Predicting Cooperation with Trembles

David K. Levine[1]

Abstract

The indefinitely repeated prisoner's dilemma game provides a paradigm for long-term cooperation in the face of short-term incentives for free riding. However, the extent to which players cooperate in the laboratory depends on the parameters of the game. To understand this, I take a simple direct approach motivated by theory. I hypothesize that players tremble and that they find the best perfect public equilibrium of the induced game of imperfect public information. I calibrate the probability of trembling using no data from repeated game experiments. This model makes sharp predictions and does a good job both qualitatively and quantitatively in explaining the experimental data.

[1]Department of Economics, RHUL

Acknowledgements: First version: March 25, 2024. I would like to thank Alessandra Casella, Gary Charness, Michalis Drouvelis, Rohan Dutta, Martin Dufwenberg, Ernst Fehr, John Mair, Andrea Mattozzi, Agnieszka Mensfelt, Salvatore Modica, Tom Palfrey, Klaus Ritzberger, Larry Samuelson, Arthur Schram, Francesco Squintani, Catharina Tilmans, and Myrna Wooders. The Leverhulme Trust provided financial support for which I am grateful. My knowledge of the theory and experiments discussed here comes from decades of work with Drew Fudenberg. A particular debt is owed to Pedro Dal Bo and Guillaume Frechette not only for their intellectual contributions to this work, but for providing me with their original experimental data and the data from their own metastudy. I am similarly grateful to Drew Fudenberg and Gustav Karreskog Rehbinder for making available the data from their metastudy.

1. Introduction

The infinitely repeated prisoner's dilemma game provides a paradigm for long-term cooperation in the face of short-term incentives for free riding. Such an environment is induced in the laboratory by ending the game with a role of the dice. In a repeated indefinite horizon prisoner's dilemma game, players are then rematched with new opponents to play another indefinite horizon prisoner's dilemma game. This enables us to see in the laboratory the long-term play of experienced players - the type of play that is most relevant in most applications outside the laboratory.

Theoretically, when the discount factor is sufficiently high, on account of the folk theorem of Fudenberg and Maskin (1986), any degree of cooperation is possible in equilibrium. However, game theorists have long believed, or at least hoped, that players will find their way to cooperative equilibria. This hope has proven false in experimental studies such as Dal Bo and Frechette (2011).

In order to explain how the degree of long-term cooperation depends on the specific payoff parameters of the game researchers, such as Dal Bo and Frechette (2011), Blonski, Ockenfels and Spagnolo (2011), Blonski and Spagnolo (2015) and Fudenberg and Rehbinder (2024), have employed a variety of models motivated by individual learning. Here, I take a simple direct approach, motivated by theory. I hypothesize first that players tremble, and second that, given these trembles, they achieve the greatest welfare possible in equilibrum. An indefinite horizon prisoner's dilemma game is then one of imperfect public signals and I study the welfare maximizing memory-one perfect public equilibria.

Based on earlier work on non-repeated games in Levine (2025), I assume that the chance that a player trembles uniformly is 1/3, that is, there is a 1/6 chance of playing an unintended action. I show that this calibration does a good job qualitatively and quantitatively of explaining cooperation levels in twelve repeated prisoner's dilemma treatments from ten papers.

I defer the literature review to section 5 where I explicitly compare the results of this theory to earlier work on cooperation in repeated games. There I show that performance of noise theory is similar to theories that are estimated from the repeated game data.

2. The Model

Two players $i \in \{1, 2\}$ repeatedly play a prisoner's dilemma stage game. They take actions $a^i \in \{C, D\}$, C for cooperate and D for defect. Payoffs are symmetric and given by $u^i(a)$ where a denotes the profile. Following the literature, payoffs are normalized so that $u^i(CC) = 1$ and $u^i(DD) = 0$. After each stage game, the match continues with probability $0 < \delta < 1$. Welfare in each period is given by $\omega(a) = (u^1(a) + u^2(a))/2$.

2.1. Trembles

Having described the environment, I turn to the theory. Players tremble uniformly with probability $1 > \phi > 0$. This makes it necessary to distinguish between intentions and actions. The game is then one of imperfect public information with players choosing intentions and observing public signals which are the outcomes. Following Fudenberg, Levine and Maskin (1994) the solution concept is perfect public equilibrium. Because of the nature of matching in the laboratory, equilibria must be symmetric.

With respect to intentions, let α^i denote the probability that i intends to cooperate. The probability that i actually cooperates is then given by $\tilde{\alpha}^i = (1 - \phi/2)\alpha^i + (\phi/2)(1 - \alpha^i)$. The intention profile of the two players induces a probability distribution over outcomes $y \in \{CC, CD, DC, DD\}$by

$$\pi(y|\alpha) = \begin{cases} \tilde{\alpha}^1\tilde{\alpha}^2 & y = CC \\ \tilde{\alpha}^1(1 - \tilde{\alpha}^2) & y = CD \\ (1 - \tilde{\alpha}^1)\tilde{\alpha}^2 & y = DC \\ (1 - \tilde{\alpha}^1)(1 - \tilde{\alpha}^2) & y = DD \end{cases}.$$

Utility in the one-shot game as a function of intentions is accordingly given by $U^i(\alpha) = \sum_{y \in A^2} \pi(y|\alpha)u^i(y)$.

2.2. One Period Memory

A one-period memory strategy is $\sigma^i(y)$ is a probability vector over intentions $a^i \in A$, conditional on the previous period outcome $y \in \{CC, CD, DC, DD\}$. It is used together with an initial strategy σ^i_0, also a probability vector over $a^i \in A$. This class of strategies is commonly used in studying experimental repeated games, for example, by Fudenberg and Rehbinder (2024).

For any initial y and one-period memory strategies σ^i expected utility is given recursively by

$$V^i(y|\sigma) = (1-\delta)U^i(\sigma(y)) + \delta \sum_{a\in A^2}\sum_{y'\in A^2} \sigma^i(a^i|y)\sigma^{-i}(a^{-i}|y)\pi(y'|a)V^i(y'|\sigma)$$

Define the 4×4 matrix for $y, y' \in A^2$ by

$$H(y,y'|\sigma) = \sum_{a\in A^2} \sigma^i(a^i|y)\sigma^{-i}(a^{-i}|y)\pi(y'|a)$$

then $V^i(\sigma) = (I-\delta H(\sigma))^{-1}(1-\delta)U^i$, where U^i is the 4-vector of utilities $U^i(\sigma(y))$.

