Strategies in the repeated prisoner's dilemma: A cluster analysis

Heller, Yuval and Tubul, Itay

Bar-Ilan University

25 May 2023

Strategies in the Repeated Prisoner's Dilemma: A Cluster Analysis

Yuval Heller[*]   Itay Tubul[†‡]

May 25, 2023

Abstract

This study uses k-means clustering to analyze the strategic choices made by participants playing the infinitely repeated prisoner's dilemma in laboratory experiments. We identify five distinct strategies that closely resemble well-known pure strategies: always defecting, suspicious tit-for-tat, grim, tit-for-tat, and always cooperating. Our analysis reveals moderate systematic deviations of the clustered strategies from their pure counterparts, and these deviations are important for capturing the experimental behavior. Additionally, we demonstrate that our approach significantly enhances the predictive power of previous analyses. Finally, we examine how the frequencies and payoffs of these clustered strategies vary based on the underlying game parameters.

Keywords: k-means clustering, machine-learning, memory, laboratory experiment, repeated games. **JEL Classification**: C73, C91.

1 Introduction

The infinitely repeated prisoner's dilemma is a widely-used model for examining the balance between self-interested actions and cooperative efforts in long-term, strategic interactions. The theoretical folk-theorem results of the game offer little predictive power in terms of the actual level of cooperation and the strategies that players use in practice. By analyzing the

[*]Department of Economics, Bar-Ilan University and UC San Diego. email: yuval.heller@biu.ac.il.

[†]Department of Economics, Bar-Ilan University. email: tubulit@biu.ac.il

[‡]This work is in partial fulfillment of the requirements for a Ph.D of IT at Bar-Ilan University. We thank seminar participants in Bologna University and Bar-Ilan University for many useful comments. Both authors are grateful the financial support of the European Research Council (#677057), US–Israel Bi-national Science Foundation (#2020022), and the Israeli Science Foundation (#2443/19, #448/22).

strategies adopted by players in the laboratory, we can bridge this gap and gain a better understanding of how mutual cooperation can be sustained in real-life settings.

Dal Bó & Fréchette (2018) have compiled a comprehensive dataset comprising 32 treatments of the infinitely repeated prisoner's dilemma experiment from 12 different papers published between 2005 and 2019. This dataset includes over 150,000 choices made by 1,734 subjects in nearly 807 supergames3). Interestingly, two previous relevant analyses of this comprehensive database (namely, Dal Bó & Fréchette, 2018 and Backhaus & Breitmoser, 2021), have yielded two substantially different results regarding the strategies employed by the players.

Both analyses agree on two key aspects: players' behavior mostly depends on the what was played in the previous round (henceforth, memory-1 strategies), and about 30% of the subjects almost always defect. The two analyses have reached opposing conclusions regarding the behavior of the remaining 70% of the subjects.

Dal Bó & Fréchette (2018, 2019) assert that these players can are categorized into four classes, where each class each follows a different pure strategy: (1) tit-for-tat, which involves cooperating in the first round and playing the opponent's previous action in subsequent rounds; (2) suspicious tit-for-tat, which is similar to Tit-for-tat, but with a defection in the first round; (3) grim, which involves cooperating in the first round and after mutual cooperation, while defecting if either player defected in the previous round; and (4) a smaller class of players who almost always cooperate.

In contrast, Backhaus & Breitmoser (2021) (and Breitmoser, 2015) argue that these players all play a *semi-grim* strategy, characterized by almost always cooperating after mutual cooperation, almost always defecting after mutual defection, and cooperating with a probability of about 35% if either player unilaterally defected in the previous round. These divergent results highlight the need for further research on this important research question.

In this study, we use a new approach using cluster analysis to re-examine the question of evaluating strategies employed in laboratory experiments. Our analysis reveals that subjects can be grouped into five distinct clusters, with each cluster displaying behavior that is relatively similar to the pure strategies identified by Dal Bó & Fréchette (2018, 2019). Furthermore, we identify systematic deviations in each clustered strategy from its corresponding pure strategy and demonstrate that these deviations play an important role in understanding and predicting players' behavior.

Brief description of our approach Following the existing literature we focus our analysis on memory-1 strategies (and we demonstrate in Appendix C the relatively small impact of allowing the players' strategies to depend on longer histories). We represent each player has a vector in the five-dimensional unit cube $[0,1]^5$, where each number corresponds to the player's average cooperation frequency after each memory-1 history (namely, in the first round of a new supergame, after both players cooperating, after unilateral opponent's defection, after unilateral player's defection, and after mutual defection).

In order to categorize the players' behavior in a parsimonious way, we apply the commonly-used machine-learning k-means algorithm (see,Larose & Larose 2014, Chapter 10.5 for a textbook exposition). The algorithm groups data points into k clusters, such that the sum of squared distances between each data point and the mean of the points in its cluster is minimal. By applying a common heuristic (the elbow method), we choose the number of clusters to be $k = 5$ (and we demonstrate the robustness of our results for nearby k-s).

Key Results Our analysis classifies the players into five distinct types. In what follows we briefly describe each type and its average frequency in the database (as detailed below, these frequencies depend on the parameters of the underlying game):

1. 30% of the players consistently defect after all histories, with a slightly lower probability of defection if they cooperated in the previous round.

2. 25% of the players play a strategy similar to tit-for-tat, but with one systematic deviation: occasionally playing their own action in the previous round instead of the opponent's action in the previous round.

3. 20% of the players play a strategy similar to suspicious-tit-for-tat, with the same systematic deviation as tit-for-tat.

4. 20% of the players play a strategy similar to grim, with two systematic deviations: occasionally defecting in the first round of supergames, and occasionally cooperating after a unilateral defection of the opponent.

