Predicting Cooperation with Learning Models: **Online Appendix**

Drew Fudenberg[1] and Gustav Karreskog Rehbinder[2]

[1]Department of Economics, MIT
[2]Department of Economics, Uppsala Universitet

November 23, 2022

A Additional Descriptions of the Data

A.1 Data Visualization

To visualize how behavior differs depending on Δ^{RD}, we put the sessions into five groups: $\delta < \delta^{SPE}$, $\delta^{SPE} < \delta < \delta^{RD}$, $0 < \Delta^{RD} < 0.15$, $0.15 < \Delta^{RD} < 0.3$, and $0.3 < \Delta^{RD}$.

Here the first 2 groups were motivated by theory, while the subdivision of the treatments with $\Delta^{RD} > 0$ was based on the data. The thresholds and relative frequencies of Δ^{RD} can be seen in figure 1.

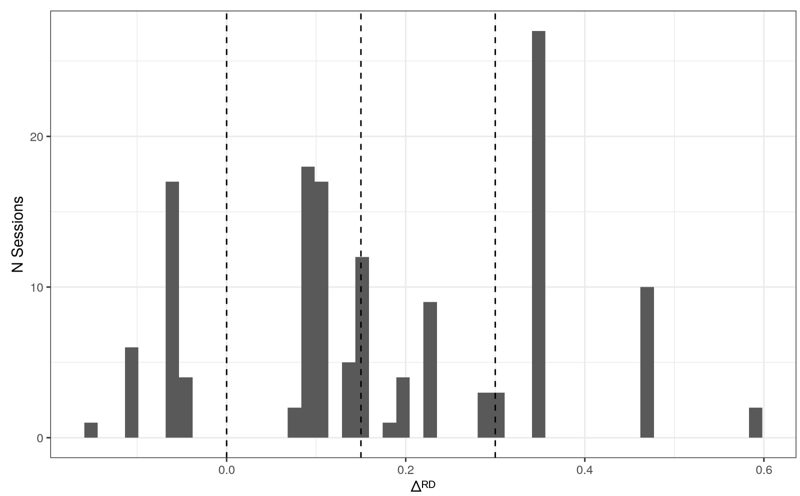

Figure 1: Distribution of Δ^{RD} for $\delta > \delta^{SPE}$

Figure 2 shows the CDF for average supergame length across the sessions.

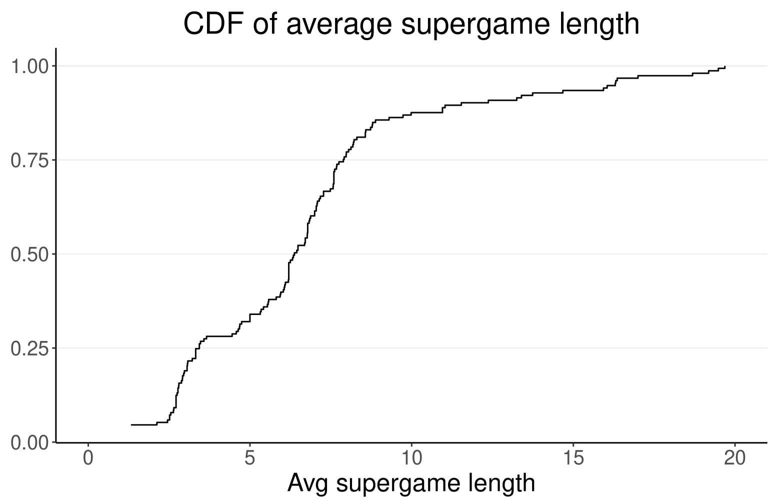

Figure 2: CDF of average realized supergame length for the different sessions in the data. The figure is truncated at an average length of 20 rounds.

Figure 3 shows the evolution of cooperation during the first 10 supergames, restricted to sessions of at least 10 supergames (134 of 161), and in figure 4 the first 20 supergames restricted to the sessions that included at least 20 supergames (93 of 161).

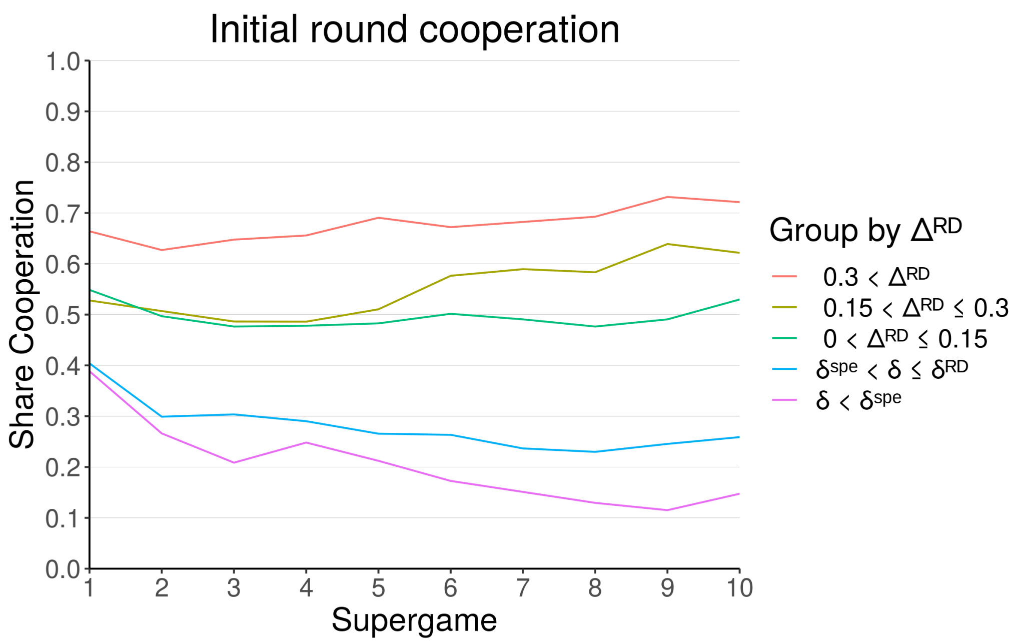

Figure 3: Cooperation in the initial round over the 10 first supergames.