2.3. Incentive Compatibility

The memory-one profile σ is *incentive compatible* if for all $i \in 1,2$, $d^i \in \{C,D\}$, $y \in \{CC,CD,DC,DD\}$

$$V^i(y|\sigma) \geq (1-\delta)U^i(d^i,\sigma^{-i}(y)) + \delta \sum_{a^{-i}\in A}\sum_{y'\in A^2} \sigma^{-i}(a^{-i}|y)\pi(y'|d^i,a^{-i})V^i(y'|\sigma).$$

A pair consisting of an initial profile and memory-one profile (σ_0,σ) is *incentive compatible* if σ is incentive compatible and for all $i\in\{1,2\}$,$d^i \in \{C,D\}$

$$\overline{V}^i(\sigma_0,\sigma) \equiv (1-\delta)U^i(\sigma_0) + \delta \sum_{a\in A^2}\sum_{y'\in A^2} \sigma_0^i(a^i)\sigma_0^{-i}(a^{-i})\pi(y'|a)V^i(y'|\sigma) \geq$$

$$(1-\delta)U^i(d^i,\sigma_0^{-i}) + \delta \sum_{a^{-i}\in A}\sum_{y'\in A^2} \sigma_0^{-i}(a^{-i})\pi(y'|d^i,a^{-i})V^i(y'|\sigma).$$

Let Σ denote the set of all symmetric incentive compatible pairs (σ_0,σ).

2.4. Welfare Maximum

Maximizing welfare subject to the incentive contraints defines $\hat{\omega} = \max_{(\sigma_0,\sigma)\in\Sigma} \overline{V}^i(\sigma_0,\sigma)$. *Noise theory* predicts that this should describe long-run average welfare in laboratory play.

2.5. Computation

Based on earlier work on non-repeated games in Levine (2025) I used the tremble parameter $\phi = 1/3$.

Maximization took place searching over a grid of $K+1$ points for each component of σ^i. In order to find mixed strategies all ϵ-equilibria were computed. Define $\Gamma = 0.33$ to be the greatest gain to defecting in any payoff matrix in Table 3.2 below. Then $\epsilon = .51\Gamma/K$ assures if there are mixed equilibria they will be found.

The first run used $K = 20$. As $\sigma^i(CC)$ was always in $\{0, 1\}$ this restriction was imposed for the second run with $K = 80$. Recall that in a Harsanyi purification individuals are restricted to pure strategies and mixing takes place only because different individuals use different pure strategies. Since the number of participants per session was less than 30, using a grid of $K = 80$ is greater precision than can be obtained with a Harsanyi purification

On an ordinary desktop computer searching grids of $K = 20$ took about thirty seconds to solve for each treatment.

3. The Data

I started with all the data in the metastudies of Dal Bo and Frechette (2018) and Fudenberg and Rehbinder (2024). As in Fudenberg and Rehbinder (2024) I excluded treatments with finite endings. As I am interested in long-run play, I used only data starting from the tenth match from experiments in which the sessions lasted more than fifteen matches. Twenty five treatments from ten papers remained: the ten papers are listed in Table 3.1 below. As indicated, some of the studies replicated treatments originally done in Dal Bo and Frechette (2011).

abbreviation	replication	citation
dal_fre	no	Dal Bo and Frechette (2011)
dre_et_al_08	no	Dreber et al (2008)
bru_kam	no	Bruttel and Kamecke (2012)
she_tar_sai	no	Sherstyuk, Tarui and Saijo (2013)
kag_sch	no	Kagel and Schley (2013)
are_cou_ran	no	Arechar, Kouchaki and Rand (2018)
dal_fre_15	yes	Dal Bo and Frechette (2015)
dal_fre_19	yes	Dal Bo and Frechette (2019)
pro_rus_sof	yes	Proto, Rustichini and Sofianos (2019)
ghi_sue	yes	Ghidoni and Suetens (2022)

Table 3.1: Studies
replication indicates if study replicates treatments in Dal Bo and Frechette (2011)

In these studies there are nine normalized payoff matrices shown below in Table 3.2.

payoffs	$u^1(CC)$	$u^1(CD)$	$u^1(DC)$	$u^1(DD)$
dal_fre:A	1.00	-1.86	3.57	0.00
dal_fre:B	1.00	-0.87	1.67	0.00
dal_fre:C	1.00	-0.57	1.09	0.00
dre_et_al_08:A	1.00	-2.00	3.00	0.00
dre_et_al_08:B	1.00	-1.00	2.00	0.00
bru_kam	1.00	-0.83	2.17	0.00
she_tar_sai	1.00	-0.25	2.00	0.00
kag_sch	1.00	-0.50	2.00	0.00
are_cou_ran	1.00	-0.33	1.33	0.00

Table 3.2: Payoff Matrices: player 1 payoffs
C for cooperate, D for defect

These payoff matrices are combined with different discount factors and give rise to twelve distinct treatments shown below in Table 3.3. The column CD-DC is the welfare from the off-diagonal where one player cooperates and one defects. I always aggregate statistics by combining all experiments which are replications of each other.

payoffs	δ	CD-DC	welfare $\overline{\omega}$	cooperation	participants	matches
dal_fre:A	0.50	0.86	0.14	0.10	196	42
dal_fre:B	0.50	0.41	0.15	0.18	80	61
dal_fre:C	0.50	0.26	0.39	0.46	306	43
dal_fre:A	0.75	0.86	0.25	0.19	146	33
dal_fre:B	0.75	0.41	0.64	0.65	68	42
dal_fre:C	0.75	0.26	0.69	0.72	260	44
dre_et_al_08:A	0.75	0.50	0.14	0.14	28	21
dre_et_al_08:B	0.75	0.50	0.46	0.46	22	27
bru_kam	0.80	0.67	0.35	0.32	36	20
she_tar_sai	0.75	0.88	0.69	0.59	56	26
kag_sch	0.75	0.75	0.58	0.53	114	30
are_cou_ran	0.13	0.50	0.19	0.19	66	21

Table 3.3: Treatments: by payoffs and δ
welfare: average welfare in the data
cooperation: fraction of times cooperation is chosen in the data
CD-DC: average payoff when one cooperates and one defects
matches: average number of matches in the data

3.1. Cooperation vs. Welfare

Welfare is predicted by the theory and has a strong economic meaning, so is a natural measure of "what happened in a treatment." By contrast earlier work has focused on cooperation rates. Of course, the reason that cooperation is interesting in the prisoner's dilemma is precisely because the dilemma is that mutual cooperation welfare dominates the dominant strategy equilibrium. That is, the dilemma it is about welfare and cooperation is correlated with welfare.

Cooperation rates do directly measure behavior. Hence, it is sensible to ask how welfare differs from the cooperation rate. First, if players play only CC or DD then the two are the same: because payoffs are normalized, welfare is equal to the cooperation rate. Second, if off-diagonal welfare is one half, as it is in Dreber et al (2008) and Arechar, Kouchaki and Rand (2018) then the two are the same. All of the off-diagonal welfares are reported in Table 3.3 above.

The cooperation rate is an imperfect measure of "what happened in a treatment" because it does not account for correlation. It views a $50-50$ alternation between CC and DD as exactly the same as a $50-50$ alternation between CD and DC. From the point of view of behavior those two alternations are quite different. In a game like dal_fre:C where off-diagonal welfare is 0.26 alternating between one cooperating and defecting is not terribly good. By contrast, in a game she_tar_sai where off-diagonal welfare is 0.88 alternating between one cooperating and one defecting is not so bad.

