Initial Play Determines Average Cooperation

Drew Fudenberg ¹  Gustav Karreskog ²

¹Department of Economics, MIT
²Department of Economics, Stockholm School of Economics

BESC, March 2, 2021
Introduction

- Repeated games often have many equilibria, especially when the players are very patient.

- Economic analyses often assume that people play the equilibrium with the most cooperation, but this is a poor fit for observed behavior in the laboratory.

- We try to develop a better understanding of how cooperation rates in experimental play of repeated games depend on their parameters.

- Treat this as a prediction problem, using data from repeated prisoner’s dilemma experiments gathered in Dal Bó and Fréchette [2018] and additional data from four papers published since then.
• In total 205,468 decisions from 2,432 participants across 143 experimental sessions in 16 papers.

• We propose a simple learning model to predict cooperation rates across treatments in the experimental play of the indefinitely repeated prisoner’s dilemma.

• The simplest version with only 6 parameters performs at least as well as more complicated models and machine learning algorithms.

• Learning only effects choices in the initial round of each supergame.

• Our results help explain past findings on the impact of risk dominance considerations.
Preliminaries and Literature Review

In the experiments we study:

- Participants played a sequence of repeated prisoner’s dilemma games with perfect monitoring. The game parameters were held fixed within each session.

- Randomly chosen partners and a random stopping time, so the discount factor $\delta$ determines the probability $(1 - \delta)$ that the current repeated game ends at the end of the current round.

- We will refer to the “rounds” of a given repeated game, and call each repeated game a new “supergame.”

- W/o loss normalize the payoff to joint cooperation to 1 and the payoff to joint defection to 0.
\[ \begin{array}{c|cc}
C & C & D \\
\hline
C & 1, 1 & -l, 1 + g \\
D & 1 + g, -l & 0, 0 \\
\end{array} \]

- “Cooperate every round” is the outcome of a subgame-perfect equilibrium if and only if

\[ 1 \geq (1 - \delta)(1 + g) \iff \delta \geq g/(1 + g) \iff \delta \geq \delta^{\text{SPE}}. \]

- Note that the loss \( l \) incurred to \((C, D)\) does not enter in to this equation!

- Applied theoretical work on repeated games often assumes that players will cooperate whenever cooperation can be supported by an equilibrium, but this hypothesis has little experimental support.

- Instead, the level of cooperation in repeated game experiments can be better predicted by measures that reflect uncertainty about the opponents’ play.
• A strategy is risk dominant in a 2x2 game if it is the best response to a 50-50 randomization.

• Grim is risk dominant in a 2x2 matrix game with the strategies Grim and Always Defect iff

\[ \delta \geq \frac{(g + l)}{(1 + g + l)} \equiv \delta^{RD}. \]

• Define

\[ \Delta^{RD} = \delta - \delta^{RD} = \delta - \frac{(g + l)}{(1 + g + l)}. \]

• Note that unlike \( \delta^{SPE} = \frac{g}{1 + g} \), \( \Delta^{RD} \) depends on \( l \) as well as \( g \), which makes sense if people aren’t sure how their partners will play.
Previous Work

- Blonski, Ockenfels, and Spagnolo [2011], Rand and Nowak [2013], and Blonski and Spagnolo [2015] show that average cooperation rates are increasing in $\Delta^{RD}$.

- Dal Bó and Fréchette [2018] shows that $\text{sign}(\Delta^{RD})$ is more correlated with high cooperation than $\text{sign}(\delta - \delta^{SPE})$.

- Dal Bó and Fréchette [2011] uses the related measure

$$
\frac{(1 - \delta)\lambda}{1 - (1 - \delta)(1 + g - l)};
$$

it is very correlated with $\Delta^{RD}$ and again reflects the role of $l$.

- [Dal Bó and Fréchette, 2018, 2011; Engle-Warnick and Slonim, 2006] find there is more cooperation in the first round increased if the realized length of the previous supergame is longer than expected.
• Past work also suggests that most participants use memory-1 strategies at least when the PD has perfect monitoring; see e.g. Dal Bó and Fréchette [2018, 2011] and Fudenberg, Rand, and Dreber [2012].

• Romero and Rosokha [2018] and Dal Bó and Fréchette [2019] elicit pure strategies from participants and confirm the finding that a small set of memory-1 strategies are enough to capture most of the strategies used.