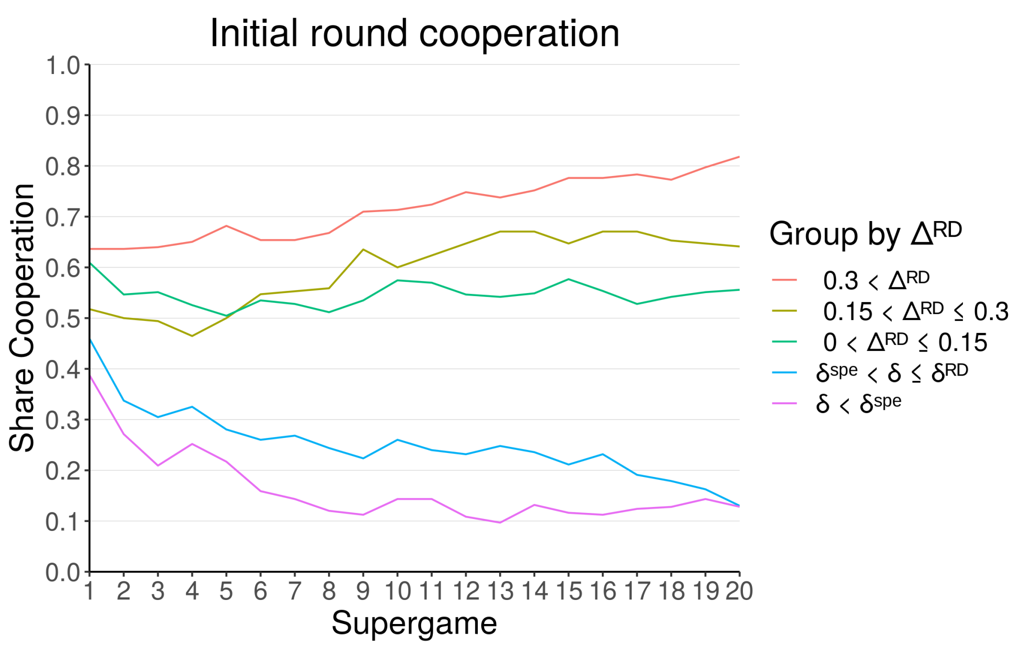

Figure 4: Cooperation in the initial round over the 20 first supergames.

In figure 5 and 6 we show the corresponding plots but for different memory-1 histories.

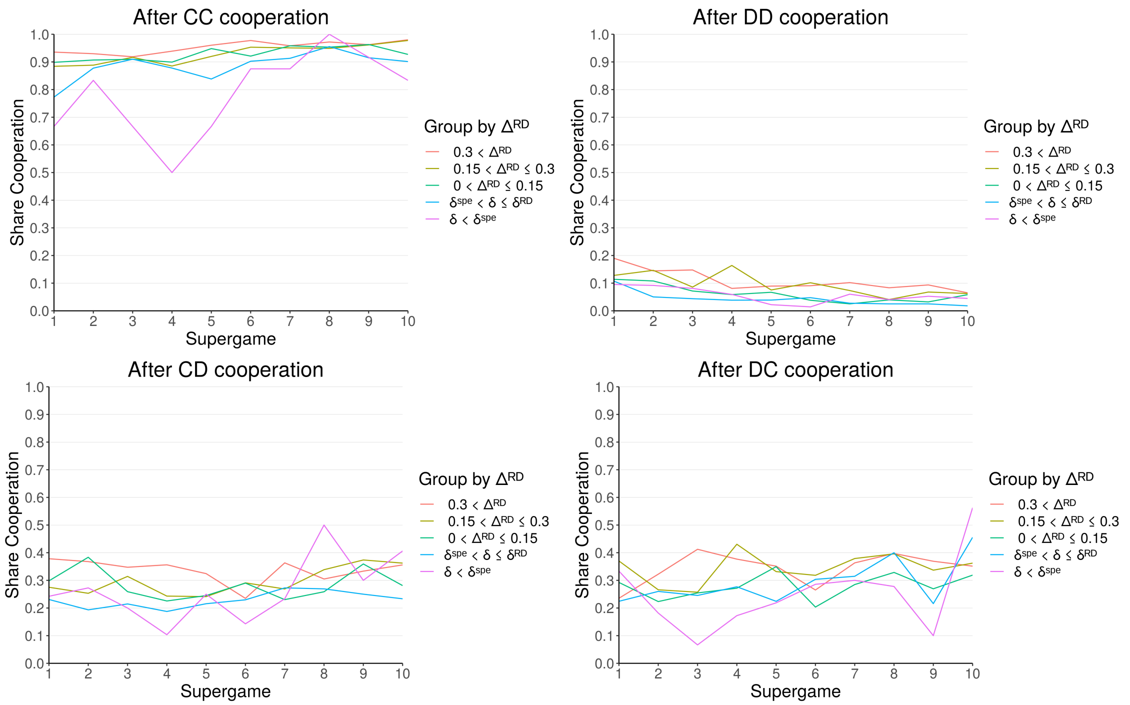

Figure 5: Average cooperation after different memory-1 histories for the first 10 supergames.