While from an economic point of view welfare is meaningful and so is arguably a better measure, the bottom line is that as a practical matter it does not make much difference: the two measures are highly correlated. Using the data in Table 3.3 I regressed cooperation rates on welfare. The results in Table 3.4 below show how highly correlated the two measures are.

	intercept	slope
coefficient	-0.02	1.01
se	0.03	0.07

Table 3.4: Regression of cooperation on welfare $\overline{\omega}$
R-squared is 0.95

Since the theory here predicts welfare and as it has strong economic meaning, I will focus on that.

4. Theory vs. Data

The results of this study are summarized in Table 4.1 below. It shows the twelve treatments, together with the welfare prediction of noise theory $\hat{\omega}$ versus the empirical welfare $\overline{\omega}$. Also shown for comparative purposes are the minimum and maximum welfare achievable with noise. The italics are cases where the theory says that mutual defection is the only equilibrium. There are two substantial anomalies in the data, both highlighted in bold: dal_fre:C;$\delta = 0.50$ and dre_et_al_08:B. The latter is based on little data - only 22 participants - but the former has the most data of any treatment - 306 participants - so is unlikely due to sampling error in measuring $\overline{\omega}$. These anomalies will be discussed in greater detail subsequently.

payoff matrix	δ	theory $\hat{\omega}$	data $\overline{\omega}$		min	max		participants
dal_fre:A	0.50	0.27	0.14		0.27	0.93		197
dal_fre:B	0.50	0.14	0.15		0.14	0.81		80
dal_fre:C	0.50	**0.67**	0.39		0.10	0.77		306
dal_fre:A	0.75	0.27	0.25		0.27	0.93		146
dal_fre:B	0.75	0.69	0.64		0.14	0.81		68
dal_fre:C	0.75	0.74	0.69		0.10	0.77		260
dre_et_al_08:A	0.75	0.17	0.14		0.17	0.83		28
dre_et_al_08:B	0.75	**0.67**	0.46		0.17	0.83		22
bru_kam	0.80	0.21	0.35		0.21	0.88		36
she_tar_sai	0.75	0.67	0.69		0.27	0.94		56
kag_sch	0.75	0.67	0.58		0.24	0.90		114
are_cou_ran	0.13	0.17	0.19		0.17	0.83		66

Table 4.1: Welfare: theory versus data
min/max: welfare of mutual defection/cooperation with noise
italics: theory $\hat{\omega}$ equal to minimum
bold: error $|\hat{\omega} - \overline{\omega}|$ of more than 0.16

The first take-away from the table is that supporting cooperation can be costly. Call the treatments where mutual-defection is not the only equilibrium *cooperative*. From Table 4.3 below, these equilibrium are all supported by a strategy that cooperates in the first period and whenever the outcome last period was CC. Never-the-less for she_tar_sai and kag_sch the maximum welfare that can be achieved by mutual cooperation with noise is at least 0.23 greater than the best equilibrium predicted by the theory. This means that there is a substantial cost from punishing noisy opponents.

I turn now to a more detailed assessment of the results in Table 4.1.

4.1. Qualitative Analysis.

I first examine how well the theory explains the data from a qualitative point of view. First, in the six italicized treatments where mutual defection is the only equilibrium, in two treatments welfare predictions are off by $0.13 - 0.14$ while in the remaining four cases welfare in the data is within 0.03 of what the theory predicts. The highest empirical welfare in these mutual-defection cases is 0.35.

By contrast, in the cooperative cases welfare in the data is at least 0.39 and exceeds the welfare of mutual-defection with noise by at least 0.29. That is, the treatments that are predicted to have mutual defection have less cooperation (as measured by welfare) than those that are predicted to be cooperative. The theory does a good job sorting out the cooperative treatments.

The theory also sorts well based on cooperation rates. The italicized treatments all have a predicted cooperation rate of 0.17 due to the assumed 1/6 probability of trembling onto an unintended action. The empirical cooperation rates for five of the treatments range from 0.10 to 0.19 with the sixth being equal to 0.32. By contrast, the empirical cooperation rates in the cooperative treatments are all at least 0.46.

4.2. Quantitative Analysis

I turn now to the quantitative fit of the theory. In section 5 below I will compare it to other theories, but in an absolute sense, how big is big? Should a difference between the theory and data of 0.05 as is the case in dal_fre:B;$\delta = 0.75$ be regarded as large or small? There is no unique answer to this question, but insight can be gained by looking at past practice.

Below in Table 4.2 I have gathered together data from games where Nash equilibrium uniquely predicts an equilibrium path of mutual defection. These are the finite horizon games from Dal Bo (2005) together with the two treatments in this paper, dal_fre:A;$\delta = 0.5$ and are_cou_ran. My reading of the literature is that mutual defection is regarded as a good description of what happens in those games. However, as reported in the second column of Table 4.2 cooperation rates are small but not zero.

game	cooperation rate
dal_bo:$H = 1$	0.06
dal_bo:$H = 2$	0.07
dal_bo:$H = 4$	0.18
dal_fre:A;$\delta = 0.5$	0.10
are_cou_ran	0.19

Table 4.2: Cooperation rate errors for unique Nash equilibrium paths
finite games: cooperation is average over final two matches, H is number of supergame periods
dal_fre and are_cou_ran: average is over the 10th and later matches from Table 3.3

I take this as an indication of what is an acceptable error for a theory. Below in Figure 4.1 below I compare the histogram from Table 4.2 with the histogram from Table 4.1.

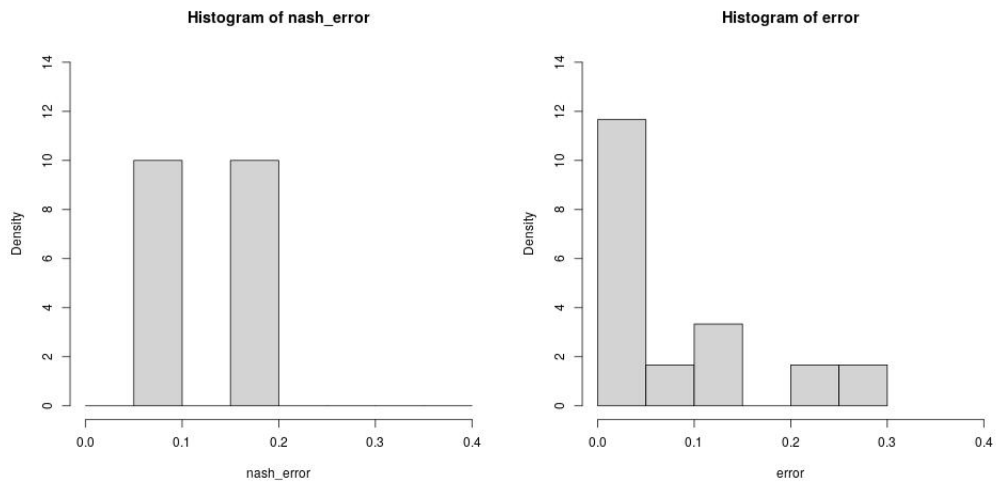

Figure 4.1: Error Histograms
left are the Nash cooperation rate errors from Table 4.2
right are the welfare errors for noise theory from Table 4.1

There are two take-aways from this. First, in general, the quantitive errors in noise theory are small compared to what has been viewed as acceptable for subgame perfect equilibrium. Second, as highlighted in Table 4.1, the two bold-faced observations, dal_fre:C;$\delta = 0.50$ and dre_et_al_08:B, are indeed anomolous, with errors larger than what would ordinarily be considered “acceptable.”