5. 5% of the players almost always cooperate, with a slightly reduced probability of cooperation if either the player or their opponent defected in the previous round.

Our model improves the predictive power compared to both existing analyses (Dal Bó & Fréchette, 2018; Backhaus & Breitmoser, 2021). Next, we examine how the frequencies and average payoffs of each clustered strategy change with the underlying parameters. Our results show that as the parameters of the underlying games change to facilitate cooperation (decreasing the value of SizeBad, a la Dal Bó & Fréchette, 2011, as defined in Section 3.1), the share and average payoff of the two most cooperative strategies, namely always cooperate and tit-for-tat, increase (65% of the players play one of these two most cooperative strategies when SizeBad is close to zero). Conversely, when the parameters change in the opposite direction, hindering cooperation, the share and average payoff of the two least cooperative strategies, namely always defect and suspicious-tit-for-tat, increase (90% of the players play one of these two least cooperative strategies when SizeBad is close to one). Notably, the grim strategy exhibits a relatively high average payoff for all underlying prisoner's dilemma games, thus demonstrating its robustness.

Structure Section 2 provides a review of the related literature. We describe our methodology and the database in Section 3. Our analysis is presented in Section 4, while Section 5 compares our findings with those of the existing literature. In Section 6, we explore how the frequencies and the average payoffs of the clustered strategies change as a function of the underlying game parameters. We conclude our paper in Section 7.

Additional analysis is available in the online appendices. Appendix A assigns each player to the pure strategy that best predicts their play. Appendix B demonstrates the robustness of our results by removing the early supergames from the data, which may be noisier due to the limited players' experience. The analysis presented in Appendix C shows that allowing the strategies to depend on longer histories has a relatively small impact. Lastly, Appendix D provides the frequency and average payoff of each clustered strategy in each treatment. The code and data used in this paper is available in the supplementary material in GitHub.

2 Related Literature

Eliciting strategies The closest related papers are those that address the same research question as ours (namely, assessing players' strategies in the infinitely-repeated prisoner's dilemma). We have discussed above Dal Bó & Fréchette (2018) and Backhaus & Breitmoser (2021) that have applied econometric methods (such as the Strategy Frequency Estimation Method (SFEM) a la Dal Bó & Fréchette, 2011) to estimate players' strategies. Three other recent papers (Dal Bó & Fréchette, 2019; Romero & Rosokha, 2018, 2023) have used an experimental design that explicitly elicits the players' strategies. In all three papers, each experimental treatment is divided to three phases. In the first phase players play in standard way (i.e., choosing their action at each round, as in our database). Next the players are asked to specify their repeated-game strategy. In the second phase of the experiment, players observe the action recommended by their strategy, but is non-binding, and they can choose a different action, and can modify their elicited strategies. In the third Phase, the computers follow the players' elicited strategies.

The three experiments differ in the elicitation interface, which determines which strategies can be elicited by the participants. Dal Bó & Fréchette (2019) allow the elicitation of pure memory-1 strategies, and an additional small number of strategies that depend on longer histories or has randomness. Romero & Rosokha (2018) allow the elicitation of pure strategies (essentially without limiting the length of memory of these elicited strategies). Romero & Rosokha (2023) allow the elicitation of any behavior memory-1 strategies. The three papers show bounds on differences in the aggregate behavior induced by the elicited strategies and the aggregate behavior when players directly choose their actions (as in the database analyzed in the current paper). Moreover, there are some similar themes across the three papers (such as, the large share of players who play strategies similar to either tit-for-tat, grim or always defect). Having said that, the elicitation process and interface do have non-negligible impact on the players' behavior, which suggests that the present analysis that offers an independent method to asses the players' strategies is valuable. In Section 5 we compare our results with the findings of these three papers (and a brief analysis of the impact of memory-2 histories, and a comparison with Romero & Rosokha, 2018, is presented in Appendix C).

Experiments excluded from our database In what follows we briefly survey the related experimental literature which is *not* included in our database (the experiments included in our database are briefly described in Section 3.3).

Our database focuses on infinitely-repeated prisoner's dilemma games. The various experiments of the finitely repeated prisoner's dilemma (in which the number of rounds in each supergame is known in advance) has been surveyed and studied in the meta analyses of Mengel (2018) and Embrey *et al.* (2018). We exclude experiments in which time is continuous, rather than discrete (see, e.g., Friedman & Oprea, 2012; Bigoni *et al.*, 2015), and those in which choices are made by a team, rather than by a single player (see, e.g., Cooper & Kagel, 2022, which provides evidence that teams tend to be more cooperative than single decision makers). In addition, we exclude experiments in which a player can choose to terminate the current supergame and opt for being rematched with a new opponent (e.g., Honhon & Hyndman, 2020; and see Fujiwara-Greve & Okuno-Fujiwara, 2009).

Our database focuses on experiments in which each player plays against the same opponent throughout the entire supergame. Other experiments (see, e.g., Duffy & Ochs, 2009; Camera & Casari, 2009; Camera *et al.*, 2012) have studies a setup in which players are randomly-matched in each round from the same matching group (typically of size of 4-14), with varying degree of information about the identity and the past behavior of the current opponent against other players. Theoretical predictions for these setups with random-matching has been presented, among others, in Kandori (1992); Ellison (1994); Heller & Mohlin (2018).

Our analysis focuses on games with in which each player perfectly monitors the past actions of her opponent. We exclude experiments with imperfect monitoring. As demonstrated in Aoyagi *et al.* (2019), repeated prisoner's dilemma with imperfect monitoring induce players to follow strategies that are more complex and more lenient than those chose with perfect monitoring. In addition, we exclude experiments in which play is implemented with execution errors (e.g., Dreber *et al.*, 2014; Rand *et al.*, 2015), and games in which players play against computers (e.g., Duffy *et al.*, 2021; Kasberger *et al.*, 2023).