• Breitmoser [2015] finds that strategies of the form “semi-grim” better fit play after the initial round than pure strategies do. These strategies depend only on play in the previous period, and satisfy \( \sigma_{CC} > \sigma_{CD} = \sigma_{DC} > \sigma_{DD} \): same probability of cooperating today if yesterday you cooperated and other defected or vice versa.

• None of these papers addresses our question of predicting average cooperation from game parameters.
Overview: Prediction Tasks

- We consider two prediction tasks: Predicting average cooperation in a given session, and predicting the time-path of cooperation over the course of a session.
- The first task is more fundamental; the second lets us make predictions about what cooperation levels would be if sessions were longer.
- We evaluate models based on their cross-validated mean squared error (MSE).
- When predicting the average cooperation level in a session, each session is a single data point.
- When predicting the time path of cooperation, a data point is the average (across participants) cooperation level on each round of each supergame.
Overview: Model

- We assume all individuals use memory-1 strategies.

- Learning influences how they play in the initial round of each supergame.

- Our baseline model supposes all agents use the same rule, and that play at other histories doesn’t change with experience, but we also consider extensions that relax this.

- We use simulations to make predictions based on the exogenous game parameters and realized game lengths, don’t use endogenous data like actions played and payoffs received.
Data Summary

Average cooperation rate after different memory-1 histories.

<table>
<thead>
<tr>
<th>History</th>
<th>Avg C</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>96.5%</td>
<td>52 769</td>
</tr>
<tr>
<td>CD</td>
<td>31.2%</td>
<td>15 254</td>
</tr>
<tr>
<td>DC</td>
<td>33.9%</td>
<td>15 256</td>
</tr>
<tr>
<td>DD</td>
<td>5.4%</td>
<td>66 359</td>
</tr>
<tr>
<td>Initial</td>
<td>47.2%</td>
<td>55 830</td>
</tr>
<tr>
<td>Total</td>
<td>44.2%</td>
<td>205 468</td>
</tr>
</tbody>
</table>
For data visualization, we split treatments into five groups:

- \( \delta < \delta^{SPE} \),

- \( \delta^{SPE} < \delta < \delta^{RD} \),

- \( 0 < \Delta^{RD} < 0.15 \),

- \( 0.15 < \Delta^{RD} < 0.3 \),

- \( 0.3 < \Delta^{RD} \). First 2 groups were motivated by theory, others based on the data.
• For $\Delta^{RD} > 0.15$ initial-round cooperation rates increase over the course of a session.

• For $\Delta^{RD} < 0$ cooperation rates decrease.

• For sessions where cooperation is only marginally risk dominant, $0 < \Delta^{RD} < 0.15$, initial-round cooperation rates remain roughly constant at around 50%.

• Our learning model will capture this behavior.
Evolution of Cooperation

Figure: Restricted to sessions with at least 20 supergames. (52% of the sessions)
• Now consider average cooperation in non-initial rounds.

• Three different regressions: condition on the outcome of the initial round, condition on the initial round and the game parameters, and condition on the game parameters only.

• The outcome of the initial round is highly predictive of the cooperation in the rest of the supergame, and taking into account the game parameters barely improves those predictions: $R^2$ changes from .561 to .565, while $R^2$ is only 0.172 without using the outcome of the initial round.

• This shows that if we can accurately predict play in the initial round we should be able to predict overall cooperation fairly well.
Prediciting Cooperation

- We evaluate models based on their cross-validated mean squared error (MSE).

- We use 10-fold cross-validation.

- Train/test splits are on the level of the session, so each observation is predicted using only data from other sessions.

- To estimate the standard errors of the estimated MSE, we do 10 different such 10-fold cross-validations. This results in 100 different MSE values from which we estimate the standard errors.

- The same cross-validations are used for all models, so we can also perform pairwise tests.
Learning Model

- All individuals use memory-1 strategies.

- Learning influences how they play in the initial round of each supergame.

- Behavior in non-initial rounds is the same semi-grim strategy in all treatments and individuals; this gives 3 parameters to estimate, $(\sigma_{CC}, \sigma_{CD/DC}, \sigma_{DD})$.

- The restriction to memory-1 strategies is motivated by past work, and also by our machine learning analysis.