Below in Figure 4.2 I show a graphical view of the theory against the data. The horizontal axis is the theoretical prediction of welfare and the vertical axis the empirical welfare.

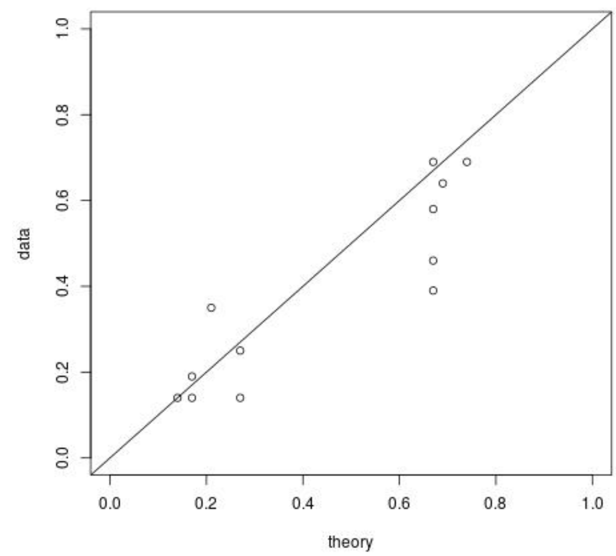

Figure 4.2: Scatter Plot of Welfare: Data $\overline{\omega}$ Against Theory $\hat{\omega}$

If the theory was perfectly correct then the dots representing observations in Figure 4.2 would lie on the 45° line. With the exception of the two anomolous outliers dal_fre:C;$\delta = 0.50$ and dre_et_al_08:B they are generally close.

4.3. Sampling Error vs. Specification Error

The difference between the theory and the data as reported in Table 4.1 has two sources: sampling error and specification error. In the case of Nash equilibrium the theory is deterministic: there is no sampling error, and so the error is entirely specification error. By contrast noise theory is stochastic due both to trembling and the use of mixed strategies. Standard errors can be computed by Monte Carlo simulation, but there is an easy upper bound that is fairly accurate.

Welfare is between zero and one. The greatest variance possible for any such random variable is 1/4 attained by the binomial with mean equal to 1/2. Because trembles and mixing are independent between periods, each match is an independent draw. Hence, for each treatment, the number of observations is half the number of participants times the average number of matches after the ninth. These can be computed from Table 3.3 above. The least number of observations is dre_et_al_08:A with 168, the greatest dal_fre:C with 4862. These give rise to upper bounds on the standard error of 0.039 and 0.007 respectively. Even the largest of these upper bounds, 0.039, is inadequate to explain the anomolous treatments. The smallest of these divergences of data from theory is 0.21 in dre_et_al_08:B. This is more than five times the uppermost possible standard error.

This is all to say that it is specification error not sampling error that is important for the anomalies.

4.4. Best Equilibrium Strategies

Below in Table 4.3 are shown strategies that support the maximum welfare reported in Table 4.1 above.

payoffs	δ	initial	CC	CD	DC	DD
dal_fre:A	0.50	0.00	0.00	0.00	0.00	0.00
dal_fre:B	0.50	0.00	0.00	0.00	0.00	0.00
dal_fre:C	0.50	1.00	1.00	0.00	0.00	1.00
dal_fre:A	0.75	0.00	0.00	0.00	0.00	0.00
dal_fre:B	0.75	1.00	1.00	0.30	1.00	0.08
dal_fre:C	0.75	1.00	1.00	0.91	1.00	0.26
dre_et_al_08:A	0.75	0.00	0.00	0.00	0.00	0.00
dre_et_al_08:B	0.75	1.00	1.00	0.01	1.00	0.00
bru_kam	0.80	0.00	0.00	0.00	0.00	0.00
she_tar_sai	0.75	1.00	1.00	0.00	0.28	0.01
kag_sch	0.75	1.00	1.00	0.00	0.53	0.01
are_cou_ran	0.13	0.00	0.00	0.00	0.00	0.00

Table 4.3: Best Equilibrium Strategies
player 1 conditional probabilities of cooperating
CC,CD,DC,DD are last the period outcome

The strategy $(0, 0, 0, 0, 0)$ is mutual-defection and is the only available equilibrium in the italicized treatments. The strategy $(1, 1, 0, 0, 0)$ is the grim-trigger strategy, which is never used. Mixed strategies are rare. All of the cooperative strategies cooperate initially and when the state is CC. They have the general form have the form $(1, 1, M, H, L)$ where $M < H$. This can be described as grim-trigger with forgiveness (L) and apology ($H > M$). When one player defects and the other cooperates the defecting player "apologizes" by cooperating with a higher probability than the other.

Each of the memory-one strategy give rise to a 2×2 game with payoffs

$$(1-\delta)U^i(a) + \delta \sum_{y \in A} \pi(y|a) V^i(y|\sigma).$$

The best equilibrium memory-one strategies give rise to three types of games. The first are prisoner's dilemma games in which it is dominant to defect. Here the only equilibrium of the supergame is mutual defection. The second are coordination games in which CC and DD are equilibria with a mixed one in between (which, however, is never optimal to use). Here the equilibrium can be chosen in a state dependent way to provide incentives for cooperation. The third are indifference games where players are indifferent between cooperating and defecting regardless of what the opponent does. In these latter games mixing sometimes takes place, and, in an off-diagonal states $y \in \{CD, DC\}$, the mixing need not be symmetric and generally is not.

5. Learning Theoretic Approach

In the last twenty years there has been extensive research into when cooperation emerges in repeated prisoner's dilemma games in the laboratory. There are four main theories: subgame perfect equilibrium, risk dominance, critical discount factor, and reinforcement learning. With the exception of reinforcement learning, none explicitly account for the presence of trembling, although there is a different literature that examines perfect public equilibrium where the experimenter introduces noise in observation: a good review of the entire literature can be found in Dal Bo and Frechette (2018).

The first approach to the issue of cooperation is that of subgame perfect equilibrium studied in Dal Bo (2005). He shows that finite games have low cooperation rates, while indefinitely repeated games with cooperative equilibria have high cooperation rates. Subsequent research has borne out the idea that if mutual defection is the only equilibrium then cooperation rates are low. This can be seen in Table 4.2 above, although, in fact, the cooperation rates range from 0.10 to 0.19, while the theory predicts 0.00.