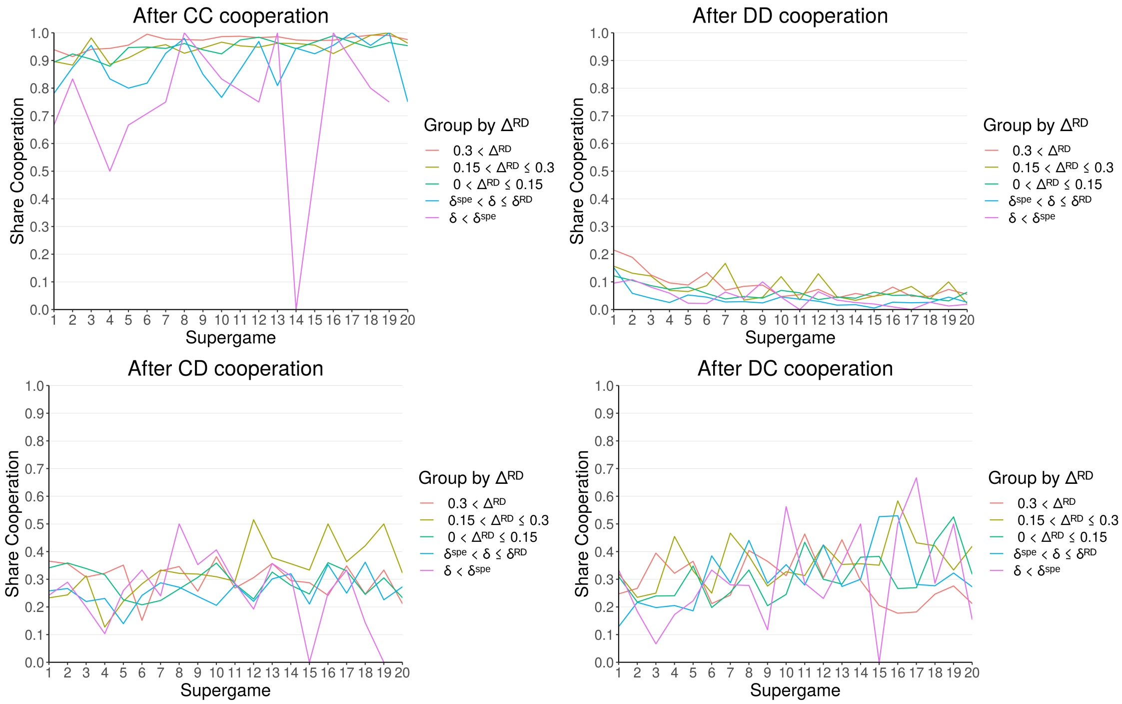

Figure 6: Average cooperation after different memory-1 histories for the first 20 supergames.

Behavior at non-initial memory-1 histories is less variable than behavior at initial histories, both between different values of Δ^{RD} and over the course of the experimental sessions. [1]

This can also be seen in Table 7, which shows average cooperation rates at each 1-period history for different rounds and Δ^{RD}. The split at $\Delta^{RD} = 0.1333$ was chosen to get as even a split as possible. We see that initial and total level of cooperation varies much more with Δ^{RD} than at any of the non-initial histories.

[1]In contrast to the initial round, different players face different distributions of the other memory-1 histories, and because these differences are not exogenous, there may be selection effects. Furthermore, some memory-1 histories are uncommon in certain treatments, e.g. there are few CC in games where the average cooperation rate is low.

All rounds with $\Delta^{RD} \leq 0.1333$

hist	Avg C	n
CC	95.0	13,228
CD	28.1	7,770
DC	32.7	7,770
DD	3.9	46,500
Initial	32.2	39,606
Total	27.7	114,874

All rounds with $\Delta^{RD} > 0.1333$

hist	Avg C	n
CC	97.1	46,208
CD	32.7	8,934
DC	33.7	8,934
DD	7.4	28,124
Initial	70.4	25,224
Total	60.1	117,424

Second round with $\Delta^{RD} \leq 0.1333$

hist	Avg C	n
CC	92.0	3,282
CD	25.5	3,968
DC	27.8	3,968
DD	6.0	10,898

Second round with $\Delta^{RD} > 0.1333$

hist	Avg C	n
CC	96.3	9,696
CD	27.5	3,573
DC	30.4	3,574
DD	15.1	2,133

Round 3 and higher with $\Delta^{RD} \leq 0.1333$

hist	Avg C	n
CC	96.0	9,946
CD	30.9	3,802
DC	37.8	3,802
DD	3.3	35,602

All rounds with $\Delta^{RD} > 0.1333$

hist	Avg C	n
CC	97.3	36,512
CD	36.1	5,361
DC	35.9	5,360
DD	6.8	25,991

Figure 7: Average behavior by round and Δ^{RD}

A.2 Importance of Initial-Round Play

If the differences in average cooperation between different treatments are driven primarily by the initial round behavior, then average cooperation after the initial round should be primarily determined by the outcome of the initial round and otherwise similar across treatments. To show this, we compare the following three regressions. The outcome variable is the average cooperation by a participant in a supergame in the rounds following the initial round, e.g., if 4 rounds were played in that particular supergame, we calculate the average cooperation by that participant in rounds 2, 3, and 4. The first regression conditions only on the outcome of the initial round. The second adds game parameters (δ, g, l, and Δ^{RD}), and the last uses only the game parameters and not the initial round.

Table 1: Rest of supergame average cooperation conditional on initial round outcome.

	(1)	(2)	(3)
initial = CD	-0.635^{***} (0.004)	-0.619^{***} (0.004)	
initial = DC	-0.638^{***} (0.004)	-0.622^{***} (0.004)	
initial = DD	-0.832^{***} (0.004)	-0.788^{***} (0.004)	
g		0.013^{***} (0.005)	-0.011^{*} (0.006)
l		-0.013^{***} (0.003)	-0.039^{***} (0.004)
δ		0.002 (0.030)	-0.028 (0.042)
Δ^{RD}		0.169^{***} (0.030)	0.813^{***} (0.041)
Constant	0.906^{***} (0.003)	0.867^{***} (0.013)	0.391^{***} (0.018)
Observations	41,080	41,080	41,080
R^2	0.582	0.586	0.187
Adjusted R^2	0.582	0.586	0.187
Note:			*p<0.1; **p<0.05; ***p<0.01

Most game parameters are statistically significant, but they explain almost no extra variance: The difference in R^2 between the model with and without game parameters is less than 0.01. In contrast, removing the outcome of the initial round lowers the R^2 to 0.0187. This is also true for second-round cooperation instead of average cooperation in the rest of the supergame.

