	0.03	0.02	1.00	0.78
			(0.23)	(0.16)	(0.18)	(0.05)	(0.03)	(0.01)	(0.01)	(0.01)	(0.24)	(0.29)			
TF2T	0.10	0.01	0.00	0.27	0.33	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.01	1.00	0.95
			(0.35)	(0.06)	(0.12)	(0.08)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.17)			
ALL			0.14	0.12	0.32	0.13	0.01	0.00	0.00	0.00	0.15	0.13			

Estimation on late supergames. SFEM estimate for β is 0.94.
Estimates in *[square brackets]* are not estimated due to collinearity.
Estimates in *(brackets)* show bootstrapped standard deviation.

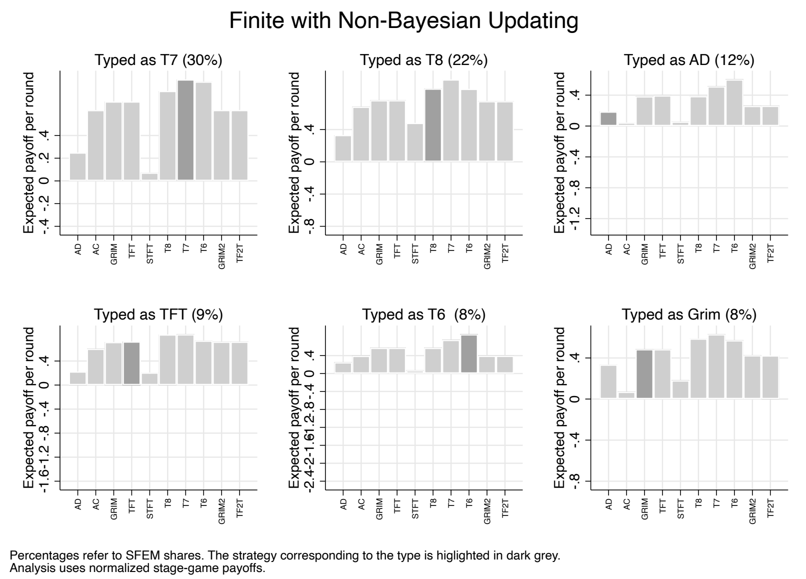


Figure 46: Normalized Expected Payoff by Type Given Estimated Beliefs (Allowing for Non-Bayesian Updating) in Late Supergames

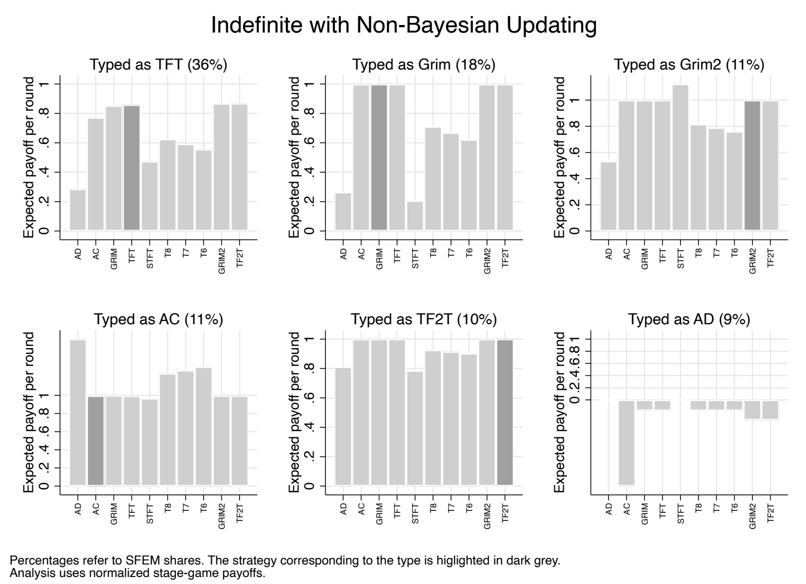


Figure 47: Normalized Expected Payoff by Type Given Estimated Beliefs (Allowing for Non-Bayesian Updating) in Late Supergames

E.4 New Indefinite Treatments

Table 23: Stage Game in New Treatments

High T	C	D
C	51, 51	22, 73
D	73, 22	39, 39

Low R	C	D
C	45, 45	22, 63
D	63, 22	39, 39

Table 24 replicates Table 5 for the new sessions. See discussion in main text around Table 5 for further information on how to read the Table. Everything was kept constant as in the original Indefinite game sessions except the changes in the stage game payoffs. Due to technical issues, we had to restrict session size to 16 subjects. We used the same seeds (to determine realization of supergame lengths) from the original Indefinite game, but the exact number of supergames played in each session showed variation relative to the original sessions, which required an adjustment of which supergames are included among the *early* and *late* supergames.[72]

Table 24: Session Summary of New Treatments

				No. of Game Rounds				Total no. of
		No. of	No. of	Actions	Actions and Beliefs			Obs.
Treatment	Session	Subjects	Supergames	Only	Early		Late	Rounds
High T	1	16	7	9, 7, 13, 7	*1,*		23	77
	2	16	8	8, 15, 7, 32	*2, 10,*		5, 1	97
	3	16	8	8, 2, 3, 14	*25,*		17, 10, 13	103
	4	16	8	9, 7, 10, 13	*32,*		7, 7, 6	96
	5	16	12	7, 22, 7, 3	*2, 5, 8,*	4, 14,	9, 3, 10	119
	6	16	8	1, 31, 4, 3	*24,*		15, 25, 3	127
	7	16	11	5, 6, 7, 14	*30, 8, 5,*	4,	9, 4,33	142
	8	14	11	11, 1, 4, 13	*9, 5, 2,*	4,	2, 2, 11	100
Low R	1	16	8	9, 7, 13, 7	*1, 2,*		23, 4	85
	2	16	8	8, 15, 7, 32	*2, 10,*		5, 1	97
	3	16	7	8, 2, 3, 14	*25,*		17, 10	90
	4	16	6	9, 7, 10, 13	*32,*		7	80
	5	16	10	7, 22, 7, 3	*2, 5, 8,*		4, 14, 9	101
	6	16	6	1, 31, 4, 3	*24,*		15	94
	7	16	10	5, 6, 7, 14	*30, 8, 5,*		4, 9, 4	109
	8	16	12	11, 1, 4, 13	*9, 5, 2*	4, 2,	2, 11, 3	108

[72] As before, we aimed for three supergames for both early and late when possible. When that was not possible, we aimed to have a division of total rounds that was as balanced as possible.