- The learning model adds 3 more parameters.
• Initial-round cooperation $p_{i}^{initial}(s)$ depends on the game parameters and the effect of individual experience $e_i(s)$:

$$p_{i}^{initial}(s) = \frac{1}{1 + \exp(-(\alpha + \beta \cdot \Delta^{RD} + e_i(s)))}.$$

• After each supergame $s$, $e_i(s)$ is updated based on the action $a_i(s) \in \{-1, 1\}$ taken, where -1 corresponds to D and 1 to C, and the total payoff $V_i(s)$ received in supergame $s$,

$$e_i(s) = \lambda \cdot a_i(s - 1) \cdot V_i(s - 1) + e_i(s - 1),$$

where $\lambda$ determines the strength of learning, and $e_i(1) = 0$ so in the initial round of the first supergame all individuals in a session randomize in the same way.

• Cooperation or defection in the initial round is thus reinforced depending on the resulting supergame payoffs, while the direct influence of $\Delta^{RD}$ is constant across supergames.
• We used various ML algorithms (and OLS) on each prediction task.

• Which one fit best varied with the task; the slides report the best performer.

• For average cooperation we used $\Delta^{RD}$, $\delta$, $g$, $l$, total #rounds, #supergames, an indicator for $\Delta^{RD} > 0$, and some interactions.

• For predicting time paths, added current supergame, current round, and difference between previous supergame length and $1/(1 - \delta)$, and an indicator for the initial round.

• In theory, ML algorithms can figure out interactions or composite variables by themselves. With limited data, including these as features can help, so we can’t rule out that other features (or ML algorithms) could yield better predictions.
Other Learning Models

• We also consider learning models based on pure strategies, such as Tit for Tat (TFT), Allways Defect (AllD) or Grim trigger.

• Dal Bó and Fréchette [2011] estimate a belief based learning model that assumes that all participants follow either TFT or AllD. Beliefs updated after each supergame.

• Choice of whether to play TFT or AllD in the following supergame is given by a logistic function of expected payoff.

• We also consider a reinforcement learning model with TFT, AllD and Grim.

• We extend these models to make predictions across treatments by letting initial beliefs/attractons be given by functions of $\Delta^{RD}$. Also allow for a mistake probability $\varepsilon$. 
### Predicting average cooperation

<table>
<thead>
<tr>
<th>Model</th>
<th>MSE</th>
<th>SE</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant prediction</td>
<td>0.0484</td>
<td>(0.0014)</td>
<td>0</td>
</tr>
<tr>
<td>OLS on $\Delta^{RD}$</td>
<td>0.0183</td>
<td>(0.0006)</td>
<td>62.2%</td>
</tr>
<tr>
<td>Lasso</td>
<td>0.0154</td>
<td>(0.0006)</td>
<td>68.2%</td>
</tr>
<tr>
<td>Pure Strategy Reinf. Learning</td>
<td>0.0180</td>
<td>(0.0007)</td>
<td>62.8%</td>
</tr>
<tr>
<td>Learning with semi-grim</td>
<td>0.0147</td>
<td>(0.0005)</td>
<td>69.8%</td>
</tr>
</tbody>
</table>

- Our learning model beats the ML algorithms because it better predicts the influence of realized supergame lengths—in particular the model predicts that there is more cooperation when the realized supergames are long.

- If we remove information about supergame lengths from both the learning model and ML algorithms, they both have prediction error .0163.
Actual (solid line) and out of sample predicted (dashed line) initial-round cooperation by supergame for sessions of at least 20 supergames.
Interpreting $\lambda$

- The estimated learning rate ($\lambda = 0.196$) implies a strong learning effect.

- Suppose $g = l = 2$ and $\delta = 0.8$, so $\Delta^{RD} = 0$. If the first supergame an individual $i$ plays goes the expected 5 rounds, and both partners cooperate all 5 rounds, $i$’s probability of cooperation $p_{i}^{initial}(2)$ goes from 42% to 66%.

- With the estimated parameters, approx 90% of the between-treatment variance in predicted cooperation in the initial round of the last supergame in a session is driven by learning.
Average **empirical** difference between total payoff in supergames where the participant cooperated and defected in the first round. Each dot corresponds to one experimental session.
Δ^{RD} and Learning

- For Δ^{RD} < 0, defection is reinforced more strongly than cooperation in all but 1 session.