Table 3 presents a similar regression for non-initial rounds. The outcome of the previous round is the main determinant of behavior; including Δ^{RD}, the round, or the supergame, plus interactions of any of these, only increases R^2 from 0.686 to 0.688.

Table 2: Second round average cooperation conditional on initial round outcome.

	(1)	(2)	(3)
initial = CD	-0.688^{***} (0.005)	-0.677^{***} (0.005)	
initial = DC	-0.662^{***} (0.005)	-0.651^{***} (0.005)	
initial = DD	-0.877^{***} (0.004)	-0.843^{***} (0.005)	
g		0.005 (0.005)	-0.020^{***} (0.007)
l		-0.021^{***} (0.003)	-0.048^{***} (0.005)
δ		0.143^{***} (0.034)	0.111^{**} (0.046)
Δ^{RD}		0.044 (0.033)	0.733^{***} (0.045)
Constant	0.952^{***} (0.003)	0.850^{***} (0.015)	0.341^{***} (0.020)
Observations	41,080	41,080	41,080
R^2	0.551	0.554	0.166
Adjusted R^2	0.551	0.554	0.166
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3: Non-initial round cooperation conditional on previous round history

	(1)	(2)	(3)
hist = CD	-0.660^{***} (0.002)	-0.571^{***} (0.021)	
hist = DC	-0.634^{***} (0.002)	-0.550^{***} (0.021)	
hist = DD	-0.914^{***} (0.002)	-0.770^{***} (0.017)	
Δ^{RD}		0.093^{***} (0.006)	0.822^{***} (0.009)
rd		0.093^{***} (0.017)	0.083^{***} (0.004)
round		0.016^{***} (0.003)	0.002^{***} (0.0002)
supergame		0.004^{***} (0.001)	0.003^{***} (0.0001)
Constant	0.966^{***} (0.001)	0.815^{***} (0.016)	0.181^{***} (0.003)
Interactions	N	Y	Y
Observations	167,468	167,468	167,468
R^2	0.686	0.688	0.142
Adjusted R^2	0.686	0.688	0.142
Note:			*p<0.1; **p<0.05; ***p<0.01

B The Pure Strategy Belief Learning Model

Here we outline the belief learning model from Dal Bó and Fréchette (2011) and our across-treatment generalization. Individuals are assumed to choose between TFT or AllD at the beginning of each supergame. The decision is made via a logit best reply based on the individual's beliefs about how likely a partner is to play TFT or AllD, and the implied expected payoffs.

The beliefs are tracked by the two values B^C_{is} and B^D_{is}, where i is the individual and s is the supergame. Since only two pure strategies are considered, and they prescribe different actions in the initial round of a supergame, the initial-round actions reveal the partner's strategy. The beliefs are updated according to

$$B^a_{is+1} = \theta B^a_{is} + \mathbb{1}\{a_{-i}(s) = a\}$$

where $a_{-i}(s)$ denotes the initial round action taken by the partner of individual i in supergame s, and θ captures recency in the beliefs. Given those two belief values, the belief that the partner will play TFT in supergame s is given by $B^C_{is}/(B^C_{is} + B^D_{is})$.

Let $u^\sigma(\text{TFT}), u^\sigma(\text{AllD})$ denote the expected payoff from following strategy σ if the partner is playing TFT and AllD respectively. The expected value of each choice is given by

$$U^a_{is} = \frac{B^C_{is}}{B^C_{is} + B^D_{is}} u^\sigma(TFT) + \frac{B^D_{is}}{B^C_{is} + B^D_{is}} u^\sigma(AllD) + \lambda_{is}\epsilon^a_{is}$$

where ϵ^a_{is} follows a type I extreme value distribution $\lambda_{is} = \lambda^F_i + (\phi_i)^s \lambda^V_i$. is a sensitivity parameter. This gives the following probability of subject i playing a in the initial round of supergame s, and thereafter following the according pure strategy,

$$p^a_{is} = \frac{\exp\left(\frac{1}{\lambda_{is}} U^a_{is}\right)}{\exp\left(\frac{1}{\lambda_{is}} U^C_{is}\right) + \exp\left(\frac{1}{\lambda_{is}} U^D_{is}\right)}.$$

B.1 Trembles

Here we modify the Dal Bó and Fréchette, 2011 learning model by supposing that individual i takes the prescribed action with probability $1 - \varepsilon_i$. Otherwise, the model remains the same, including using the theoretical values for the value of TFT against TFT, etc.

C Evaluation of the Procedure on Simulated Data

To test our estimation approach, we simulate the data using three different models: IRL-SG, IRL-SG with noisy individual parameters drawn from a normal distribution, and the pure strategy reinforcement learning model. The parameters for each model are taken as the average of our parameter estimates on the actual data. When we add noise to the IRL-SG, we draw each individual's parameters from a normal distribution where $\alpha \sim N(-0.268, 0.5)$, $\beta \sim N(1.291, 1)$, $\lambda \sim N(0.182, 0.1)$, $p_{CC} \sim N(0.995, 0.1)$, $p_{CD/DC} \sim N(0.355, 0.1)$ and $p_{CC} \sim N(0.012, 0.1)$. Here the means are the estimated parameters from the main analysis, and the standard deviations were set ad-hoc to what we thought were reasonable and quite large sizes. The sampled probabilities are then cut-off to be in the interval $(0, 1)$.