Table 25: Correlated Random Effects Probit
Determinants of Cooperation in Round One

	Low R	Low R	Low R	High T	High T	High T
Beliefs Are Elicited	0.369***	0.267	0.286	0.936***	0.133	0.155
	(0.119)	(0.173)	(0.208)	(0.144)	(0.192)	(0.204)
Supergame		0.0225	0.0261		0.180***	0.165***
		(0.0151)	(0.0275)		(0.0486)	(0.0493)
Other Cooperated in Previous Supergame			0.00136			0.505***
			(0.147)			(0.167)
Cooperated in Supergame 1			2.247***			1.766***
			(0.264)			(0.328)
Risk Measure			-0.0112			0.0230***
			(0.00762)			(0.00536)
Length of Previous Supergame						0.00382
						(0.00814)
Constant	-0.442**	-0.498**	-0.910*	0.501***	0.0689	-2.104***
	(0.203)	(0.214)	(0.469)	(0.188)	(0.236)	(0.265)
Observations	1072	1072	944	1146	1146	1020

Standard errors clustered (at the session level) in parentheses. ***1%, **5%, *10% significance.
All variables refer to behavior in Round 1.
Risk Measure is equal to the number of boxes collected in the bomb task.

Table 26: Correlated Random Effects Probit (Marginal Effects)
Determinants of Cooperation in Round One

	Low R	Low R	Low R	High T	High T	High T
Beliefs Are Elicited	0.0747***	0.0540	0.0501	0.175***	0.0239	0.0242
	(0.0228)	(0.0340)	(0.0344)	(0.0326)	(0.0354)	(0.0325)
Supergame		0.00455	0.00457		0.0324***	0.0257***
		(0.00311)	(0.00501)		(0.00870)	(0.00639)
Other Cooperated in Previous Supergame			0.000238			0.0788***
			(0.0257)			(0.0237)
Cooperated in Supergame 1			0.393***			0.276***
			(0.0323)			(0.0505)
Risk Measure			-0.00196			0.00359***
			(0.00134)			(0.000753)
Length of Previous Supergame						0.000596
						(0.00130)
Observations	1072	1072	944	1146	1146	1020

Standard errors clustered (at the session level) in parentheses. ***1%, **5%, *10% significance.
All variables refer to behavior in Round 1.
Risk Measure is equal to the number of boxes collected in the bomb task.

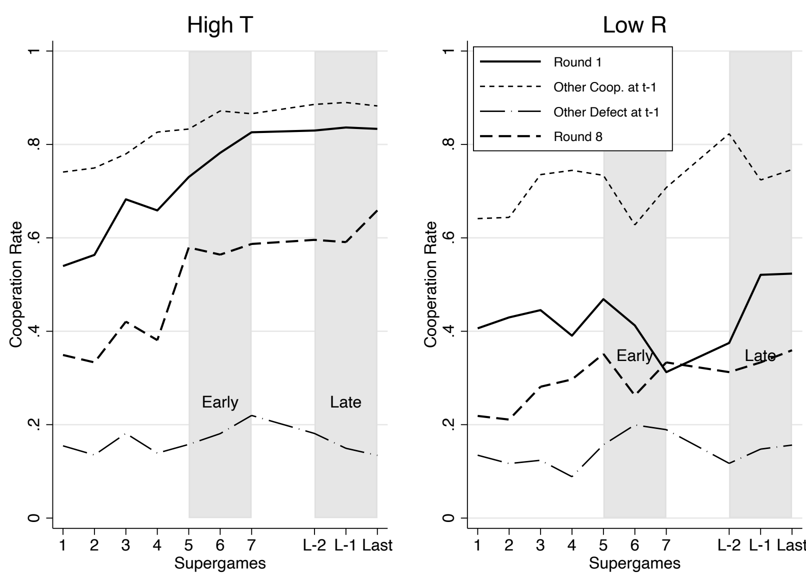
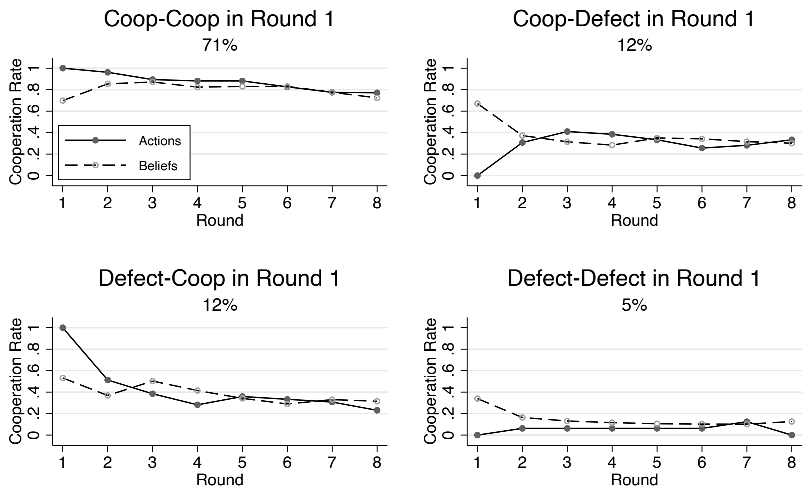
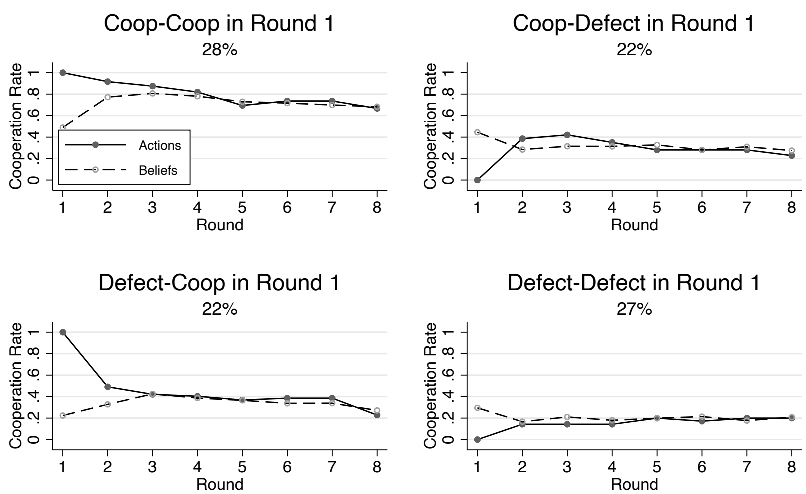


Figure 48: Cooperation Rate over Supergames



Own choice listed first.
Late supergames.
Percentage of cases under the title.
By construction the 'Actions' rate is 1 in round 1 of the left panels and 0 in the right panels.

Figure 49: Beliefs Conditional on Round One Action Pair, *Hight T*



Own choice listed first.
Late supergames.
Percentage of cases under the title.
By construction the 'Actions' rate is 1 in round 1 of the left panels and 0 in the right panels.