Proto *et al.* (2019) study how differences in intelligence and personality traits affect the behavior and the level of cooperation in various repeated games, including the prisoner's dilemma. Their study found that higher intelligence, conscientiousness, or agreeableness, predict a higher cooperation rate. Similarly, Gill & Rosokha (2020) find that trusting subjects tend to cooperate more in the repeated prisoner's dilemma. We leave the interesting question of the correlation between intelligence and personality traits and the player's clustered strategy to future research.

Cluster Analysis and Economic Research The k-means clustering algorithm is widely applied in data science and computer science due to its versatility and effectiveness across various domains. One important application of the algorithm is image segmentation, where it significantly enhances output image quality and performance measurements (see, e.g., **?**). **?** has combined the k-means algorithm with watershed segmentation to substantially improve medical image segmentation. Another application is molecular biology (see, e.g., **?**), where the k-means algorithm has been used to partition genes into groups based on the similarity of their expression profiles.

Recently, cluster analysis (and related machine-learning algorithms) have been applied to study economic questions, as surveyed in Mullainathan & Spiess (2017), **?**; see also **? ?** for surveys of the finance literature. These recent economic applications include segmentation of credit card customers (Umuhoza *et al.*, 2020), evaluating behavioral models of choice under risk and ambiguity (**?**; see **?** for further uses of machine-learning models in behavioral economics), predicting futile surgeries (**?**), classification and evaluation of NBA players (**?**), improving bail decisions (**?**), and assessing the predictive ability of corporate governance features (**?**).

3 Background and Database

3.1 Repeated Prisoner's Dilemma

Table 1 presents the normalized payoff matrix of the prisoner's dilemma game. Each player has two actions, denoted by c and d, representing cooperation and defection, respectively. When both players cooperate they both get a relatively high payoff (normalized to one), and when they both defect they both get a relatively low payoff (normalized to zero). When a single player defects she obtains a payoff of $1 + g$ (i.e., an additional payoff of g) while her opponent gets $-l$, where $g, l > 0$.[1]

	c	d
c	1 , 1	$-l$, $1+g$
d	$1+g$, $-l$	0 , 0

Table 1: Prisoner's Dilemma Payoff Matrix ($g, l \in (0, 1 + l)$)

In the (infinitely-)repeated (discounted) prisoner's dilemma, players are randomly matched to play repeated *rounds* of the underlying game, with a discount factor of $\delta \in (0, 1)$. This is typically implemented in experiments by having an independent continuation probability δ after each round; with the remaining probability of $1 - \delta$ a new *supergame* begins (i.e., the players are matched with new anonymous opponents) or the session ends (see the discussion of various experimental implementations of the discount factor in Fréchette & Yuksel, 2017). It is well-known that for low discount factors, the unique Nash equilibrium outcome is always defecting. Higher discount factors introduce additional equilibria. In particular, mutual cooperation is a subgame perfect equilibrium iff $\delta \geq \delta^{SPE} \equiv \frac{g}{1+g}$. When the session ends, each subject gets a real monetary payoff (typically, a few dozen dollars), which depends on her payoffs in a few randomly selected rounds.

Next we describe two parameters introduced in the existing literature to predict the rate of cooperation. The first parameter, $\delta^{RD} = \frac{g+l}{1+g+l}$, is the minimal discount factor for which mutual cooperation is risk dominant (Harsanyi & Selten, 1988) in the sense of being the best reply against an opponent who with probability 50% always defects and with probability 50% plays the *grim* strategy (i.e., defects iff the opponent has ever defected).[2] The

[1] It is also often required that $1 + g - l < 2$, which implies that mutual cooperation is the efficient action profile, maximizing the sum of payoffs. In most of the lab experiments, the participants observe a (non-normalized) payoff matrix with integer payoffs: R for mutual cooperation, P for mutual defection, T for a sole defector, and S for a sole cooperator, where $T > R > P > S > 0$. The standard normalization of these four parameters to the two parameters of g and l is as follows: $g = \frac{T-R}{R-P}, l = \frac{P-S}{R-P}$.

[2] δ^{RD} was introduced in Blonski *et al.* (2011) with an axiomatic foundation (where it was denoted by δ^*), and its risk-dominant interpretation was presented in Blonski & Spagnolo (2015).

second parameter, *SizeBad*, is defined by Dal Bó & Fréchette (2011, 2018) as the maximum probability of the opponent following the grim strategy such that always defecting is the best reply against such an opponent:

$$\text{SizeBad} = \begin{cases} 1 & \text{if} \delta < \delta^{\text{SPE}} \equiv \frac{\text{g}}{1+\text{g}} \\ \frac{(1-\delta)l}{(1-(1-\delta)(1+g-l)} & \text{otherwise.} \end{cases} \tag{3.1}$$

3.2 Memory-1 Strategies

The repeated prisoner's dilemma admits an infinite number of strategies. Tractability requires focusing on a small finite subset of strategies. We follow most of the existing literature (e.g., Breitmoser, 2015; Dal Bó & Fréchette, 2018; Romero & Rosokha, 2023) and focus on *memory-1 strategies*, which are strategies in which the player's behavior depends only on the actions played in the previous round of the current match. This focus can be justified by the following findings in experiments with strategy elicitation:

1. 91% of the players choose memory-1 strategies when the elicitation interface allows choosing variants of grim and tit-for-tat strategies with longer memories (Dal Bó & Fréchette, 2019); and

2. 72% of the players choose strategies that are very similar to the one of the four most popular memory-1 strategies (AD, GR, TFT and STFT, in the notation of Table 2) when the elicitation interface allows choosing pure strategies of arbitrary length of memory (Romero & Rosokha, 2018, Result 3).