- For positive but low values of Δ^{RD}, the difference in reinforcement \( \pi(C) - \pi(D) \) is centered around 0, so cooperating and defecting are on average reinforced equally.

- This helps explain why there aren’t clear time trends in the sessions where 0 < Δ^{RD} < 0.15.
Average simulated difference between total payoff in supergames where the participant cooperated and defected. Each dot corresponds to one simulated session.
Extrapolating to Longer Experiments

• We are interested in what would happen over a longer time scale than feasible in the lab.

• Will use our learning model to make predictions about that.

• But first we test how well we can extrapolate from the first half of the sessions to the second half.

• As before each session is in either a training or a test fold but not both.

• But now we use the first halves of the training sessions to predict play in the second halves of the test sets.
<table>
<thead>
<tr>
<th>Model</th>
<th>2nd half MSE for avg C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant Prediction</td>
<td>0.066 (0.002)</td>
</tr>
<tr>
<td>OLS on $\Delta^{RD}$</td>
<td>0.029 (0.001)</td>
</tr>
<tr>
<td>Lasso</td>
<td>0.028 (0.001)</td>
</tr>
<tr>
<td>GBT</td>
<td>0.028 (0.001)</td>
</tr>
<tr>
<td>Learning with semi-grim</td>
<td>0.024 (0.001)</td>
</tr>
</tbody>
</table>

Prediction loss (MSE) from estimating on first half and evaluating on second half of the experimental sessions.

- The learning model is better at extrapolating to longer supergames than our atheoretical black-box algorithms.
• We now extrapolate long-run play in the six treatments in Dal Bó and Fréchette [2011].

• For each treatment, simulated 1000 populations with 14 participants were simulated for 10 000 supergames, with randomly drawn supergame lengths.

• Wide 90% intervals: randomness in both behavior and the differing experiences can lead to substantially different outcomes.

  – For $\Delta^{RD} < 0$, we predict and see less than 50 % cooperation.

  – For $\Delta^{RD} = 0.11$, even after 10 000 supergames the 90% interval goes from 0% to 75%, and the average is just 45%.

  – For $\Delta^{RD} = 0.36$ relatively certain prediction of high rates of cooperation.
Long-run predictions and actual behavior for six different treatments.

The solid black line corresponds to the average actual cooperation.

The dashed red line is the average of 1000 simulated populations, the dotted red lines depict the middle 90% interval.
• Wide 90% intervals for intermediate values of $\Delta^{RD}$ due to the randomness of behavior and small population size.

• Randomness comes from random initial play in a finite population: When $\Delta^{RD} = 0.11$ and population size is 100, the 90% interval is [0.01%, 60%]; with 1000 participants it is [24%, 56%].

• Randomness also comes from the realized supergame lengths: If population size is 1000 and all of the simulated supergames have their expected number of rounds, the 90% interval shrinks to [41%, 48%].

• The intervals are smaller in treatments where $\Delta^{RD}$ is more extreme in either direction.
Identifying Strategies and Predicting Actions

- We compare the performance of the finite mixture model (SFEM) of Dal Bó and Fréchette [2011], a memory-1 mixed strategy model, our two learning models, and a GBT.

- Evaluate models by how well they predict the next action taken by a participant.

- The GBT does almost as well using only last period’s actions as inputs as using the last 3 periods, which supports our assumption that people use memory-1 strategies.

- Learning alone captures most of the heterogeneity in behavior and gives better predictions than pure strategies.

- Neglecting the influence of learning might lead researchers overemphasize heterogeneity in "types."

- Skip most details today to save time.
Conclusions

• The key to predicting cooperation in a given match is the prediction of play in the initial round.

• Initial-round play depends on the game parameters and on experience in previous matches.

• A simple learning model that holds play fixed except in the initial round of each supergame, and has only 6 parameters.
• Our results show that the main way $\Delta^{RD}$ influences cooperation is through the probability of cooperation in the initial rounds of each match.

• Initial cooperation trends up or down depending on whether it is positively or negatively reinforced, which depends on $\Delta^{RD}$.

• Our model lets us predict what average cooperation rates would be with longer lab sessions (assuming the participants did not lose focus on the task).