Figure 50: Beliefs Conditional on Round One Action Pair, *Low R*

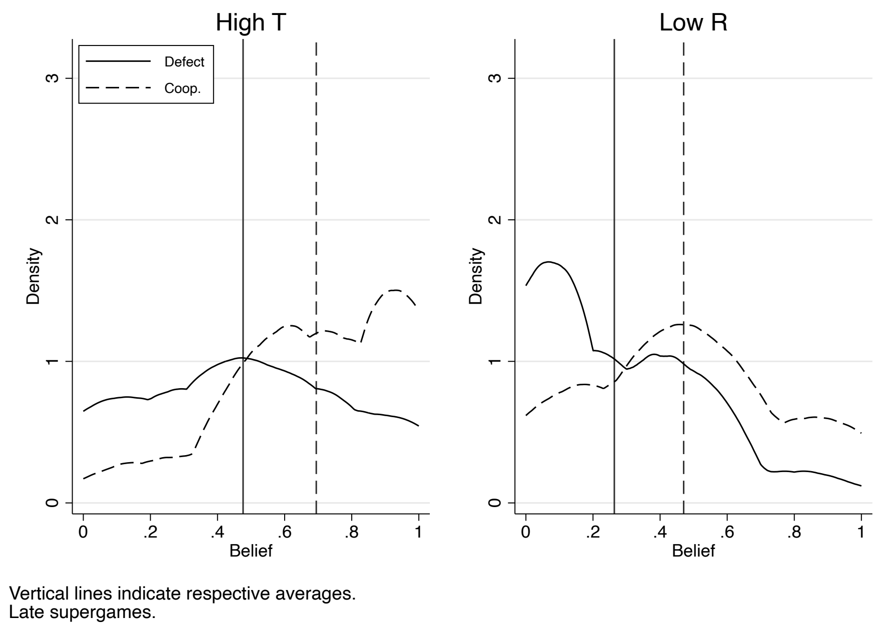


Figure 51: Beliefs of Defectors vs. Cooperators in Round One by Treatment

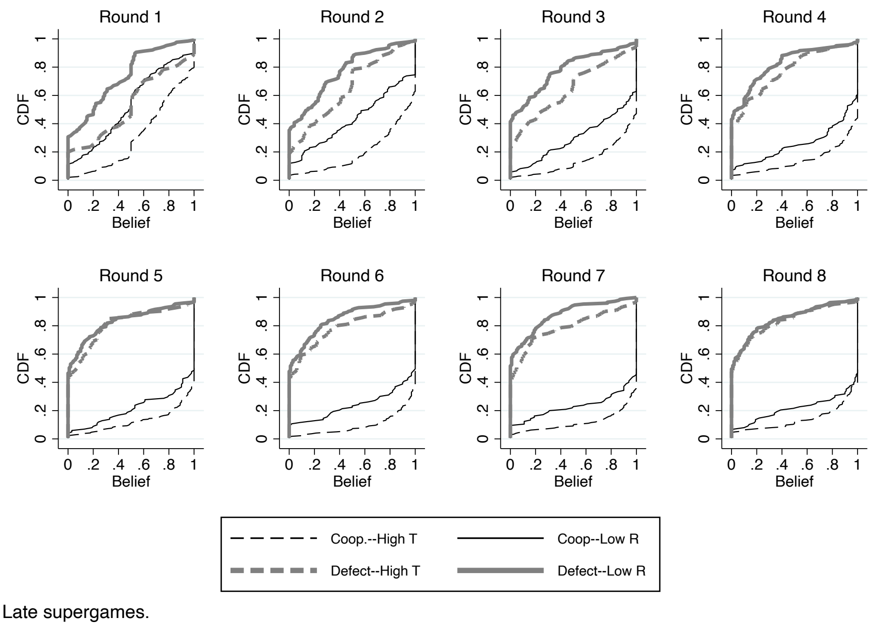


Figure 52: Beliefs by Action and Treatment: Rounds One through Eight

Table 27: Estimates for *High T* on Late Supergames

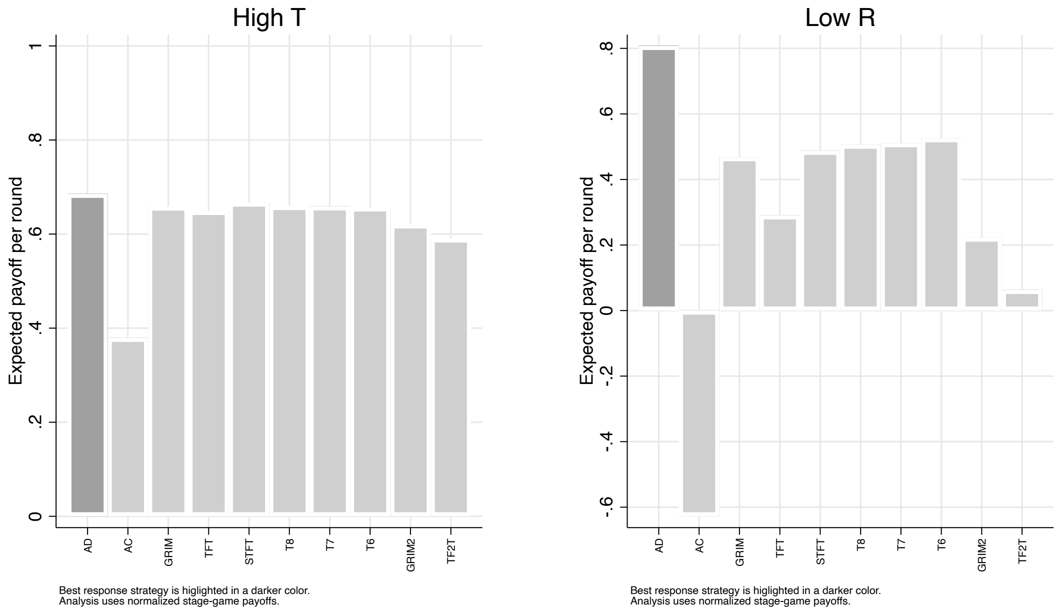
	Share		Estimated Beliefs - $\tilde{p}$											
	SFEM	TYPING	AD	AC	GRIM	TFT	STFT	T8	T7	T6	GRIM2	TF2T	ν	$\tilde{\beta}$
AD	0.11	0.13	0.40	0.14	0.26	0.00	0.04	*[0.00]*	*[0.00]*	*[0.00]*	0.08	0.08	0.05	1.00
			(0.19)	(0.1)	(0.15)	(0.1)	(0.06)				(0.08)	(0.08)		
AC	0.06	0.03	0.40	0.05	0.27	0.24	0.00	0.00	0.00	0.00	0.02	0.02	0.06	1.00
			(0.19)	(0.15)	(0.24)	(0.21)	(0.09)	(0.02)	(0.02)	(0.02)	(0.05)	(0.03)		
GRIM	0.24	0.13	0.22	0.00	0.50	0.23	0.06	0.00	0.00	0.00	0.00	0.00	0.02	1.00
			(0.08)	(0.03)	(0.16)	(0.15)	(0.04)	(0.02)	(0.01)	(0.01)	(0.03)	(0.05)		
TFT	0.29	0.46	0.26	0.00	0.41	0.19	0.00	0.00	0.00	0.00	0.00	0.13	0.02	1.00
			(0.08)	(0.03)	(0.12)	(0.09)	(0.02)	(0)	(0)	(0)	(0.04)	(0.06)		
STFT	0.02	0.02	0.44	0.00	0.00	0.00	0.56	*[0.00]*	*[0.00]*	*[0.00]*	0.00	0.00	0.07	1.00
			(0.22)	(0.05)	(0.03)	(0.03)	(0.28)				(0.07)	(0.07)		
T8	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
T7	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
T6	0.01	0.01	-	-	-	-	-	-	-	-	-	-	-	-
GRIM2	0.09	0.03	0.00	0.10	0.13	0.10	0.00	0.00	0.14	0.26	0.09	0.19	0.09	1.00
			(0.06)	(0.16)	(0.11)	(0.05)	(0.07)	(0.06)	(0.08)	(0.11)	(0.07)	(0.11)		
TF2T	0.18	0.18	0.10	0.00	0.42	0.24	0.10	0.00	0.00	0.00	0.00	0.13	0.06	1.00
			(0.09)	(0.09)	(0.25)	(0.23)	(0.1)	(0.03)	(0)	(0)	(0.03)	(0.18)		
ALL			0.23	0.03	0.37	0.18	0.05	0.00	0.01	0.02	0.02	0.09		