Appendix C presents evidence suggesting that the players' behavior mainly depends on the most recent round. Although we find and report some systematic impact for the actions played in the penultimate round, this impact is relatively small.

We denote a memory-1 strategy by a vector $\sigma = (\sigma_0, \sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd}) \in [0,1]^5$, where the first component σ_0 describes the probability of cooperation in the first round of a supergame, and each of the four remaining components σ_{ab} describes the probability of cooperation when in the previous round the player played a and her opponent played b (for example, σ_{cd} denotes the cooperation probability of a player when in the previous round she cooperated and the opponent defected). There are $32 = 2^5$ *pure memory-1 strategies* (in which, each component is either 0 or 1). In addition, there is a continuum of *behavior memory-1 strategies*, in which each component is between 0 and 1. Table 2 presents the commonly analyzed memory-1 strategies and some variants of them that will play a role in our analysis: 8 pure strategies (namely, always defect, always cooperate, grim, suspicious grim, tit-for-tat, suspicious tit-for-tat, and win-stay-lose-shift), and the behavior strategy of semi-grim (Backhaus & Breitmoser,

Table 2: Memory-1 Strategies of Special Interest

Strategy	Abbr.	$(\sigma_0, \sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd})$	Description
Always defect	AD	$(0,0,0,0,0)$	Always defects
Always cooperate	AC	$(1,1,1,1,1)$	Always cooperates
Grim	GR	$(1,1,0,0,0)$	Only cooperates in R1 and after CC
Noisy Grim	NGR	$(1-\epsilon, 1-\epsilon, \epsilon, \epsilon, \epsilon)$	Grim with error probability $\epsilon \leq 0.5$
Suspicious grim	SGR	$(0,1,0,0,0)$	Only cooperates after CC (off the equilibrium path)
Tit-for-tat	TFT	$(1,1,0,1,0)$	Start with c, then copy opponent
Suspicious TFT	$STFT$	$(0,1,0,1,0)$	Start with d, then copy opponent
Win-stay-lose-shift	$WSLS$	$(1,1,0,0,1)$	Cooperate in R1, cc and dd
Semi-grim (behavior strategy)	SmG	$(\sigma_0, 1, \sigma_d, \sigma_d, 0)$	Cooperate after cc, defect after dd. Mix in 1st round, and mix after a single defection ($\sigma_d \equiv \sigma_{cd} = \sigma_{dc}$)

2021). In addition we demonstrate how a noisy variant of a pure strategy looks like with noisy grim, in which a player follows the pure strategy grim with portability $1 - \epsilon$ in each round, and plays the opposite action with probability $\epsilon < 0.5$, where this in interpreted as a "trembling-hand" error.

3.3 Database

The experimental research on the repeated prisoner's dilemma is usually done in a laboratory environment, and usually includes undergraduate students where most of them are without a game theory background. Participants are randomly divided into pairs and are given verbal and written instructions concerning the rules and payoffs of the game. The instructions are phrased in terms of the individual's payoffs as a function of his own decision (to cooperate or not) and the decisions made by the opponent.

We analyze in this paper the large database kindly constructed and publicly shared by Dal Bó & Fréchette (2018), which comprises the modern experiments on repeated (and one-shot) Prisoner's Dilemmas with perfect monitoring. These modern experiments are characterized by anonymous matching between human players, neutral framing of the actions and the game (i.e., presenting the actions as a and b, and not as cooperation and defection), being truthful to the subjects, and providing monetary incentives to the subjects, which are proportional to their game payoffs. After removing experimental treatments with one-shot prisoner's dilemma games, the remaining dataset includes a total number of 1734 players that do 145,800 choices in 32 different treatments taken from 12 experimental papers, as summarized in Table 3.

In order to maximize the amount of data for each subject, we have chosen in the main-text analysis to study all actions played by each subject. In Appendix B we present an alternative specification that removes from the data the early supergames in which the experimental behavior might be nosier (a methodology commonly used in the existing literature, see, e.g., Dal Bó & Fréchette, 2018; Backhaus & Breitmoser, 2021). The appendix demonstrates that the main results of our analysis remain qualitatively similar when removing up to 50% of the early supergames in each treatment.

4 Analysis

4.1 Memory-1 Strategy Space

The first step of the analysis is to represent each player i by a vector $\sigma = (\sigma_0^i, \sigma_{cc}^i, \sigma_{cd}^i, \sigma_{dc}^i, \sigma_{dd}^i) \in [0,1]^5$, which captures her average memory-1 behavior. Recall that the first component σ_0^i describes her frequency of cooperation in the first round of her supergames, and each component σ_{ab}^i describes her frequency of cooperation when in the previous round player i played action a and her opponent played action b. For example, if player i had encountered 20 rounds following a memory-1 history of cd (player i cooperating and her opponent defecting), and she had cooperated in 8 of these rounds, then we set $\sigma_{cd}^i = \frac{8}{20} = 40\%$).

Some of the players have not encountered all four memory-1 histories (for example, a player who defects with high probability is likely to never encounter the history in which both players have cooperated in the previous round).[3] In such cases, we set the cooperation probability after a never-encountered memory-1 history as the player's mean frequency of cooperation (e.g., if player i has played a total number of 300 rounds, out of which she cooperated in 30 rounds, and if she had never encountered the history cc, then we set $\sigma_{cc}^i = \frac{30}{300} = 10\%$).

The representation of the 1734 players as points in the memory-1 strategy space is illustrated in Figure 4.1. Each panel illustrates the frequency of cooperation in three out of the five dimensions of the memory-1 strategy space. Both panels show in the x-axis the frequency of cooperation in the first round. The left panel shows the frequency of cooperation after history cd in the y-axis and after history dc in the vertical z-axis, The right panel shows the frequency of cooperation after history cc in the y-axis and after history dd in the z-axis. As can be seen in the right panel, most players cooperate with high probability after cc,

[3] Specifically, 68% of the subjects encountered all four memory-1 histories, 16% encountered three of these histories, 14% encountered two of these histories, and 2% encountered only one memory-1 history.