Subsequent research by Dal Bo and Frechette (2011) (the results of which are included here) showed that the reverse need not be true. Even if cooperation is an equilibrium cooperation rates can be low. This can be seen in Table 4.1 above, while dal_fre:B;$\delta = 0.50$, dal_fre:A;$\delta = 0.75$ and dre_et_al_08:A all admit cooperative equilibria (without noise), all have low cooperation rates - welfare is no higher than 0.25 and cooperation rates no more than 0.19.

The anomaly of games in which there are cooperative equilibria, yet cooperation does not emerge, led to learning theoretic considerations. Theories of learning and evolution such as Kandori, Mailath and Rob (1993) and Young (1993) suggest that in a 2×2 game risk dominance plays a key role in determining which equilibrium will emerge when there are multiple equilibria. To make this operational in an indefinitely repeated game where there are 32 memory-one strategies the literature has focused on just two such strategies: mutual-defection and grim-trigger. Before discussing how this works and the findings, I want to point out that this choice incorporates the basic hypothesis of this paper - that players are trying to achieve a good outcome. Specifically, if the opponent is thought to randomize equally over all 32 memory-one strategies, for all the treatments here, always-defect is a best response. By limiting attention to grim-trigger, the question becomes: if players are trying to find a good solution using grim-trigger, will they be able to learn to do so? Note that this methodology can be extended to other games by considering the welfare worst and best equilibria.

Dal Bo and Frechette (2011) examine risk dominance in the first six games of Table 3.3. They give mixed reviews. Their assessment is based on whether cooperation rose or fell over time, and they do not have a clear theory of what constitutes high and low cooperation - standard theories being limited to 0.00 and 1.00. Moreover, rising and falling is a difficult criteria because initial match cooperation ranges from 28% to 56% and doing well in the early matches may mean doing less well in the later matches.

Before looking at the data, I first introduce two key theoretical constructs connected with risk dominance. The first is the basin: the greatest proportion of players who can be playing always defect for grim-trigger to be a best response. This can be computed as

$$\beta = \max\left\{0, \frac{(1-\delta)(u^1(DC)-1)-\delta}{(1-\delta)(u^1(DC)+u^1(CD)-1)-\delta}\right\}.$$

Note that the right-hand expression will be negative when grim-trigger is not an equilibrium at all, in which case the size of the basin is 0. If this is greater than 0.50 then grim-trigger is risk dominant, if it is smaller then mutual-defection is risk dominant.

A related idea is that of the critical discount factor. This was originally introduced by Blonski, Ockenfels and Spagnolo (2011) based on axiomatic ideas, the crucial one being that for any given payoff matrix there is a critical discount factor above which there will be cooperation and below which there will not. They have additional less persuasive axioms concerning additive separability and so forth, but the bottom line is that they compute

$$\Delta = \delta - \frac{u^1(DC) - u^1(CD) - 1}{u^1(DC) - u^1(CD)}$$

where the negative term is the critical discount factor. In other words, there should be cooperation if Δ is positive and should not be if it is negative. Blonski and Spagnolo (2015) subsequently pointed out that the critical discount factor is the discount factor that makes players indifferent between grim-trigger and always-defect when the population is $50 - 50$ between the two, so this has a similar flavor to the basin.

Finally, in a recent paper, Fudenberg and Rehbinder (2024) consider a six parameter reinforcement learning model using Δ as an explanatory variable. They undertake the ambitious program of predicting play in every round of every match. The details of that model are well described in the paper. For the purposes here I generated artificial data using the parameters from their Table 4. For each treatment in Table 3.3 I conducted a Monte Carlo simulation with 1000 trials. I used the greatest even number less than the mean number of participants per session (14) and the mean number of matches for the treatment from Table 3.3. Averaging welfare over the tenth and subsequent matches generated predictions $\hat{\hat{\omega}}$ of welfare for each treatment.

I turn now to how β, Δ, $\hat{\omega}$ and $\hat{\hat{\omega}}$ fare with respect to the data.

5.1. Sorting Cooperation Levels

As indicated those treatments predicted to have mutual defection by $\hat{\omega}$ have strictly lower welfare than those that are cooperative. How do β,Δ do on this score? The β theory pred and Δ theory when $\Delta < 0$. These generate the same predictions about mutual defection as $\hat{\omega}$ except that $\hat{\omega}$ predicts mutual defection for bru_kam and the other two do not. As welfare in bru_kam is higher than any of the other italicized treatments, this means that while β, Δ disagree with $\hat{\omega}$ about where the dividing line is, they also correctly yield lower welfare in the treatments they predict to have mutual-defection than those they predict to have cooperation.

5.2. Bifurcation

As just indicated, all four theories do a good job sorting the data into high and low welfare treatments. It is a characteristic of the static theories - $\hat{\omega}, \beta, \Delta$ - that there is a bifurcation. That is, as the parameters are gradually changed from a game in which the only equilibrium is mutual defection, welfare should jump when cooperative equilibria emerge. In other words, the prediction is that a histogram of welfare should be U-shaped, exhibiting low and high welfare, but nothing in between. Below in Figure 5.1 is the histogram of empirical welfare.

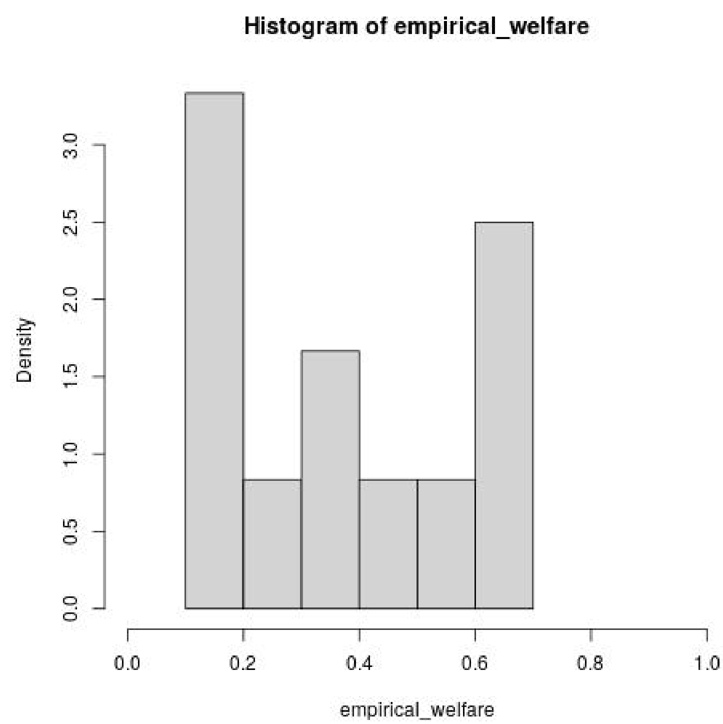

Figure 5.1: Histogram of Empirical Welfare

The first thing to observe is that the histogram is compressed: the low welfare treatments give welfare well greater than zero and the high welfare treatments are even further away from one. Consequently the β, Δ theories in their purest form cannot contend with the data quantitatively. They predict welfare of 0 when they predict mutual defection, while it ranges from 0.10 to 0.25 in the data. They predict grim-trigger with welfare of 1.00 in the remaining cases, while empirical welfare in those cases ranges from 0.35 to 0.69.