Estimation on late supergames. SFEM estimate for β is 0.92.
Estimates in *[square brackets]* are not estimated due to collinearity.
Estimates in *(brackets)* show bootstrapped standard deviation.

Table 28: Estimates for *Low R* on Late Supergames

	Share		Estimated Beliefs - $\tilde{p}$											
	SFEM	TYPING	AD	AC	GRIM	TFT	STFT	T8	T7	T6	GRIM2	TF2T	ν	$\tilde{\beta}$
AD	0.29	0.32	0.56	0.00	0.14	0.00	0.25	*[0.00]*	*[0.00]*	*[0.00]*	0.02	0.03	0.06	1.00
			(0.22)	(0)	(0.09)	(0.03)	(0.23)				(0.03)	(0.04)		
AC	0.02	0.03	0.89	0.00	0.00	0.00	0.11	0.00	0.00	0.00	0.00	0.00	0.02	1.00
			(0.38)	(0.02)	(0.02)	(0.04)	(0.09)	(0.02)	(0.02)	(0.02)	(0.02)	(0.04)		
GRIM	0.18	0.16	0.48	0.00	0.50	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.06	1.00
			(0.18)	(0.05)	(0.19)	(0.07)	(0.04)	(0.03)	(0.02)	(0.02)	(0.04)	(0.09)		
TFT	0.12	0.06	0.57	0.00	0.16	0.00	0.00	*[0.00]*	0.28	0.00	0.00	0.00	0.13	1.00
			(0.27)	(0.09)	(0.12)	(0.05)	(0.07)		(0.16)	(0.11)	(0.07)	(0.13)		
STFT	0.17	0.15	0.38	0.01	0.02	0.05	0.55	*[0.00]*	*[0.00]*	*[0.00]*	0.00	0.00	0.07	1.00
			(0.18)	(0.01)	(0.02)	(0.02)	(0.19)				(0.02)	(0.02)		
T8	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
T7	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
T6	0.00	0.00	-	-	-	-	-	-	-	-	-	-	-	-
GRIM2	0.18	0.26	0.31	0.00	0.31	0.00	0.19	0.00	0.00	0.00	0.00	0.19	0.02	1.00
			(0.04)	(0.01)	(0.05)	(0.02)	(0.03)	(0)	(0)	(0)	(0.06)	(0.06)		
TF2T	0.04	0.02	0.93	0.00	0.00	0.00	0.00	*[0.00]*	0.00	0.00	0.07	0.00	0.39	1.00
			(0.45)	(0.02)	(0.05)	(0.07)	(0.26)		(0.04)	(0.03)	(0.13)	(0.06)		
ALL			0.49	0.00	0.21	0.01	0.20	0.00	0.03	0.00	0.01	0.04		

Estimation on late supergames. SFEM estimate for β is 0.89.
Estimates in *[square brackets]* are not estimated due to collinearity.
Estimates in *(brackets)* show bootstrapped standard deviation.



(a) (b)

Figure 53: Normalized Expected Payoff by Type Given Strategy Distribution in Late Supergames

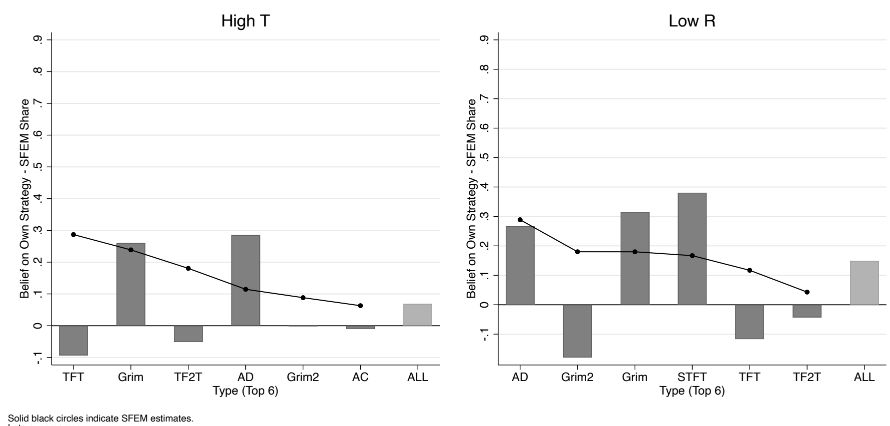


Figure 54: Overestimation in Beliefs of the Prevalence of One's Own Startegyt

F Instructions

INSTRUCTIONS

You are about to participate in an experiment on decision-making. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance. Please turn off cell phones and similar devices now. Please do not talk or in any way try to communicate with other participants.

We will start with a brief instruction period. If you have any questions during this period, raise your hand and your question will be answered so everyone can hear.

This experiment has three parts; these instructions are for the first part. Once this part is over, instructions for the next part will be given to you. Your decisions in this part have no influence on the other parts.

General Instructions

1. In this experiment you will be repeatedly matched with a randomly selected person in the room. During each match, you will be asked to make decisions over a sequence of rounds.

2. The points you can obtain in each round of a match depend on your choice and the choice of the person you are paired with. The table below represents all the possible outcomes:

Your Choice	**Other's Choice**	
	1	***2***
1	51, *51*	22 *63*
2	63, *22*	39, *39*

The table shows the points associated with each combination of your choice and choice of the person you are paired with. The first entry in each cell represents the points you obtain for that round, while the second entry (in italics) represents the points obtained by the person you are paired with.