Table 3: Description of the Analyzed Dataset (Dal Bó & Fréchette, 2018)

Experimental Paper	Treat. ID	#Super games	#Cho -ices	#Pla -yers	Coop. Freq.	Parameters δ	g	l	δ_{SPE}	δ_{RD}	Size Bad
Aoyagi & Fréchette (2009)	AF1	10	3,050	38	74%	0.9	0.33	0.11	0.01	0.31	1%
Blonski *et al.* (2011)	BOS2	11	740	20	42%	0.75	0.75	1.25	0.36	0.67	36%
	BOS3	11	860	20	25%	0.75	0.83	0.5	0.19	0.57	19%
	BOS4	8	1,200	20	24%	0.875	2	2	0.29	0.80	29%
	BOS5	11	1,040	20	19%	0.75	1	1	0.33	0.67	33%
	BOS6	11	1,860	40	17%	0.75	2	2	0.67	0.80	67%
	BOS7	11	520	20	14%	0.5	2	2	1	0.80	100%
	BOS8	30	700	20	11%	0.875	0.5	3.5	0.35	0.80	35%
	BOS9	11	620	20	5%	0.75	0.5	3.5	0.58	0.80	58%
	BOS10	11	960	20	1%	0.75	1	8	0.8	0.90	80%
Bruttel & Kamecke (2012)	BK11	20	3,552	36	31%	0.8	1.17	0.83	0.23	0.67	23%
Dal Bó (2005)	D12	10	1,050	60	38%	0.75	0.83	1.17	0.35	0.67	35%
	D13	7	1,920	42	36%	0.75	1.17	0.83	0.31	0.67	31%
Dal Bó & Fréchette (2011)	DF14	35	6284	44	76%	0.75	0.09	0.57	0.16	0.40	16%
	DF15	77	6490	46	35%	0.5	0.09	0.57	0.38	0.40	38%
	DF16	33	6,448	44	20%	0.75	2.57	1.86	0.81	0.82	81%
	DF17	72	6,904	50	18%	0.5	0.67	0.87	0.72	0.61	72%
	DF18	71	5,736	44	10%	0.5	2.57	1.86	1	0.82	100%
	Df19	47	5180	38	59%	0.75	0.67	0.87	0.27	0.61	27%
Dal Bó & Fréchette (2019)	DF20	24	10,830	164	62%	0.75	0.09	0.57	0.16	0.40	16%
	DF21	7	4,356	36	50%	0.95	2.57	1.86	0.1	0.82	10%
	DF22	46	8,750	140	48%	0.5	0.09	0.57	0.38	0.40	38%
	DF23	21	16,156	168	33%	0.9	2.57	1.86	0.22	0.82	22%
	DF24	25	8,424	114	21%	0.75	2.57	1.86	0.81	0.82	81%
	DF25	37	3,076	50	9%	0.5	2.57	1.86	1	0.82	100%
Dreber *et al.* (2008)	Dea26	27	1,914	22	43%	0.75	1	1	0.33	0.67	33%
	Dea27	21	1,988	28	21%	0.75	2	2	0.67	0.80	67%
Duffy & Ochs (2009)	DO28	13	9,146	102	54%	0.9	1	1	0.11	0.67	11%
Fréchette & Yuksel (2017)	FY29	12	2,368	50	65%	0.75	0.4	0.4	0.13	0.44	13%
Fudenberg *et al.* (2012)	Fea30	9	3,252	48	74%	0.875	0.33	0.33	0.05	0.40	5%
Kagel & Schley (2013)	KS31	39	14,772	114	48%	0.75	1	0.5	0.2	0.60	20%
Sherstyuk *et al.* (2013)	Sea32	29	5,656	56	55%	0.75	1	0.25	0.11	0.56	11%
Total / Average		807	145,802	1734	41%	-	-	-	-	-	-

and defect with high probability after dd, which implies that the heterogeneity in these two dimensions is limited. By contrast, there is substantial heterogeneity in the players behavior in the remaining three dimensions (i.e., in the first round of a supergame, and after histories in which one of the players defected).

Figure 4.1: Cooperation rate for each participant after each memory-1 history

`the diff.png`

Each panel shows the frequency of cooperation of each player for 3 of the 5 memory-1 histories. Both panels show in the x-axis the frequency of cooperation in the first round. The left (resp., right) panel shows the frequency of cooperation after cd (resp., cc) in the y-axis and after dc (resp., dd) in the vertical z-axis. The shape and color of each point represents how many of the 3 dimensions include histories that the player has experienced.

4.2 Cluster Analysis

k-Means Algorithm In what follows we briefly describe the algorithm that we used to classify the players into clusters. The k-means algorithm is a commonly-used method to partition the dataset into a pre-defined number (k) of disjoint clusters/groups (see, e.g., Larose & Larose 2014, Chapter 10.5 for a textbook exposition). The objective of the k-means algorithm is to obtain k clusters for which the sum of squared distances between each data point (subject) and the mean of the points in its cluster is minimal (henceforth called Within-Cluster Sum of Square, and abbreviated as WCSS).