The second observation is that while the histogram is generally U-shaped, there is also a middle-range spike between 0.35 and 0.46. These correspond to the two anomolous treatments in noise theory - dal_fre:C;$\delta = 0.5$ and dre_et_al_08:B, plus bru_kam. This poses a challenge for all the theories, about which, more below.

5.3. Monotonicity

According to the relevant theories empirical welfare should be increasing in $\hat{\omega}, \beta, \Delta, \hat{\hat{\omega}}$. If empirical welfare is increasing in the theoretical constructs, then the theoretical constructs should be increasing in the data. I examine this below in Figure 5.2 plotting the welfare data on the horizontal axis against $\hat{\omega}$ and $\hat{\hat{\omega}}$ and values of β and Δ normalized so they will conveniently fit on the graph.

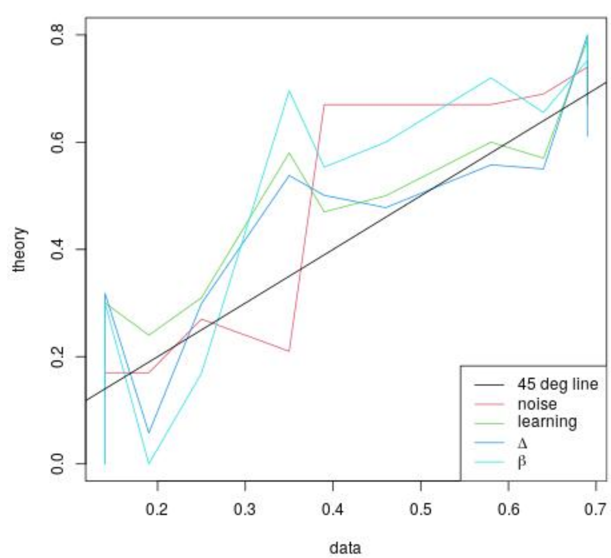

Figure 5.2: Theory Against Data

As can be seen, while all curves are generally upward sloping, none are increasing. Note also that, consistent with discussion of bifurcation above, all of the theories are quite challenged in the middle range between 0.35 and 0.46.

To investigate further, observe that two theories, $\hat{\omega}$ and $\hat{\hat{\omega}}$ make specific predictions of welfare. Unless we take the discredited theory of 0 for mutual-defection and 1 for cooperation, the theories β, Δ do not. However, we can regress the data on the theoretical constructs to get reasonable predictions.[2] The results are shown in Table 5.1 below.

	parameters	mean abs	R^2	intercept (se)	slope (se)
$\hat{\omega}$	0	0.09	0.75	0.00	1.00
β	2	0.09	0.77	0.09 (0.06)	0.58 (0.10)
Δ	2	0.09	0.75	0.36 (0.03)	0.98 (0.18)
$\hat{\hat{\omega}}$	6	0.10	0.75	0.00	1.00

Table 5.1: Predictors of Welfare ϖ
parameters: number of parameters estimated from repeated game data
mean abs: mean absolute error
R^2 not adjusted

The bottom line is that there is little difference in accuracy between the theories. All have similar mean square error (R^2) and mean absolute error. Put another way: we might think of $\beta, \Delta, \hat{\hat{\omega}}$ as the best we can do looking at the data. Perhaps imperfection is unavoidable because of difference in subject populations, laboratory conditions, locations, and other factors not accounted for by the theories. Noise theory, which does not look at the data at all, does just as well. However, as I will indicate below, it is possible to do better.

5.4. Predicted Values

As indicated $\hat{\omega}$ and $\hat{\hat{\omega}}$ make specific predictions about welfare in each treatment and for β, Δ we can use the fitted values from the regression in Table 5.1 as predictions. Below in Table 5.2 I report the predictions of the four theories along with the data.

[2] As the relationship might not be linear I also ran regressions with quadratic terms, but these added little.

payoff matrix	δ	$\hat{\omega}$	β	Δ	$\hat{\hat{\omega}}$	data $\overline{\omega}$
dal_fre:A	0.50	0.27	0.09	0.05	0.12	0.14
dal_fre:B	0.50	0.14	0.25	0.25	0.22	0.15
dal_fre:C	0.50	**0.67**	0.45	0.46	0.47	0.39
dal_fre:A	0.75	0.27	0.20	0.30	0.31	0.25
dal_fre:B	0.75	0.69	0.52	0.50	0.57	0.64
dal_fre:C	0.75	0.74	0.58	0.70	0.79	0.69
dre_et_al_08:A	0.75	0.17	0.29	**0.31**	0.30	0.14
dre_et_al_08:B	0.75	**0.67**	0.48	0.44	0.50	0.46
bru_kam	0.80	0.21	**0.54**	0.49	**0.58**	0.35
she_tar_sai	0.75	0.67	0.61	0.55	0.67	0.69
kag_sch	0.75	0.67	0.56	0.51	0.60	0.58
are_cou_ran	0.13	0.17	0.09	0.10	0.24	0.19

Table 5.2: Welfare Predictions vs. Data
italics: noise $\hat{\omega}$ equal to minimum
bold: error $|\hat{\omega} - \overline{\omega}|$ or $|\hat{\hat{\omega}} - \overline{\omega}|$ of more than 0.16

The takeaway is this. Noise theory has more outliers, with two major anomalies, while each of the other theories has only a single anomaly. For the five low welfare treatments, with the exception of dal_fre:A, noise theory does better than any of the other theories. Indeed: what 5.2 highlights is that noise theory does very well most of the time, but fails badly twice - for dal_fre:C;$\delta = 0.50$ and dre_et_al_08:B. This is suggestive that there might be some missing ingredient that explains what is different about those treatments.

5.5. Noise and Learning as Complements

Noise and learning theory need not be in competition. Step back for a moment. Whatever participants are doing in the laboratory they are not searching grids of hundreds of thousands of points to find a best solution to a problem. Instead they are in a noisy environment trying to learn by trial and error. Indeed, it seems likely that they understand the need to punish bad behavior and reward good behavior and that they seek to do so in a way that leads to a good social outcome. Remarkably, in ten out of twelve treatments they do nearly as well as if they had solved the problem on the grid.

What about the other two treatments, the anomalies? It may well be that some solutions are harder to find by trial and error than others, and here it seems as if learning theory should be able to provide a guide. Unlike noise theory, Table 5.1

shows that the learning theories provide a relatively accurate description of welfare in the two anomalous cases. Perhaps for an intermediate range of learning values, indicating that it is not so obvious whether cooperation can be achieved, learning is hard, and so fails to find the best solution? Indeed, Table 5.1 shows exactly that. For β predictions the two anomolous treatments have predicted values between 0.45 and 0.48,while for all the other treatments predicted values lie outside this range. Similarly for Δ the anomolous treatments have predicted values between 0.44 and 0.46 while all the other treatments have predicted values outside this range. Table 5.3 summarizes the situation and also maps the predicted values (an increasing linear transformation of the underlying values) from Table 5.1 back to the underlying values of β and Δ.