• Many real-world settings have implementation errors or imperfect monitoring. Not yet enough experimental studies of these games to test cross-treatment predictions. Once there are it would be useful to extend our analysis to them.
New Comparative Statics Prediction

• In the lab, there is typically a tradeoff between specifying high discount factors and having participants play many supergames.

• Consider varying $\delta$ and $g$ holding $\Delta^{RD}$ fixed.

• And hold fixed the expected number of *rounds* per session, so that the expected number of supergames per session is inversely proportional to their expected length.

• Our model predicts that when $\Delta^{RD} > .15$, the treatments with more, shorter, supergames have more cooperation.

• And it predicts these treatments will have less cooperation than when $\Delta^{RD} < 0$. 
Thank you!
Estimation and Prediction OSAP

- Let $D = \{(h_i(t), a_i(t)) | i \in I, t \in T(i)\}$, be the collection of all individual histories and actions, and $m(h_i(t)) = \hat{a}_i(t)$ be model $m$’s predicted probability of cooperation.
- Model $m$ consists of a set of types $\sigma^j$, and corresponding shares in the mixture $\phi^j$.
- Allow us to calculate probability of $h_i(t)$ given $m$

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

- Bayes’ Rule give us the probability of $i$ being of type $j$ conditional on history $h_i(t)$

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

- The resulting prediction is

$$m(h_i(t)) = \sum_j \sigma^j(h_i(t)) \Pr(\sigma^j | h_i(t)).$$
• We consider two different measures of predictive performance.

• The *prediction loss* is the average cross-entropy

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

• The *accuracy* is the fraction of the predictions where the action taken was one that was predicted to be most likely.

• The *relative accuracy*

\[
\frac{\text{Accuracy}(m|D, K) - \text{Accuracy}(m^{naive}|D, K)}{1 - \text{Accuracy}(m^{naive}|D, K)},
\]

measures a model’s improvement over the naive benchmark.
## Results for OSAP

<table>
<thead>
<tr>
<th>Model</th>
<th>N</th>
<th>AllD</th>
<th>Loss</th>
<th>Accuracy</th>
<th>Rel. Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive</td>
<td></td>
<td></td>
<td>0.435</td>
<td>84.3%</td>
<td></td>
</tr>
<tr>
<td>Pure</td>
<td></td>
<td></td>
<td>0.323</td>
<td>87.6%</td>
<td>21.1%</td>
</tr>
<tr>
<td>Learning with semi-grim</td>
<td>1</td>
<td>Yes</td>
<td>0.328</td>
<td>87.2%</td>
<td>18.2%</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>Yes</td>
<td>0.308</td>
<td>87.4%</td>
<td>19.9%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>0.294</td>
<td>87.9%</td>
<td>22.7%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>Yes</td>
<td>0.288</td>
<td>87.9%</td>
<td>23.2%</td>
</tr>
<tr>
<td>Learning with memory-1</td>
<td>1</td>
<td></td>
<td>0.328</td>
<td>87.2%</td>
<td>18.2%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td></td>
<td>0.301</td>
<td>87.9%</td>
<td>22.6%</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td></td>
<td>0.289</td>
<td>88.1%</td>
<td>24.0%</td>
</tr>
<tr>
<td>GBT memory-1</td>
<td></td>
<td></td>
<td>0.225</td>
<td>90.88%</td>
<td>41.9%</td>
</tr>
<tr>
<td>GBT memory-3</td>
<td></td>
<td></td>
<td>0.223</td>
<td>90.93%</td>
<td>42.2%</td>
</tr>
</tbody>
</table>

**Table:** OSAP predictive performance
## Time-path Predictions

<table>
<thead>
<tr>
<th>Model</th>
<th>Time-path</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant prediction</td>
<td>0.071 (0.002)</td>
</tr>
<tr>
<td>OLS on $\Delta^{RD}$</td>
<td>0.039 (0.001)</td>
</tr>
<tr>
<td>Lasso</td>
<td>0.033 (0.001)</td>
</tr>
<tr>
<td>GBT:time-path</td>
<td>0.033 (0.001)</td>
</tr>
<tr>
<td>Learning with semi-grim</td>
<td>0.032 (0.001)</td>
</tr>
</tbody>
</table>

Out of sample prediction loss (MSE) of predictions for time path for different learning models.