That is, in each round of a match, if:
- (1,1): Your choice is 1 and the other's choice is 1, you each make 51.
- (1,2): Your choice is 1 and the other's choice is 2, you make 22 while the other makes 63.
- (2,1): Your choice is 2 and the other's choice is 1, you make 63 while the other makes 22.
- (2,2): Your choice is 2 and the other's choice is 2, you each make 39.

3. At the end of each round, you will see your choice (**1** or **2**) and the choice of the person you were paired with (**1** or **2**).

4. Each match will last for 8 rounds.

5. Once a match ends, you will be paired randomly with someone for a new match. You will not be able to identify who you've interacted with in previous or future matches.

6. Each part of the experiment will generate points that count towards your final payoff. In this part, one match will be randomly selected to count towards your final payoff. Points earned in this match will be converted to dollars at a rate of 3 cents per point. You will receive an additional $8 show up fee for your participation. You will only be informed of your payoffs at the end of the experiment.

7. This part will last for four matches.

Are there any questions?

Before we start, let me remind you that:

- Each match will last for 8 rounds. You will interact with the same person for the entire match.
- Your choice and the choice of the person you are paired with will be shown to both of you at the end of the round.
- Points obtained in each round depend on these choices.
- After a match is finished, you will be randomly paired with someone for a new match.

General Instructions for Part 2

The basic structure of this part is very similar to part 1. How the match proceeds and how you are paired with others will remain the same.

However, in this part, you will have one more task. In each round of a match, after you make a choice, we will ask you to submit your belief about the choice of the person you are paired with.

To indicate your beliefs, you will use a slider. Where you move the slider will represent your best assessment of the likelihood (expressed as chance out of 100) that the person you are paired with chose **1** or **2.**

Two different matches from this part will be randomly selected to count towards payment. For one of these, you will receive the points associated with your choices as in part 1. For the other, the computer will randomly choose one round from that match for payment for beliefs. The belief that you report in that round will determine your chance of winning a prize of 50 points.

To determine your payment, the computer will randomly draw two numbers. For each draw, all numbers between 0 and 100 (including decimal numbers) are equally likely to be selected. Draws are independent in the sense that the outcome of the first draw in no way affects the outcome of the second draw.

If the person you are paired with chose **1** in that round and the number you indicated as the likelihood that the other chose **1** is larger than either of the two draws, you will win the prize.

If the person you are paired with chose **2** in that round and the number you indicated as the likelihood that the other chose **2** is larger than either of the two draws, you will win the prize.

The rules that determine your chance of winning this prize were purposefully designed so that you have the greatest chance of winning the prize when you answer the question with your true assessment on how likely the person you are paired with chose **1** or **2**.

The first match to end after 60 minutes of play (including the first part) will mark the end of the experiment.

General Instructions for Part 3

On the screen, you see a field composed of 100 boxes, as shown below (the numbers on each box will not be visible):

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

There is also a Start button—please do not click on this button until we finish reading the instructions. Once the Start button is clicked, the experiment begins. Every two seconds, a box will be collected, beginning with Box #1 (top left) and ending with Box #100 (bottom right).

You earn 3 cents for every box that is collected. Once collected, the box changes from dark grey to light grey, and your earnings are updated accordingly. At any moment, on the information box, you can see the number of boxes collected so far and the amount earned up to that point.

Such earnings are only ***potential***, however, because behind one of these boxes a bomb is hidden that destroys everything that has been collected in this part of the experiment. You do not know the location of the bomb. Moreover, even if you collect the bomb, you will not know it until the end of the experiment. Your task is to choose when to stop the collecting process. You stop the process by hitting 'Stop' at any time.

Payoffs: If at the moment you hit 'Stop' none of the boxes you have collected contain the bomb, you will receive the amount of money you have accumulated. If at the moment you hit 'Stop' you happen to have collected the box with the bomb, then you will earn $0. Remember that you will not be told if a box that you have collected has or does not have the bomb until after you hit the 'Stop' button. So the earnings you see on the screen are only potential earnings, and you will earn those earnings only if none of the boxes you have collected had the bomb.

Location of the Bomb: The interface will randomly choose a number between 1 and 100. All numbers are equally likely. The interface will then place the bomb in the box with the randomly chosen number.

G Proof of the Cooperativeness Order

When each strategy is denoted by a finite automaton, we assume that an *implementation error* is made in the choice of an action in each state, and not in transition from the current state to the next. We also assume that the errors are independent and identically distributed between the players and across rounds. Denote by $\varepsilon \in [0, \frac{1}{2}]$ the probability of such an error.[73] For the analytical comparison of cooperative levels, we assume that ε is small. In some cases considered below, this implies that we treat ε^2 as negligible. In other cases, however, we need to consider the difference in the order of ε^2 and treat ε^3 as negligible. Let $p = (1-\varepsilon)^2$, $q = \varepsilon(1-\varepsilon)$ and $r = \varepsilon^2$. The normalized stage payoffs with implementation errors are given by

$$\begin{aligned} g_{CC} &= p + q(1+g-\ell), & g_{CD} &= p(-\ell) + q + r(1+g), \\ g_{CD} &= p(1+g) + q + r(-\ell), & g_{DD} &= q(1+g-\ell) + r, \end{aligned}$$

where $g = 1$ and $\ell = 17/12 \approx 1.416$ in our implementation. Define

$$g = \begin{bmatrix} g_{CC} \\ g_{CD} \\ g_{DC} \\ g_{DD} \end{bmatrix}.$$

We consider a Markov process induced by a pair of the same strategy implemented with errors ε. Let Θ be the set of states of this Markov process. For each strategy that can be expressed as an S-state automaton, Θ can have up to $S \times S$ elements. The Markov process is defined over the set $\Delta\Theta$ of distributions over those states. Let $\omega^1 \in \Delta\Theta$ be the row vector representing the initial distribution and $A = (a_{st})_{s,t\in\Theta}$ be the transition matrix: a_{st} is the probability that the next state is t when the current state is s. The distribution ω^2 over round 2 states is given by $\omega^2 = \omega^1 A$, and the distribution ω^t over round t states is given by $\omega^t = \omega^1 A^{t-1}$. With the distribution ω over states, the expected stage payoff to a player is given by ωg. In the case of the finite games, the average payoff over eight rounds can be computed as

$$\frac{1}{8}\sum_{t=1}^{8} \omega^t g = \frac{1}{8}\omega^1 \left(I + A^1 + \cdots + A^7\right) g. \tag{12}$$

[73] Hence, $\varepsilon = 1 - \beta$ for the parameter β in SFEM.