The k-Means algorithm selects k data points randomly and uses them as initial values for the cluster means.[4] The algorithm then iterates between the following two steps: (1) assigning each point to the cluster with the nearest mean, and (2) recalculate the mean of each cluster as the mean location of the observations assigned to the cluster. The algorithm continues to iterate until the clusters and their means do not undergo any change, and calculate the obtained WCSS. Finally, in order to avoid the risk of the outcome being affected by a convergence to a local (rather than global) minimum, the algorithm repeats the above process 300 times (each time with new random initial values for the cluster means) and returns the minimal outcome.[5]

Choosing *k* Next we describe the method that we used to determine the number of clusters (k). The *elbow method* is a heuristic commonly used to determine the optimal k in k-means clustering (see, e.g., Yuan & Yang, 2019; Umargono *et al.*, 2020). The heuristics calculates the WCSS induced by the k-means algorithm as a (decreasing) function of k, presents it graphically, and aims to choose the value of k that is the "elbow of the curve"(a turning point in which the curve bends from a high slope to a low slope).

The right panel in Figure 4.2 presents the WCSS induced by the k-means algorithm as a function of k. The left panel presents the log-likelihood obtained by k-mean clustering for each value of k. That is, for each k, and each cluster, we assume that all players assigned to the cluster follow the cluster's average behavior strategy (as detailed in Table 4 below), and calculate the (absolute value of) the log-likelihood of the observed play in all experiments under the assumption that all players follow the average behavior of their cluster.

The WCSS analysis suggests that the number of clusters should be 5 ± 1 (as the slope of the WCSS is approximately constant for the values above 6, while the slope substantially decreases till 4). The likelihood analysis suggests the number of clusters should be 4 (or slightly above 4), as adding more clusters above 4 only marginally improve the log-likelihood. As further discussed below, we have chosen $k = 5$ as the number of clusters in our main analysis (and we test the robustness of the results to close values of k around 5).

The Clusters It turns out the key properties of the clusters are robust, in the sense that increasing k by 1 typically adds a new cluster, without substantially changing the properties of the pre-existing clusters. Table 4 presents the average frequency of cooperation after

[4] We have followed the algorithm of Vassilvitskii & Arthur (2006), according to which, the first cluster mean is chosen uniformly among all the data points. The $k-1$ other initial cluster means are chosen sequentially, using a weighted probability distribution, where each point is chosen with probability proportional to the square of its distance to the nearest existing cluster center.

[5] We have manually run the algorithm at-least 10 times for each analysis presented in the paper (with 300 iterations in each time as described above), and we have confirmed that in all cases the algorithm returns the same WCSS (our Python code is detailed in the supplementary material in Github).

Figure 4.2: Log-likelihood and WCSS as a Function of the Number of Clusters k

MLE and WCSS.png

each memory-1 history in each cluster for each value of k between 1 and 7 (e.g., if the total number of times in which the players within a cluster faced history ab 300 times, and they cooperated in 180 of the rounds following this history, then this average probability is equal to 60%). Table 5 presents the five clusters induced by our preferred specification of $k = 5$, and demonstrates their robustness to changing the number of clusters to 4, 6 and 7.

When $k = 2$, about 30% of the population is clustered into AD_c. These agents almost always defect in the first round, and their probability of cooperation remains low for all histories (although, being slightly higher if the player cooperated in the previous round). This cluster remains essentially the same for all values of k between 2 and 7. Increasing the number of clusters to 4 divides the remaining population into three additional clusters, each with a size of 20%-25% of the population. The average behavior of each of these clusters remain stable for all values of k between 4 and 7. In each of these clusters agents almost always cooperate after mutual cooperation and almost always defect after mutual defection.

The agents in the second cluster, $STFT_c$, present a behavior similar to suspicious tit-for-tat. The agents typically defect in the first round, but they tend to shift to cooperation if the opponent cooperates, and they keep cooperating as long as both players cooperated in the previous round. If the opponent unilaterally defects (resp., cooperates), players following $STFT_c$ are relatively likely to defect (resp., cooperate) in the next round, though this reaction is substantially weaker than of the pure strategy $STFT$.

The third cluster is the grim-like cluster GR_c, which corresponds to agents who initially cooperates with a relatively high probability, and they continue cooperating as long as no player has defected, while they defect if any player defected in the previous round (with a somewhat lower defection probability after a unilateral opponent's defection). The fourth cluster, TFT_c presents a tit-for-tat-like behavior (usually playing the action played by the opponent in the previous round) that mainly differs from $STFT_c$ by having the agents starting most of the supergames by cooperation. Similar to $STFT_c$, the clustered strategy TFT_c differs from its pure counterpart by having a weaker response to the opponent's action after a recent history in which the players played different actions.

Finally, when increasing the number of clusters to 5, a new small cluster, AC_c, appears, and the agents in this cluster are always more likely to cooperate than defect. Unlike the pure strategy of AC, the players have some tendency to defect if either player has defected in the previous round. Because the size of this cluster is small (5%), its impact on the likelihood and on the WCSS is relatively small (see Figure 4.2). Despite this we choose to include this fifth cluster in our main analysis because we think that it does truly capture the behavior of a group of subjects in prisoner's dilemma experiments (the cluster remains essentially the same for higher values of k in our analysis, and its pure-strategy counterpart AC has been found to describe the behavior of a small, yet statistically significant, share of agents in the existing literature, see, e.g., Dal Bó & Fréchette, 2018, 2019).

Figure 4.3 illustrates the locations of the 5 clusters in the two three-dimensional sub-spaces, as introduced in Figure 4.1. Figure 4.4 illustrates the locations of the five clusters in each of the 20= $\frac{5 \cdot 4}{2}$ two-dimensional sub-spaces. In addition, the panels in the main diagonal of Figure 4.4 show the distribution of the players' frequency of cooperation in each cluster after each history.

5 Comparison with Existing Analyses

In the next section, we will compare our results with the existing analyses of the same database (or of a large subset of this database) in Dal Bó & Fréchette (2018), Backhaus & Breitmoser (2021), and Breitmoser (2015).