It is not computationally feasible to replicate the complete analysis a large number of times, as each iteration of the analysis takes a couple of days. Instead, we generate 10 different data sets for each of the three different assumptions we consider. Each session is then simulated with an actual sequence of supergame lengths, with 16 participants in each session. On each of these 10 data sets we perform a 10-fold cross-validation with the two models: IRL-SG and pure strategy reinforcement learning. We then sample prediction errors, with replacement, of the same size as the original data to get a sense of how often we would correctly infer the underlying model. On these samples, we perform the same bootstrapped pairwise test to see if one of the models is significantly better. This sampling procedure is iterated a 1000 times, and the share of correct and incorrect inferences is calculated.

	Data generating model		
Estimated model	IRL-SG	IRL-SG with noise	Pure reinf.
IRL-SG	0.0101	0.0133	0.0077
	(0.0013)	(0.0017)	(0.0008)
Pure reinf. learning	0.0156	0.0166	0.0048
	(0.0017)	(0.0019)	(0.0008)
Share correct difference	100.0%	99.2%	100.0%
Share correct significant	99.9%	70.7%	97.8%
Share incorrect significant	0.0%	0.0%	0.0%

Table 4: Comparison of the IRL-SG and the pure strategy reinforcement learning model estimated on populations simulated under different assumptions.

Table 4 shows the MSE of the IRL-SG and the pure strategy reinforcement learning model evaluated on the three different simulated data sets. The standard deviations show the variation of MSE across samples. Below are three rows that show how often the analysis draws the right conclusion. The first row indicates how often the difference in MSE goes in the correct direction, which is almost always the case. The second row shows how often the difference is both in the right direction and significant. The last row shows how often the wrong model is identified as significantly better. The wrong model is not significantly better in any of our samples. For the models without individual noise, the correct model is almost always significantly better. For the IRL-SG with individual level noise in the parameters, the IRL-SG is significantly better in 70.7% of the cases, and the pure strategy reinforcement learning model is almost never better, and never significantly so. This suggests that our estimation and evaluation approach should be able to correctly identify the underlying model, with low risk of drawing the wrong conclusion.

D One Step Ahead Prediction and Maximum Likelihood Estimation

We here consider the question of how well we can predict the next action taken by a participant given their actions so far. If each participant uses a fixed strategy or learning rule, and the relative shares in the population are known, it should be possible to accurately predict the next action a given individual will take at a given history. Thus, following the literature, we assume that there are a finite number of different "strategic types" (i.e. strategies or learning rules) used in the population, and estimate the parameters and the shares of these strategies by maximum likelihood. We then see how well the different types match the individual's behavior up to that point, and then make the corresponding prediction.

We consider the pure strategy mixture model, IRL-SG, learning with memory-1, and learning at all h, and compare their performance both to a naive benchmark that predicts the previous action taken by the individual, and to the predictions made by a gradient boosting tree.

We focus on out-of-sample predictions. This allows us to compare models of different complexities, because overly complex models may be penalized by cross-validation.

Several interesting conclusions arise from this exercise. First, in contrast to the problem of predicting the populations behavior, explicitly modeling heterogeneity does improve predictions here. Moreover, as above, learning allows us to make better out-of-sample predictions. Finally, as in past work we see no evidence of the participants using strategies of memory greater than 1.

D.1 The General Prediction Problem

Consider the complete data set of observations

$$D = \{(h_i(t), a_i(t)) | i \in I, t \in T(i)\},$$

each pair consisting of the history and the action taken for individual $i \in I$, in time period t, where $T(i)$ denotes all the rounds played by individual i, and we track the game parameters Γ_i as part of the history. The action taken $a_i(t)$ is 1 for cooperation and -1 for defection.

A predictive model is a function $m : \mathcal{H} \to [0, 1]$, where $\mathcal{H}$ is the space of all individual histories in an experimental session. This function predicts the probability that an individual with a given history cooperates. A model comes with a set of parameters θ and we write

$$m(h_i(t)|\theta) = \widehat{a}_i(t)$$

to the denote model m's predicted probability of cooperation given history $h_i(t)$.

Two different measures of predictive performance are used, prediction loss and accuracy. The prediction loss is based on the cross-entropy of the predicted probability of the taken action. For a data set $D' \subset D$, the average prediction loss is given by

$$\mathcal{L}(m|D', \theta) = \frac{-1}{|D'|} \sum_{(h_i(t), a_i(t)) \in D'} \log(m(h_i(t)|\theta)) \cdot \mathbb{1}\{a_i(t) = 1\} + \log(1 - m(h_i(t)|\theta)) \cdot \mathbb{1}\{a_i(t) = -1\}.$$

or if we simplify the notation, by letting m and θ be implicit, with

$$\mathcal{L}(D') = \frac{-1}{|D'|} \sum_{(h_{isr}, y_{isr}) \in D'} \log(\widehat{y}_{isr}) \cdot y_{isr} + \log(1 - \widehat{y}_{isr}) \cdot (1 - y_{isr}).$$

The models are always optimized with respect to the prediction loss, however, it is also interesting to look at the accuracy of the predictions. The accuracy is the share of observations where the taken action was predicted to be the most likely, i.e.

$$\begin{aligned} Acc(m|D', \theta) = \frac{1}{|D'|} \sum_{(h_i(t), a_i(t)) \in D'} \Big(&\mathbb{1}\{a_i(t) = 1\} \cdot \mathbb{1}\{m(h_i(t)|\theta) \geq 0.5\} \\ &+ \mathbb{1}\{a_i(t) = -1\} \cdot \mathbb{1}\{m(h_i(t)|\theta) < 0.5\} \Big). \end{aligned}$$

D.2 Finite Mixture models

When estimating the models, we assume that the population can be divided into different types, where individuals of the same type behave in the same way. Depending on the model, these types are parameterized in different ways. The learning models presented are the same ones used in the main text, with the difference that experience at a given supergame is calculated using the actual observed data up to that supergame, and not by simulation.