In the case of the indefinite games, the average discounted payoff can be computed as

$$\begin{aligned}(1-\delta)\sum_{t=1}^{\infty}\omega^t\delta^{t-1}\,g &= (1-\delta)\,\omega^1\left(I+\delta A^1+\cdots+\delta^t A^t+\cdots\right)\,g\\ &= (1-\delta)\,\omega^1(I-\delta A)^{-1}\,g,\end{aligned}\tag{13}$$

where $\delta = 7/8$ in our implementation. If we denote by v_θ the average discounted payoff in the indefinite games along the Markov process with the initial state θ (i.e., the initial distribution ω^1 places probability one on state θ), and by $v = (v_\theta)_{\theta\in\Theta}$ the corresponding column vector, then (13) implies the recursive equation

$$v = (1-\delta)\,(I-\delta A)^{-1}\,g \quad\Leftrightarrow\quad v = (1-\delta)\,g + \delta A v.\tag{14}$$

G.0.1 Indefinite games with small implementation errors

1. TFT and STFT: These strategies have two states 0 and 1. Both strategies play C in state 0, and D in state 1. Because the implementation errors occur independently between the two players, state transitions do not synchronize between them. Accordingly, the Markov process has four states $\Theta = \{(0,0),(0,1),(1,0),(1,1)\}$. The initial distribution is $\omega^1 = (1,0,0,0)$ if both play TFT and $\omega^1 = (0,0,0,1)$ if both play STFT. We hence have $v^{\mathrm{TFT}} = v_{00}$ and $v^{\mathrm{STFT}} = v_{11}$. The transition matrix is given by

$$A = \begin{bmatrix} p & q & q & r\\ q & r & p & q\\ q & p & r & q\\ r & q & q & p\end{bmatrix}.$$

   Ignoring the terms of order ε^2, we can write (14) as

$$\begin{bmatrix} v_{00}\\ v_{01}\\ v_{10}\\ v_{11}\end{bmatrix} = (1-\delta)\begin{bmatrix} g_{CC}\\ g_{CD}\\ g_{DC}\\ g_{DD}\end{bmatrix} + \delta\begin{bmatrix} 1-2\varepsilon & \varepsilon & \varepsilon & 0\\ \varepsilon & 0 & 1-2\varepsilon & \varepsilon\\ \varepsilon & 1-2\varepsilon & 0 & \varepsilon\\ 0 & \varepsilon & \varepsilon & 1-2\varepsilon\end{bmatrix}\begin{bmatrix} v_{00}\\ v_{01}\\ v_{10}\\ v_{11}\end{bmatrix}.\tag{15}$$

   It follows from the second and third rows of (15) that

$$\begin{aligned}\begin{bmatrix} v_{01}\\ v_{10}\end{bmatrix} &= (1-\delta)\begin{bmatrix} g_{CD}\\ g_{DC}\end{bmatrix} + \delta\begin{bmatrix} v_{10}\\ v_{01}\end{bmatrix} + \delta\varepsilon\begin{bmatrix} v_{00}+v_{11}-2v_{10}\\ v_{00}+v_{11}-2v_{01}\end{bmatrix}\\ &= (1-\delta)\begin{bmatrix} g_{CD}\\ g_{DC}\end{bmatrix} + \delta\begin{bmatrix} v_{10}\\ v_{01}\end{bmatrix} + O(\varepsilon),\end{aligned}$$

where $O(\varepsilon)$ is the term of order ε. Hence,

$$\begin{bmatrix} 1 & -\delta \\ -\delta & 1 \end{bmatrix} \begin{bmatrix} v_{01} \\ v_{10} \end{bmatrix} = (1-\delta) \begin{bmatrix} g_{CD} \\ g_{DC} \end{bmatrix} + O(\varepsilon).$$

Solving this, we get

$$\begin{bmatrix} v_{01} \\ v_{10} \end{bmatrix} = \frac{1}{1+\delta} \begin{bmatrix} 1 & \delta \\ \delta & 1 \end{bmatrix} \begin{bmatrix} g_{CD} \\ g_{DC} \end{bmatrix} + O(\varepsilon).$$

Substituting this into the first and fourth rows of (15), we obtain

$$\begin{aligned} \begin{bmatrix} v_{00} \\ v_{11} \end{bmatrix} &= (1-\delta) \begin{bmatrix} g_{CC} \\ g_{DD} \end{bmatrix} + \delta(1-2\varepsilon) \begin{bmatrix} v_{00} \\ v_{11} \end{bmatrix} + \delta\varepsilon(1+\delta) \begin{bmatrix} v_{01}+v_{10} \\ v_{01}+v_{10} \end{bmatrix} \\ &= (1-\delta) \begin{bmatrix} g_{CC} \\ g_{DD} \end{bmatrix} + \delta(1-2\varepsilon) \begin{bmatrix} v_{00} \\ v_{11} \end{bmatrix} + \delta\varepsilon \begin{bmatrix} g_{CD}+g_{DC} \\ g_{CD}+g_{DC} \end{bmatrix} + O(\varepsilon^2). \end{aligned}$$

This can be rewritten as

$$\begin{aligned} &\begin{bmatrix} 1-\delta+2\delta\varepsilon & 0 \\ 0 & 1-\delta+2\delta\varepsilon \end{bmatrix} \begin{bmatrix} v_{00} \\ v_{11} \end{bmatrix} \\ &= (1-\delta) \begin{bmatrix} g_{CC} \\ g_{DD} \end{bmatrix} + \delta\varepsilon \begin{bmatrix} g_{CD}+g_{DC} \\ g_{CD}+g_{DC} \end{bmatrix} + O(\varepsilon^2). \end{aligned}$$

Ignoring the terms involving ε^2, we hence obtain

$$\begin{bmatrix} v^{\mathrm{TFT}} \\ v^{\mathrm{STFT}} \end{bmatrix} = \begin{bmatrix} v_{00} \\ v_{11} \end{bmatrix} = \frac{1}{1-\delta+2\delta\varepsilon} \begin{bmatrix} (1-\delta)\, g_{CC} + \delta\varepsilon(g_{CD}+g_{DC}) \\ (1-\delta)\, g_{DD} + \delta\varepsilon(g_{CD}+g_{DC}) \end{bmatrix}.$$

2. Grim: The strategy has two states 0 and 1 where it chooses C and D, respectively. State transitions are synchronized between the two players when they both play Grim so that the Markov process has only two states $\Theta = \{(0,0),(1,1)\}$. We have $\omega^1 = (1,0)$ so that $v^{\mathrm{Grim}} = v_{00}$. The transition matrix is given by

$$A = \begin{bmatrix} p & 1-p \\ 0 & 1 \end{bmatrix}.$$

Ignoring the terms of order ε^2, we can write (14) as

$$\begin{bmatrix} v_{00} \\ v_{11} \end{bmatrix} = (1-\delta) \begin{bmatrix} g_{CC} \\ g_{DD} \end{bmatrix} + \delta \begin{bmatrix} 1-2\varepsilon & 2\varepsilon \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v_{00} \\ v_{11} \end{bmatrix}$$