	β predict	β	Δ predict	Δ
highest below low	0.29	0.33	0.30	-0.06
low anomaly	0.45	0.61	0.44	0.08
high anomaly	0.48	0.67	0.46	0.10
lowest above high	0.52	0.72	0.50	0.14

Table 5.3: β, Δ ranges
predicted welfare and underlying values of β, Δ
low and high anomaly: the lowest and highest values for anomolous treatments
highest below low: highest value of non-anomolous treatment below high anomaly
lowest above high: lowest value of non-anomolous treatment above low anomaly

From this we may conclude that if β is between 0.61 and 0.71 or Δ between 0.08 and 0.10 there is an anomaly. These range are not narrowly pinned down in the data: they could be as large as 0.34 to 0.71 and -0.05 to 0.13.

If a theory does not work all the time it is useful to be able to predict when it will and will not work. Here learning theory enables the prediction of the anomolous cases - and both β and Δ are computed directly from the experimental instructions and involve no estimation. As a crude approximation we might conjecture that in the anomolous range of β, Δ participants are only able to make it halfway from the minimum welfare with noise to the highest equilibrium welfare. In both cases this predicts a welfare of 0.42 as against the empirical welfare of 0.39 and 0.46, so is reasonably accurate.

Note that bru_kam poses something of a puzzle. According to noise theory there is no better equilibrium than mutual-defection, while the learning theories indicate that learning is easier than in the anomolous range - and so should lead to a cooperative equilibrium. In fact the data lies about midway in between $\hat{\omega}$ and the predictions of the learning theories, although a bit closer to noise theory. Perhaps when noise predicts mutual-defection but learning theory is above the anomolous range it would make sense to make a prediction in between. A related fact about the bru_kam treatment is that it sensitive to a small change in the amount of noise. As discussed in Section 6 below there is a bifurcation point in $\phi \in (0.332, 0.333)$ and when ϕ is reduced to 0.332, theoretical welfare jumps from 0.21 to 0.69, surrounding the empirical welfare of 0.35. As there is no clear cut recipe for what to do here, I have not pursued this.

Returning to the hybrid theory, how does it fare? In the anomolous range we take the average of the minimum with noise and the highest equilibrium welfare, and otherwise we predict the highest equilibrium welfare. This involves, in some sense, the estimation of three parameters: the bottom and top of the anomolous range and the weight on the minimum welfare and equilibrium welfare. This hybrid theory results in a mean absolute error of 0.05 and an unadjusted R^2 of 0.91. That is to say: this hybrid theory does extremely well and much better than any of the theories in Table 5.1.

5.6. Application to Other Classes of Games

A good theory ought to be able to predict behavior in more than one class of games. In previous work, in Levine (2025), I studied non-repeated games using the model described here augmented to allow for social preferences that are irrelevant in the repeated prisoner's dilemma setting. In that paper there was a single significant anomaly. This occured in an experiment by Nikiforakis and Normann (2008). They studied a four player public goods game with punishment, varying the efficacy of punishment from 1 to 4. While contributions in the treatments from 2 to 4 are well predicted by the theory, in the treatment with an efficacy of 1 contributions to the public good fall well below what the theory predicts. Could the hybrid theory described above also be a possible explanation for this anomoly?

As indicated, the basin β can be defined in any game with multiple equilibria. In the case of the public goods game with punishment, the worst Nash equilibrium is no contribution and the best is the maximum contribution by all four players combined with maximum off-the-path punishment for failing to contribute the full amount. In the earlier working paper version of this paper Levine (2024) (available online), I computed the basins for the public good game treatments. I found that the basin for treatment 1, the anomolous treatment, is $\beta = 0.57$. For the non-anomolous treatment with efficacy 2, the basin is β =0.71. The higher efficacy treatments $3, 4$ have even larger basins. Above in Table 5.3 the lower bound for anomolous outcomes is in $[0.34, 0.61]$ and the upper bound in $[0.68, 0.71]$. Hence the public good game treatment 1, with $\beta = 0.57$, is plausibly in the anomolous range. By contrast the the treatment 2 , with $\beta = 0.71$, is plausibly in the non-anomolous range.

Another way to say this is to take the lower bound for anomalies to be in $[0.34, 0.56]$, and the upper bound to be in $[0.68, 0.70]$. Then β explains the anomalies and lack of anomalies in both the games studied here and in the public good game studied in Levine (2025).

6. Robustness

The noise theory is governed by a single parameter ϕ. I have not tried to estimate ϕ. The strategy of estimation tries to fit parameters to particular treatments, experiments, or classes of treatments. I believe it is more useful to look for parameters that are stable across similar populations and stakes - in this case college students playing for normal stakes, and $\phi = 1/3$, in fact, does the job. Never-the-less it is useful to understand how the theory changes with ϕ.

There are two types of changes as ϕ changes. The first is a gradual change in which strategies do not change much, but the change in noise leads to a change in welfare. The second is bifurcation where strategies change abruptly as the availability of cooperative equilibria changes. The latter is substantially more important for the quality of the fit.

For gradual change, as noise is reduced, welfare decreases for the mutual-defection treatments and increases for the cooperative treatments. In five out of the six mutual-defection treatments predicted welfare is either highly accurate or understates true welfare. In the cooperative treatments predicted welfare is generally too high already. This means that reducing noise will reduce the quality of the fit.

The converse, as noise is increased, is not true. In five out of the six mutual-defection treatments predicted welfare is either highly accurate or overstates true welfare. Hence increasing noise continues to hurt the fit for the mutual-defection treatments. By contrast it improves the fit for the cooperative treatments. Worse for mutual-defection treatments, better for cooperative treatments.

Bifurcations are reported as ranges in Table 6.1 below. The top of the range has mutual-defection as the only equilibrium. The bottom of the range has a best equilibrium that is cooperative. Hence, the bifurcation takes place in the range shown. For those treatments with bifurcation points nearer $\phi = 1/3$ the ranges were computed with a higher degree of precision. As expected, dal_fre_A and are_cou_ran have only mutual-defection as an equilibrium regardless of noise.

payoff matrix	δ	bifurcation range for ϕ
dal_fre:A	0.50	none
dal_fre:B	0.50	$(0.1, 0.2)$
dal_fre:C	0.50	$(0.5, 0.6)$
dal_fre:A	0.75	$(0.0, 0.1)$
dal_fre:B	0.75	$(0.35, 0.4)$
dal_fre:C	0.75	$(0.6, 0.7)$
dre_et_al_08:A	0.75	$(0.1, 0.2)$
dre_et_al_08:B	0.75	$(0.339, 0.340)$
bru_kam	0.80	$(0.332, 0.333)$
she_tar_sai	0.75	$(0.338, 0.339)$
kag_sch	0.75	$(0.338, 0.339)$
are_cou_ran	0.13	none

Table 6.1: Bifurcation Ranges
for ϕ at the top mutual-defection is the only equilibrium
for ϕ at the bottom the best equilibrium is cooperative
none indicates for all ϕ mutual-defection is the only equilibrium

Most of the treatments are quite robust, with their bifurcation points lying well away from $\phi = 1/3$. Four of the treatments, however, dre_et_al_08:B, bru_kam, she_tar_sai and kag_sch, have bifurcations close to $\phi = 1/3$.