Pure Strategy Model In this model we assume that each type σ^j follows a pure strategy with a fixed mistake probability ε_j. If we let $\omega^j : \mathcal{H} \to \{0,1\}$ denote a pure strategy, e.g., Tit for Tat or Grim, a type can be described by a tuple $(\omega^j, \varepsilon_j)$. We start with an exogenous list of 11 different pure strategies, taken from the pure strategies estimated to have positive share in Fudenberg, Rand, and Dreber, 2012 [2], and estimate the share ϕ_j and mistake probability ε_j and for each such pure strategy. The mistake probabilities ε_j and the shares ϕ_j are explicitly estimated, while the 11 available pure strategies remain fixed. In the standard SFEM a common error rate ε is used, we relax this assumption in order to give the pure strategy model a better chance of performing well.

Estimating any finite mixture model gives us a set of types and their relative shares. To make a prediction of $a_i(t)$ based on $h_i(t)$, we first calculate the probability of $h_i(t)$ under the different types. For simplicity, we represent the different types with σ^j for each type j.

$$\Pr(h_i(t)|\sigma^j) = \prod_{\tau<t} \sigma^j(h_i(t))^{\mathbb{1}\{a_i(t)=1\}} \cdot \Big(1 - \sigma^j(h_i(t))\Big)^{\mathbb{1}\{a_i(t)=-1\}}.$$

Given the estimated shares ϕ the conditional probability of individual i being of type j at time t is given by

[2]While Fudenberg, Rand, and Dreber, 2012 studies interactions with exogenous noise, these 11 strategies contain those strategies often found to be used in games without noise e.g. Dal Bó and Fréchette, 2018

$$\Pr(\sigma^j|h_i(t)) = \frac{\phi^j \Pr(h_i(t)|\sigma^j)}{\sum_l \phi^l \Pr(h_i(t)|\sigma^l)}.$$

Given these estimated probabilities, the prediction of model m is given by

$$m(h_i(t)) = \sum_j \sigma^j(h_i(t)) \Pr(\sigma^j|h_i(t)).$$

D.3 Evaluating the Models

To evaluate out-of-sample performance we again use 10-fold cross-validation. Because we are now predicting individual and not aggregate play, here the partitions are at the level of individuals, so that each individual is in exactly one test set. The splits are balanced over the treatments so that roughly 10% of the participants from each treatment are in each fold. For each such partition k, we find the parameters θ_k^{train} with the smallest prediction loss on the training set,

$$\theta_k^{train} = \arg\min_{\theta \in \Theta} \mathcal{L}(m|D_k^{train}, \theta)$$

and calculate the prediction loss on the test set, $\mathcal{L}(m|D_k^{test}, \theta_k^{train})$. The prediction loss from the 10-fold cross-validation that will be reported, and used to compare the models, is given by averaging over all such splits,

$$\text{PredictionLoss}(m|D, K) = \frac{1}{K} \sum_{k=1}^{K} \mathcal{L}(m|D_k^{test}, \theta_k^{train})$$

and the accuracy is similarly given by

$$\text{Accuracy}(m|D, K) = \frac{1}{K} \sum_{k=1}^{K} Acc(m|D_k^{test}, \theta_k^{train}).$$

There is no canonical way to capture across-treatment differences in the pure strategy model. Instead, we follow the literature and make a separate estimation for each treatment. For the other models, we use a single finite mixture model for all treatments.

D.4 List of Pure Strategies Considered for Predicting the Next Action

Strategy	Description
Always Cooperate	Always play C.
Always Defect	Always play D.
Tit-for-Tat	Play C unless partner played D the last round.
Tit-for-2-Tats	Play C unless partner played D the last 2 rounds.
Tit-for-3-Tats	Play C unless partner played D the last 3 rounds.
Exploitative Tit-for-Tat	Play D in first round, then play TFT.
2-Tits-for-1-Tat	Play C unless partner played D the last round and punish for 2 rounds.
2-Tits-for-2-Tat	Play C unless partner played D the last 2 rounds and punish for 2 rounds.
Grim	Play C until either player plays D, then defect forever.
Lenient Grim 2	Play C until two consecutive rounds occur in which either player played D, then play D forever.
Lenient Grim 3	Play C until three consecutive rounds occur in which either player played D, then play D forever.

Table 5: List of pure strategies considered for predicting the next action played.

D.5 Results

Tble 6 reports the prediction errors of the different models. In this and later tables of this section, we include two additional models, *Memory-1* and *flexible memory-1*. These two models are similar to the learning models, but without learning, essentially setting $\lambda = 0$. In the memory-1 model, we simply estimate a mixture of memory-1 strategies, that are the same in all treatments. In the flexible memory-1 model, the cooperation probabilities at different memory-1 histories depend on Δ^{RD}.

Model	N types	Loss	Accuracy	Relative Accuracy
Naive		0.429	84.6%	
Pure		0.311	88.2%	23.0%
Memory-1	1	0.389	79.7%	-32.3%
	2	0.315	87.7%	20.2%
	3	0.298	87.2%	16.7%
Flexible memory-1	1	0.36	83.8%	-5.8%
	2	0.308	87.4%	17.7%
	3	0.294	87.6%	19.1%
IRL-SG	1	0.321	87.6%	18.9%
	2	0.297	88.0%	22.1%
	3	0.288	88.3%	23.9%
Initial round learning with memory-1	1	0.321	87.6%	18.9%
	2	0.292	88.5%	25.1%
	3	0.282	88.4%	24.5%
Initial round learning with flexible memory-1	1	0.319	87.6%	18.9%
	2	0.293	88.0%	22.0%
	3	0.28	88.4%	24.3%
Learning at all memory-1	1	0.322	87.5%	18.6%
	2	0.283	88.3%	23.5%
	3	0.28	88.5%	25.2%
GBT with memory-1		0.225	90.8%	40.2%
GBT with memory-3		0.222	90.9%	41.0%

Table 6: out-of-sample prediction errors.

As we see, a single type of the IRL-SG performs only slightly worse than fitting 11 different pure strategy types on each treatment. Allowing for heterogeneity in the learning model makes it slightly better than the pure strategies. The further improvement from allowing any memory-1 strategy is small, and little is gained by extending to flexible memory-1 behavior that adjusts to Δ^{RD} or extending learning to all h.