This yields

$$v^{\mathrm{Grim}} = v_{00} = \frac{(1-\delta)g_{CC} + 2\delta\varepsilon g_{DD}}{1-\delta+2\delta\varepsilon}.$$

3. Grim2: The strategy has three states 0, 1 and 2, where it chooses C, C, and D, respectively. State transitions are synchronized between the two players so that the Markov process has three states $\Theta = \{(0,0),(1,1),(2,2)\}$. We have $\omega^1 = (1,0,0)$ so that $v^{\text{Grim2}} = v_{00}$. The transition matrix is given by

$$A = \begin{bmatrix} p & 1-p & 0 \\ p & 0 & 1-p \\ 0 & 0 & 1 \end{bmatrix}.$$

We can write (14) as

$$\begin{bmatrix} v_{00} \\ v_{11} \\ v_{22} \end{bmatrix} = (1-\delta) \begin{bmatrix} g_{CC} \\ g_{CC} \\ g_{DD} \end{bmatrix} + \delta \begin{bmatrix} (1-\varepsilon)^2 & \varepsilon(2-\varepsilon) & 0 \\ (1-\varepsilon)^2 & 0 & \varepsilon(2-\varepsilon) \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_{00} \\ v_{11} \\ v_{22} \end{bmatrix}.$$

Solving this, we obtain

$$v^{\text{Grim2}} = v_{00} = \frac{(1-\delta)\{1+\delta\varepsilon(2-\varepsilon)\}g_{CC} + 4\delta^2\varepsilon^2 g_{DD}}{(1-\delta)\{1+\delta\varepsilon(2-\varepsilon)\} + 4\delta^2\varepsilon^2}.$$

4. TF2T: The strategy has three states 0, 1 and 2, where the action choices are C, C, and D, respectively. Since state transitions are not synchronized, the Markov process has $3 \times 3 = 9$ states $\Theta = \{(0,0),\dots,(2,2)\}$. We have $\omega^1 = (1,0,\dots,0)$ so that $v^{\text{TF2T}} = v_{00}$. The transition matrix is given by

$$A = \begin{bmatrix} p & q & 0 & q & r & 0 & 0 & 0 & 0 \\ p & 0 & q & q & 0 & r & 0 & 0 & 0 \\ q & 0 & r & p & 0 & q & 0 & 0 & 0 \\ p & q & 0 & 0 & 0 & 0 & q & r & 0 \\ p & 0 & q & 0 & 0 & 0 & q & 0 & r \\ q & 0 & r & 0 & 0 & 0 & p & 0 & q \\ q & p & 0 & 0 & 0 & 0 & r & q & 0 \\ q & 0 & p & 0 & 0 & 0 & r & 0 & q \\ r & 0 & q & 0 & 0 & 0 & q & 0 & p \end{bmatrix}.$$

Using (14), we have

$$\begin{aligned} v_{11} &= (1-\delta)g_{CC} + \delta v_{00} + O(\varepsilon) \\ v_{02} &= (1-\delta)g_{CD} + \delta v_{10} + O(\varepsilon) \\ v_{20} &= (1-\delta)g_{DC} + \delta v_{01} + O(\varepsilon). \end{aligned} \qquad (16)$$

Substituting these into the recursive equations for v_{01} and v_{10}, we obtain

$$\begin{aligned}\begin{bmatrix} v_{01} \\ v_{10} \end{bmatrix} &= (1-\delta)\begin{bmatrix} g_{CC} \\ g_{CC} \end{bmatrix} + \delta(1-2\varepsilon)\begin{bmatrix} v_{00} \\ v_{00} \end{bmatrix} + \delta(1-\delta)\varepsilon\begin{bmatrix} g_{CD} \\ g_{DC} \end{bmatrix} \\ &\quad + \delta\varepsilon\begin{bmatrix} 0 & 1+\delta \\ 1+\delta & 0 \end{bmatrix}\begin{bmatrix} v_{01} \\ v_{10} \end{bmatrix} + O(\varepsilon^2),\end{aligned}$$

which yields

$$\begin{aligned}\begin{bmatrix} v_{01} \\ v_{10} \end{bmatrix} &= \frac{1-\delta}{1-\delta^2\varepsilon^2(1+\delta)^2}\begin{bmatrix} 1 & \delta\varepsilon(1+\delta) \\ \delta\varepsilon(1+\delta) & 1 \end{bmatrix}\begin{bmatrix} g_{CC}+\delta\varepsilon g_{CD} \\ g_{CC}+\delta\varepsilon g_{DC} \end{bmatrix} \\ &\quad + \frac{\delta(1-2\varepsilon)v_{00}}{1-\delta^2\varepsilon^2(1+\delta)^2}\begin{bmatrix} 1 & \delta\varepsilon(1+\delta) \\ \delta\varepsilon(1+\delta) & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + O(\varepsilon^2).\end{aligned}$$

It then follows that

$$\begin{aligned}v_{01}+v_{10} &= \frac{(1-\delta)\{1+\delta\varepsilon(1+\delta)\}}{1-\delta^2\varepsilon^2(1+\delta)^2}\{2g_{CC}+\delta\varepsilon(g_{CD}+g_{DC})\} \\ &\quad + \frac{2\delta(1-2\varepsilon)\{1+\delta\varepsilon(1+\delta)\}}{1-\delta^2\varepsilon^2(1+\delta)^2}v_{00}+O(\varepsilon^2) \\ &= \frac{(1-\delta)}{1-\delta\varepsilon(1+\delta)}\{2g_{CC}+\delta\varepsilon(g_{CD}+g_{DC})\} \\ &\quad + \frac{2\delta(1-2\varepsilon)}{1-\delta\varepsilon(1+\delta)}v_{00}+O(\varepsilon^2).\end{aligned} \tag{17}$$

On the other hand, the recursive equation for v_{00} yields

$$v_{00} = \frac{(1-\delta)g_{CC}+\delta\varepsilon(1-\varepsilon)(v_{01}+v_{10})+\delta\varepsilon^2 v_{11}}{1-\delta(1-\varepsilon)^2}. \tag{18}$$

Substituting (16) and (17) into (18) and ignoring the terms of order ε^3, we obtain

$$v^{\mathrm{TF2T}} = v_{00} = \frac{\{1+\delta(1-\delta)\varepsilon-\delta\varepsilon^2\}g_{CC}+\delta^2\varepsilon^2(g_{CD}+g_{DC})}{1+\delta(1-\delta)\varepsilon-\delta(1-2\delta)\varepsilon^2}.$$

As for the strategies AC, AD, and T6-T8, it can be readily verified that their cooperativeness is given as follows.