5.1 Log-Likelihood and Completeness

Substantial parts of out comparisons will present the (absolute value of the) log-likelihood induced by the various models. As an intuitive scale for these likelihoods we adapt Fudenberg *et al.*'s (2022) notion of completeness of models to the current setup. Our database includes

Figure 4.3: Graphic Illustration of the Five Clusters in Two 3-Dimensional Spaces

3D Kmeans.png

The location of each point describes the frequency of cooperation of each player as a function of the memory-1 history, and the color of the point shows its cluster. Both panels show in the x-axis the frequency of cooperation in the first round. The left (resp., right) panel shows the frequency of cooperation after cd (resp., cc) in the y-axis and after dc (resp., dd) in the z-axis. The larger circles show the average probability of cooperation after each history in each cluster.

145,802 binary choices of players. A trivial model that predicts that players always play uniformly would achieve a log-likelihood of 101,062 for any observable behavior. We have done a similar analysis focusing only on each half of the data separately (where, the first half of the data including the first half of the supergames played in the treatment, rounding up). The analogous calculation of the log-likelihood induced by a trivial prediction of uniform play yields 52,708 for the first half (with 76,042 binary choices) and 48,354 for the second half (with 52,708 binary choices).

Next, we calculate the lower bound on the log- likelihood that can be obtained by any prediction rule that relies on memory-1 histories. This lower bound is achieved by predicting for each player and for each memory-1 history a probability of cooperation that exactly matches the player's empirical frequency of play after this memory-1 history (i.e., this is the log-likelihood induced by our clustering algorithms, if one were using 1,734 clusters, one for each player). Applying this lower bound on our data yields log-likelihood of 31,399. A similar lower bound applied to each half of data (where the values of each player's l cluster

Figure 4.4: Graphic Illustration of the Five Clusters in all 2-Dimensional Sub-Spaces

save_as_a_png.png

The five main-diagonal panels show the distribution of the players' frequency of cooperation in each cluster after each history. The twenty off-diagonal panels illustrate the locations of the 5 clusters in all the 20 two-dimensional sub-spaces of the memory-1 space.

is adjusted to best fit her behavior in this half of the data) yields 17,228 for the first half and 9,264 for the second half.

Thus, we estimate the level of completeness of a model that induces log-likelihood of X in our database to be $\frac{101{,}062-X}{101{,}062-31{,}399}$. For example, the completeness of our k-means model with five clusters is $\frac{101{,}062-X}{101{,}062-39{,}303} = \frac{101{,}062-41{,}660}{101{,}062-31{,}399} \approx 85.3\%$.

5.2 Comparison with Dal Bó & Fréchette (2018)

The main conclusion of Dal Bó & Fréchette's (2018) strategy analysis of the data has been that the behavior of the large majority of subjects can be explained by them following one out of five pure strategies: AD, AC, GR, TFT and STFT (and the conclusions are similar in Dal Bó & Fréchette, 2019).[6] Our analysis has followed a substantially different classification methodology (namely, the k-means algorithm), and has yielded similar results: players can be clustered to five classes of behaviors, and each such behavior corresponds to a behavior strategy that is relatively close to one of the five pure strategies of Dal Bó & Fréchette. This provides a significant independent support to their main finding.

Dal Bó & Fréchette's analysis implicitly assumes that players' deviations from these five pure strategies are the result of a random noise. Our analysis suggest that the clustered strategies has some systematic differences with respect to their counterpart pure strategy. In what follows, we evaluate these systematic differences, and how much the log-likelihood of the observed data is improved by incorporating these systematic deviations.

We have clustered each subject to one of the 5 pure strategies $\{AD, AC, GR, TFT, STFT\}$ that predicts the maximal share of the subject's behavior in the experiment (breaking ties uniformly). For each strategy $\sigma \in \{AD, AC, GR, TFT, STFT\}$ we have chosen the error probability ϵ_σ (which is assumed to be the same for all histories; henceforth, history-independent noise) that maximizes the likelihood of the observed behavior of all agent who have clustered to this strategy. Table 6 presents the size of each cluster and the mean frequency of cooperation in each cluster after each memory-1 history, when using the five pure strategies of Dal Bó & Fréchette (2018), and when using the five clusters induced by our analysis of the k-means algorithm. Observe that both methods yield very similar cluster sizes. In addition, Table 6 highlights the systematic deviation of each clustered strategy from its pure/noisy counterpart strategy.

Table 7 presents the log-likelihood and completeness measure induced by clustering the players according to Dal Bó & Fréchette's pure strategies (with likelihood-maximizing history-independent noise) and compare it to the log-likelihood induced by our five clusters.

[6] Dal Bó & Fréchette (2018) strategy analysis relies on SFEM (other paper that use SFEM are).

The table shows that the systematic deviations of our clustered strategies substantially improves the log-likelihood (the model's completeness improves from 77.8% to 85.3%). This substantial improvement is achieved even though the parameters of our clusters have been chosen to minimize the WCSS (and not to maximize the log-likelihood). This suggests that the systematic deviations of our clustered strategies from their pure counterparts are likely to capture important aspects of the experimentally observed behavior. We have done an analogous analysis for each half of the data separately (where the centers of the clusters remain the same as calculated for the whole data, and each player is assigned to one of the five clusters that maximizes the likelihood of her play in this half of the data). This analogous analysis yields similar results. We note that the gap in favor of our model's predictions somewhat narrows for the second half of the data, as the behavior of most players is closer to pure strategies in the second half of the treatments.

Another exercise that we did was to calculate the likelihood induced by clustering the players to 32 clusters, corresponding to all feasible memory-1 pure strategies. That is, we classify each subject to the pure strategy that bests predicts her behavior (with an arbitrary tie-breaking rule). We then assign to each pure strategy σ the level of history-independent noise ϵ_σ that minimizes the likelihood of the observed behavior of all agent who have clustered to this strategy. This classification to all possible pure strategies yields log-likelihood of 44,581, which corresponds to completeness level of 81%. Thus our k-means model with five clusters substantively improve the induced likelihood of the classification to 32 pure strategies (i.e., it increases completeness from 81% to 85%), which gives demonstrates the importance of the moderate systematic deviations from pure strategies mentioned above to understanding and predicting the players' behavior.