Getting the classification wrong is extremely costly in both mean absolute and mean square error. For example, as noise is increased, the she_tar_sai and kag_sch treatments are the first to bifurcate. This takes place in the range $\phi \in [0.338, 0.339]$ for both. Predicted welfare drops from 0.67 in both cases to 0.27 for she_tar_sai and 0.24 for kag_sch. The empirical welfares are 0.69 and 0.58 respectively. Consequently, the absolute error increases from 0.02 and 0.09 to 0.42 and 0.34, increasing the mean absolute error for the entire sample by about 0.06. In other words, any estimation procedure will need to get the classifications of dre_et_al_08:B, bru_kam, she_tar_sai and kag_sch as right as possible. This means that estimates of ϕ must lie between 0.332 and 0.338.

While it seems that ϕ is tightly pinned down by the data, I think this is deceptive. Near a bifurcation point, the theory is sensitive to small changes in the noise parameter and also small changes in the payoffs and discount factors. For example, in the she_tar_sai treatment predicted welfare is 0.67 and empirical welfare is 0.69. If $u^i(DC)$ is increased slightly from 2.00 to 2.02 then the only equilibrium is mutual defection and predicted welfare crashes to 0.27.

It is unlikely that human behavior is as sensitive to small changes as the theory suggests. Indeed, the central region in Figure 4.1 suggests behavior has substantial continuity. Hence, in the neighborhood of a bifurcation, varying parameters, or varying noise by changing subject pools, is likely to create anomalies. I feel confident that if we conducted additional experiments with various parameters close to the bifurcation point across a range of different laboratories we would indeed see many more anomalies.

While it is desirable to have a theory that predicts well across a broad range of experiments, it is unreasonable to expect that such a theory is going to get the bifurcation points exactly right. From this point of view, we might want to ignore the treatments dre_et_al_08:B, bru_kam, she_tar_sai and kag_sch on the grounds that the theory is not really capable of making good predictions so close to the bifurcation point. Even if we do so, the overall picture does not change much, as we are left with the anomaly dal_fre:C;$\delta = 0.50$ which is strong and robust.

7. Conclusion

I have examined a simple theory of noisy play in which there is a 1/6 probability of a participant choosing an unintended action. This number is not estimated from repeated game data, but was used successfully in an earlier study of non-repeated games. I hypothesize that in the resulting game of imperfect public information welfare is the highest achievable in perfect public equilibrium with memory one strategies.

With this theory, I examine predicted and actual welfare starting from the tenth match in twelve different treatments - all of those from Fudenberg and Rehbinder (2024) that last more than fifteen matches. I find that in ten out of the twelve treatments the theory is accurate to within 0.16, and in seven out of the twelve within 0.05. In two of the treatments, the theory does well, qualitatively, predicting substantial cooperation, but fares poorly quantitatively, with participants achieving in one case 0.34 less than is predicted by the theory.

I compare the theory to existing theories that are primarily motivated by learning. These theories are estimated from data on repeated prisoner's dilemma games, but do no better than noise theory. Finally, I find that if learning theory is combined with noise theory the resulting theory is highly accurate and considerably better than any of the other theories.

References

Arechar, A. A., M. Kouchaki, and D. G. Rand (2018): "Examining spillovers between long and short repeated prisoner's dilemma games played in the laboratory," *Games*, 9(1), 5.

Blonski, M., P. Ockenfels and G. Spagnolo (2011): "Equilibrium Selection in the Repeated Prisoner's Dilemma: Axiomatic Approach and Experimental Evidence," *American Economic Journal: Microeconomics* 3: 164–92.

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

Bruttel, L. and U. Kamecke (2012): "Infinity in the lab. How do people play repeated games?" *Theory and Decision* 72: 205-219.

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

Dal Bo, P. and G. R. Frechette, G. R. (2011): "The evolution of cooperation in infinitely repeated games: Experimental evidence," *American Economic Review* 101: 411-429.

Dal Bo, P. and G. R. Frechette (2015): "Strategy Choice in the Infinitely Repeated Prisoner's Dilemma," unpublished

Dal Bo, P., and G. R. Frechette (2019): "Strategy choice in the infinitely repeated prisoner's dilemma," *American Economic Review* 109: 3929-3952.

Dreber, A., D. G. Rand, D. Fudenberg and M. A. Nowak(2008): "Winners don't punish," *Nature* 452: 348-351.

Dal Bo, P., and G. R. Frechette, G. R. (2018): "On the determinants of cooperation in infinitely repeated games: A survey," *Journal of Economic Literature* 56: 60-114.

Fudenberg, D., D. Levine and E. Maskin (1994): "The Folk Theorem With Imperfect Public Information," *Econometrica* 62: 997-1039.

Fudenberg, D., and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econom*etrica 54: 533.

Fudenberg, D., and G. Karreskog Rehbinder (2024): "Predicting Cooperation with Learning Models," *American Economic Journal: Microeconomics* 16: 1-32.

Ghidoni, R. and S. Suetens(2022): "The effect of sequentiality on cooperation in repeated games," *American Economic Journal: Microeconomics* 14: 58-77.

Kagel, J. H. and D. R. Schley (2013): "How economic rewards affect cooperation reconsidered," *Economics Letters* 121: 124-127.

Kandori, Michihiro, George Mailath, and Rafael Rob (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica* 61: 29-56.

Levine, D. K. (1986): "Modeling altruism and spitefulness in experiments," *Review of Economic Dynamics* 1: 593-622.

Levine, D. K. (2024): "Behavioral Mechanism Design in the Repeated Prisoner's Dilemma," Levine's Working Paper Archive

Levine, D. K. (2025): "Behavioral Mechanism Design as a Benchmark for Experimental Studies," mimeo RHUL.

Mailath, G. J., and L. Samuelson, L. (2006), *Repeated Games and Reputations: Long-run Relationships*, Oxford university press.

Nikiforakis, N. and H. T. Normann (2008): "A Comparative Statics Analysis of Punishment in Public-good Experiments," *Experimental Economics* 11: 358-369.

Proto, E., A. Rustichini and A. Sofianos (2019): "Intelligence, personality, and gains from cooperation in repeated interactions," *Journal of Political Economy* 127: 1351-1390.

Sherstyuk, K., N. Tarui and T. Saijo (2013): Payment schemes in infinite-horizon experimental games, Experimental Economics 16: 125-153.

Young, P. (1993): "The Evolution of Conventions," *Econometrica* 61: 57-83