5. AD: $v^{\mathrm{AD}} = g_{DD}$.

6. AC: $v^{\mathrm{AC}} = g_{CC}$.
7. T8: $v^{\mathrm{T8}} = (1-\delta^7)\, g_{CC} + \delta^7\, g_{DD} + O(\varepsilon)$.
8. T7: $v^{\mathrm{T7}} = (1-\delta^6)\, g_{CC} + \delta^6\, g_{DD} + O(\varepsilon)$.
9. T6: $v^{\mathrm{T6}} = (1-\delta^5)\, g_{CC} + \delta^5\, g_{DD} + O(\varepsilon)$.

Combining the above cases, we can rank the ten strategies from the least cooperative to the most cooperative in the indefinite games as follows:

$$\begin{aligned}&\mathrm{AD} \ll \mathrm{STFT} \lll \mathrm{T6} \lll \mathrm{T7} \lll \mathrm{T8}\\ &\lll \mathrm{Grim} \ll \mathrm{TFT} \ll \mathrm{Grim2} < \mathrm{TF2T} < \mathrm{AC},\end{aligned}$$

where $\lll$, $\ll$ and $<$ represent domination in the orders of $\varepsilon^0 (=1)$, ε, and ε^2, respectively.

G.0.2 General implementation errors

When the probability $\varepsilon \in [0, \frac{1}{2}]$ of implementation errors is not necessarily small, the cooperativeness of the strategies TFT, STFT, Grim, Grim2, and TF2T can be computed numerically using (12) for the finite games and by (13) for the indefinite games, whereas the cooperativeness of AC and AD equals g_{CC} and g_{DD}, respectively, as above. Consider now the strategy Tk ($k = 6, 7, 8$). In the indefinite games, its cooperativeness can be computed as

$$v^{\mathrm{T}k} = (1-\delta)\,\frac{1-(\delta p)^{k-1}}{1-\delta p}\, g_{CC} + \delta \left\{(1-p)\frac{1-(\delta p)^{k-2}}{1-\delta p} + (\delta p)^{k-2}\right\} g_{DD}.$$

In the finite games, suppose that $t < k$ and let v_t denote the sum of stage payoffs in rounds $t, t+1, \ldots, 8$ when Tk still specifies action C in round t. We have the following recursive equations:

$$\begin{aligned}v_{k-1} &= g_{CC} + (9-k) g_{DD},\\ v_{k-2} &= g_{CC} + p v_{k-1} + (1-p)(10-k) g_{DD},\\ &\vdots\\ v_2 &= g_{CC} + p v_3 + (1-p)\cdot 6 g_{DD},\\ v_1 &= g_{CC} + p v_2 + (1-p)\cdot 7 g_{DD}.\end{aligned}$$

The cooperativeness of Tk then equals $v^{\mathrm{T}k} = \frac{v_1}{8}$.

References

Masaki Aoyagi, V Bhaskar, and Guillaume R Fréchette. The impact of monitoring in infinitely repeated games: Perfect, public, and private. *American Economic Journal: Microeconomics*, 11 (1):1–43, 2019.

Ananish Chaudhuri, Tirnud Paichayontvijit, and Alexander Smith. Belief heterogeneity and contributions decay among conditional cooperators in public goods games. *Journal of Economic Psychology*, 58:15 – 30, 2017.

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

David J Cooper and John H Kagel. Using team discussions to understand behavior in indefinitely repeated prisoner's dilemma games. *American Economic Journal: Microeconomics*, 15:114–145, 2023.

Miguel A. Costa-Gomes and Georg Weizsäcker. Stated Beliefs and Play in Normal-Form Games. *The Review of Economic Studies*, 75(3):729–762, 07 2008.

Miguel A. Costa-Gomes, Steffen Huck, and Georg Weizsäcker. Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect. *Games and Economic Behavior*, 88:298 – 309, 2014. ISSN 0899-8256.

Pedro Dal Bó and Guillaume R Fréchette. On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, 56(1):60–114, 2018.

Pedro Dal Bó and Guillaume R. Fréchette. Strategy choice in the infinitely repeated prisoners dilemma. *American Economic Review*, 2019. Forthcoming.

David N. Danz, Dietmar Fehr, and Dorothea Kübler. Information and beliefs in a repeated normal-form game. *Experimental Economics*, 15(4):622–640, Dec 2012.

Matthew Embrey, Guillaume R Fréchette, and Sevgi Yuksel. Cooperation in the finitely repeated prisoner's dilemma. *The Quarterly Journal of Economics*, 133(1):509–551, 2018.

Urs Fischbacher and Simon Gächter. Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, 100(1):541–56, March 2010.

Merrill M. Flood. Some experimental games. *Research Memorandum RM-789, The Rand Corporation, Santa Monica*, 1952.

Simon Gächter and Elke Renner. The effects of (incentivized) belief elicitation in public goods experiments. *Experimental Economics*, 13(3):364–377, September 2010.

Kyle Hyndman, Antoine Terracol, and Jonathan Vaksmann. Strategic interactions and belief formation: an experiment. *Applied Economics Letters*, 17(17):1681–1685, 2010.

Kyle Hyndman, Erkut Y. Özbay, Andrew Schotter, and Wolf Ehrblatt. Belief formation: an experiment with outside observers. *Experimental Economics*, 15(1):176–203, Mar 2012a.

Kyle Hyndman, Erkut Y. Ozbay, Andrew Schotter, and Wolf Zefev Ehrblatt. Convergence: An experimental study of teaching and learning in repeated games. *Journal of the European Economic Association*, 10(3):573–604, 2012b.

John H Kagel and Peter McGee. Team versus individual play in finitely repeated prisoner dilemma games. *American economic Journal: Microeconomics*, 8(2):253–76, 2016.

Martin G. Kocher, Peter Martinsson, Dominik Matzat, and Conny Wollbrant. The role of beliefs, trust, and risk in contributions to a public good. *Journal of Economic Psychology*, 51:236 – 244, 2015. ISSN 0167-4870.

Tibor Neugebauer, Javier Perote, Ulrich Schmidt, and Malte Loos. Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of Economic Psychology*, 30(1):52 – 60, 2009.

Yaw Nyarko and Andrew Schotter. An experimental study of belief learning using elicited beliefs. *Econometrica*, 70(3):971–1005, 2002.

Pedro Rey-Biel. Equilibrium play and best response to (stated) beliefs in normal form games. *Games and Economic Behavior*, 65(2):572 – 585, 2009. ISSN 0899-8256.

Alexander Smith. Estimating the causal effect of beliefs on contributions in repeated public good games. *Experimental Economics*, 16(3):414–425, Sep 2013.

Alexander Smith. Modeling the dynamics of contributions and beliefs in repeated public good games. *Economics Bulletin*, 35:1501–1509, 2015.