We also see that including learning increases performance compared to the models without learning. Furthermore, when learning is included, the gains from allowing type heterogeneity are much more modest, suggesting that learning itself captures much of the observed heterogeneity in the data.

D.6 Maximum Likelihoods

Our analysis of the one-step-ahead prediction problem focuses on the prediction errors of the next action taken by individuals, since this allows for straightforward comparisons between models of different complexities. However, since it is more common in the literature to consider likelihoods instead of prediction errors, we report these likelihoods here for completeness. For every history $h_i(t)$ the behavior of type j is captured by a function $\sigma^j : \mathcal{H} \to [0, 1]$ that takes a history and assigns a probability to cooperate. Each model comes with a set of parameters. We will go through the different models in the following subsections but first, present the general estimation procedure.

If we let $a_i(t) \in \{-1, 1\}$ denote the action taken by individual i at time t, the likelihood of the observed behavior for participant i if she was of type σ^j with parameters is given by

$$\Pr_i(\sigma^j|\theta_j) = \prod_{t \in T(i)} \sigma^j(h_i(t))^{\mathbb{1}\{a_i(t)=1\}}(1 - \sigma^j(h_i(t)))^{\mathbb{1}\{a_i(t)=-1\}}.$$

Let $\theta = (\theta^j)_{j=1}^J$ denote the parameters of the different types, and let $\phi \in \Delta(J)$ denote their relative share. A is then a pair $m = (\theta, \phi)$, and its likelihood is

$$\mathcal{L}\left(m|\theta, \phi, I\right) = \sum_{i \in I} \log\left(\sum_{j=1}^{J} \phi^j \Pr_i(\sigma^j|\theta_j)\right).$$

The model is then estimated by maximum likelihood.

Our main learning model only has six parameters per type, and these six parameters are the same across treatments. In comparison, the pure strategy model incorporates 11 different pure strategies, each with a different mistake probability, and these are estimated separately for each of the 28 treatments. If we were to directly compare the pure strategy model's loglikelihoods with the initial round learning model's loglikelihoods, we would be comparing a model with 736 parameters and one with 6.

To make the comparison more meaningful, here we consider the models estimated separately on each treatment as well as on the overall data, and we include BIC values to compensate for model complexity.

We consider two versions of each model (except the 11-type pure strategy model): A single type model and mixture of three types.

Table 7 reports the loglikelihoods, estimated using D.6, of the different models, estimated and evaluated on the full supergames, and table 8 reports this evaluated on the last third of the supergames in each session.

Model	N types	Estimated on	Loglikelihood	BIC
Pure	11	Each treat	-72104	149058
Memory-1	1	Each treat	-81023	163201
Memory-1	3	Each treat	-65488	134903
Memory-1	1	All treat	-89869	179800
Memory-1	3	All treat	-67775	135761
Flexible memory-1	1	All treat	-82320	164764
Flexible memory-1	3	All treat	-66914	134223
IRL-SG	1	Each treat	-72676	146507
IRL-SG	3	Each treat	-64882	133690
IRL-SG	1	All treat	-74559	149193
IRL-SG	3	All treat	-66919	134086
Learning with flexible memory-1	1	Each treat	-72341	146069
Learning with flexible memory-1	3	Each treat	-63109	130836
Learning with flexible memory-1	1	All treat	-74121	148378
Learning with flexible memory-1	3	All treat	-65242	130917
Learning at all memory-1	1	Each treat	-72068	145522
Learning at all memory-1	3	Each treat	-62699	130018
Learning at all memory-1	1	All treat	-74774	149685
Learning at all memory-1	3	All treat	-64984	130400

Table 7: Maximum likelihood log-likelihoods evaluated on the complete set of supergames.

In the literature, it is common to focus on the latter part of the experiment, under the assumption that behavior then has become more stable.

Model	N types	Estimated on	Loglikelihood	BIC
Pure	11	Each treat	-17153	38449
Memory-1	1	Each treat	-22188	45363
Memory-1	3	Each treat	-16207	35769
Memory-1	1	All treat	-25819	51695
Memory-1	3	All treat	-17400	34991
Flexible memory-1	1	All treat	-22711	45535
Flexible memory-1	3	All treat	-17062	34484
IRL-SG	1	Each treat	-19091	39168
IRL-SG	3	Each treat	-16158	35672
IRL-SG	1	All treat	-19804	39675
IRL-SG	3	All treat	-16877	33980
Learning with flexible memory-1	1	Each treat	-18916	39016
Learning with flexible memory-1	3	Each treat	-15478	34902
Learning with flexible memory-1	1	All treat	-19697	39518
Learning with flexible memory-1	3	All treat	-16449	33291
Learning at all memory-1	1	Each treat	-18775	38734
Learning at all memory-1	3	Each treat	-15004	33954
Learning at all memory-1	1	All treat	-20005	40134
Learning at all memory-1	3	All treat	-15921	32236

Table 8: Maximum likelihood log-likelihoods evaluated on the last third of the supergames.

As shown in the tables above, the maximum likelihood results on all supergames and on the last third are both consistent with the primary analysis. According to the BIC, the best model is the learning model that extends to all memory-1 histories, while the pure strategies model is one of the worst. We also see that the difference between the model with learning and semi-grim, and the possible extensions, is quite small.

We also see that we accurately capture the between-treatment variation within our models. The loglikelihood is often similar for the models estimated jointly for all treatments, with logistic functions of Δ^{RD} capturing the variation between treatments, and the ones estimated separately for each treatment. And the lowest BIC is given by such joint estimation.

E Decomposing the Prediction Errors

The analysis in this paper suggests that correctly predicting initial round behavior is of first order importance in order to predict average cooperation: Conditional on initial round outcome, there is little variation in behavior across treatments. We here compare at the prediction error in initial and non-initial rounds.

To get these prediction errors, we take the predictions of the time-path of cooperation from a single 10-fold cross-validation. We then have out-of-sample predictions for each round of each supergame in all sessions. Given these, we calculate the mean squared errors and the standard errors of the mean squared errors.

Rounds	IRL-SG	GBT	Lasso
Initial rounds	0.0261	0.0310	0.0296
Non-initial rounds	0.0329	0.0327	0.0334

Table 9: MSE for time-path predictions separated to initial and non-initial rounds.

We see that the differences between the IRL-SG model and ML-methods are larger for the initial rounds than the non-initial rounds. This suggests that our model outperforms the ML-methods because it accurately predicts initial round behavior.

To more explicitly test if there is some additional regularity the IRL-SG does not pick up, we can combine the predictions made by simulating IRL-SG and the ML-methods. We do so by adding the predictions from the IRL-SG as a feature to be used by the Lasso and GBT algorithms. We generate the predictions from the IRL-SG with a single 10-fold cross-validation, and then perform ten 10-fold cross-validations as in the mean text with those predictions as features. In Table 10 we see that the best combination (GBT + IRL-SG), has minor and non-significant improvement over just IRL-SG. This further strengthens the conclusion that IRL-SG captures the predictable regularity in the data.

Model	MSE
IRL-SG	0.0138
Lasso with IRL-SG	0.0137
GBT with IRL-SG	0.0131

Table 10: Prediction errors from combining IRL-SG and ML methods.

Another way to try to improve the predictions from the IRL-SG and model is to combine the initial-round predictions from the IRL-SG with other predictions for the rest of supergame, conditional on the initial-round outcome, similar to the exercise in Table 1 of the Online Appendix.

For a given supergame of a given session, let $\hat{y}(s)$ be the IRL-SG model's predicted average cooperation in the initial round of that supergame, the predicted share of CC outcomes in the initial round is $\hat{y}(s)^2$, the predicted share of DC outcomes $\hat{y}(s)\cdot(1-\hat{y}(s))$, etc. We can then combine these predicted likelihoods of different initial-round outcomes with predictions for the rest of supergame cooperation conditional on the initial round outcome and possibly other information about for example game parameters or current supergame.

The simplest way to do this is to use the values from model 1 in Table 1, i.e. average values conditional only on the initial round outcome. These fixed conditionals are given by $\hat{y}_{CC} = 90.7\%, \hat{y}_{CD} = 27.1\%, \hat{y}_{DC} = 26.8\%, \hat{y}_{DD} = 7.3\%$. In an attempt to improve those estimates, we can include the features used for the time-path problem, dropping those features that have to do with the round of a given supergame. These give us conditional predictions for cooperation rates in the remainder of the supergame, which we call Lasso conditionals and GBT conditionals. In the Table 11 we report the results. For computational reasons, we use only a single 10-fold cross-validation split, and don't report standard errors. There is, however, no reason to expect these to be substantially different from the ones from the main analysis.

Model	MSE
IRL-SG with fixed conditionals	0.0139
IRL-SG with Lasso conditionals	0.0137
IRL-SG with GBT conditionals	0.0135

Table 11: Prediction errors from combining initial round predictions from IRL-SG and conditional predictions for the rest of supergame cooperation.

There are two main takeaways from this table. The first is that the predictive power of our model indeed comes from its ability to predict initial round behavior. Using averages conditional on only the initial round outcome gives essentially the same MSE as using the actual model to predict the non-initial rounds as well. The second takeaway is that using more complicated predictions for conditional cooperation rates yields at best a minor improvement. This further reinforces the conclusion that IRL-SG captures most of the predictable regularity in average cooperation across treatments.

F SFEM on Simulated Data

The table below presents the results from the Strategy Frequency Estimation Method (SFEM) conducted both on the actual data and data simulated according to the IRL-SG model. When we simulate the data, all individuals in all treatments have the same parameters. We consider the 6 treatments first introduced in Dal Bó and Fréchette (2011), which have later been extensively used in other papers. In total, we have data from 1,312 individuals on these 6 treatments. As is common in the literature, the SFEM is performed on the last third of the supergames in each session.

Table 12 shows that the estimated frequencies on the actual and simulated data are similar. There are some slight deviations, in particular on the simulated data DTFT is estimated with a slightly higher frequency and AllD with a slightly lower frequency. Crucially, however, we see the same amount of heterogeneity of estimated pure strategies in both the actual and simulated data.

Δ^{RD}	Data	AllD	DTFT	TFT	Grim	Remaining
-0.316	Actual	0.698	0.256	0.042	0.0	0.004
	Simulated	0.546	0.454	0.0	0.0	0.0
-0.105	Actual	0.557	0.329	0.034	0.031	0.049
	Simulated	0.457	0.458	0.048	0.037	0.0
-0.066	Actual	0.463	0.284	0.108	0.061	0.084
	Simulated	0.372	0.39	0.121	0.078	0.039
0.105	Actual	0.411	0.105	0.146	0.265	0.073
	Simulated	0.303	0.246	0.184	0.182	0.085
0.145	Actual	0.088	0.103	0.313	0.289	0.207
	Simulated	0.141	0.213	0.338	0.151	0.157
0.355	Actual	0.122	0.034	0.328	0.36	0.156
	Simulated	0.071	0.083	0.321	0.301	0.224

Table 12: Estimated frequency of different pure strategies performed on the actual data and on data simulated with the IRL-SG model.

References

Dal Bó, P. and G. Fréchette (2011). “The Evolution of Cooperation in Infinitely Repeated Games: Experimental Evidence.” *American Economic Review*, 101, 411–29.

— (2018). “On the Determinants of Cooperation in Infinitely Repeated Games: A Survey.” *Journal of Economic Literature*, 56, 60–114.

Fudenberg, D., D. Rand, and A. Dreber (2012). “Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World.” *American Economic Review*, 102, 720–49.