5.3 Comparison with Backhaus & Breitmoser (2021)

Backhaus & Breitmoser (2021) identify three different types of subjects:

1. Always defect – AD_B= $(\epsilon_A, \epsilon_A, \epsilon_A, \epsilon_A, \epsilon_A)$: Players who always defect with a history-independent noise of ϵ_A.

2. Cooperative semi-grim – CSG_B= $(\sigma_{0C}, 1-\epsilon_{C_1}, \sigma_C, \sigma_C, \epsilon_{C_2})$: Players who cooperate with high probability $\sigma_{0C} > 50\%$ in the first round, and follow a semi-grim strategy in later rounds: almost always cooperate after mutual cooperation (with error probability ϵ_{C_1}), almost always defect after mutual defection (with error probability ϵ_{C_2}), and cooperate with interior probability σ_C after exactly one of the players defected.

3. Suspicious semi-grim – SSG_B= $(\sigma_{0S}, 1-\epsilon_{S_1}, \sigma_S, \sigma_S, \epsilon_{S_2})$: Players who cooperate with

low probability $\sigma_0 < 50\%$ in the first round, and follow a semi-grim strategy in later rounds: almost always cooperate after mutual cooperation (with error probability ϵ_{S_1}), almost always defect after mutual defection (with error probability ϵ_{S_2}), and defect with interior probability σ_S after exactly one of the players defected.

We initiated the parameters above according to the average estimated values of Backhaus & Breitmoser's Table 9; that is, Backhaus & Breitmoser allow different estimation for each parameter in each treatment, and we took the weighted averages of these estimations, such that they will be the same in all treatments. Specifically, the initial values were: $\epsilon_A = 6\%$, $\sigma_{0C} = 87\%$, $\sigma_C = \sigma_S = 35\%$, $\epsilon_{C_1} = \epsilon_{S_1} = \epsilon_{C_2} = \epsilon_{S_2} = 7\%$, and $\sigma_{0S} = 35\%$. Next we have clustered each subject to the strategy among these 3 strategies that maximizes her likelihood of play. Finally, we have adapted the values of the various parameters of these 3 strategies, such that each strategy will maximize the likelihood of the behavior of subjects clustered to this strategy. The final values of the parameters of the 3 strategies are presented in Table 8.

Table 9 presents the log-likelihood when clustering the players à la Backhaus & Breitmoser's model (namely, players are clustered to one of the 3 strategies described in Table 8 with the parameters that maximizes the log-likelihood), and compare it with the log-likelihood obtained by the k-means clustering for $k = 3, 4, 5$. The table shows that the clustering to the 3 strategies of Backhaus & Breitmoser yields lower log-likelihood than the once induced by k-means clustering with our parameters (recall, that the parameters of our model are chosen to minimize the WCSS without regarding the induced log-likelihood). Increasing the number of the clusters in the k-means algorithm from 3 to 4 substantially decrease the log-likelihood, such that the level of completeness of our model with 4 clusters is 1.3 percentage points higher than Backhaus & Breitmoser 's. Increasing the number of clusters to 5 further increase the gap in favor of our clustering to 2.3 percentage points (85.3% vs. 83%). Similar results hold when focusing on the behavior in the second half of each treatment. Thus our analysis suggests that capturing the behavior of the non-AD players requires more than two clusters, and that the two semi-grim strategies a la Backhaus & Breitmoser are induced by averaging different clusters, as detailed in the next subsection.

5.4 Comparison with Breitmoser (2015)

Breitmoser (2015) has observed that the aggregate behavior in experiments of the prisoner's dilemma is close to semi-grim behavior, where player almost always cooperate after mutual cooperation almost always defect after mutual defection, and they mix with the same probability after a unilateral defection of either player. We replicate the same result when classifying all players to a single cluster (i.e., for $k = 1$). By contrast, none of our five clustered strategies is close to a semi-grim behavior. In order to shed more light on this, we present in Table 10 the frequency of cooperation after each memory-1 history and the number of occurrences of each such history for each clustered strategy.

Table 10 shows that the aggregate semi-grim-like behavior, in which the frequency of cooperation after *cd* is similar to the frequency of cooperation after *dc* is induced by two opposing factors:

1. The players who have been classified to the two variants of tit-for-tat (namely, TFT_c and $STFT_c$) are substantially more likely to cooperate after a recent history *cd* (in which the player was the sole defector) relative to a recent history of *dc* (in which the opponent was the sole defector). As a result, a calculation of the average probability of cooperation that gives equal weight to all players would yield a higher value after *cd* than after *cd*.

2. Agents who follow less cooperative strategy (and, in particular, those clustered to AD_c) tend to face the recent history of *dc* much more often than the recent history *cd* (in particular, the former history is 4 times more frequent for those clustered to AD_c). This and the fact that these players cooperate less often than other player implies that the weighted average probability of cooperation (which gives more weight to players who face a history more often) after *dc* is substantially reduced, and becomes similar to the average probability of cooperation after *cd*.

6 Frequencies and Payoffs

In this section we explore the relations between the frequencies and payoffs of the clustered strategies and the parameters of the underlying prisoner's dilemma game.

6.1 Frequencies of the Clustered Strategies

The analysis of the previous sections has classified each subject to one of five clustered strategies, and it has presented the average frequency of these five strategies across all treatments. In what follows we study how these frequencies depend on the underlying parameters of the game. Specifically, Figure 6.1 shows the frequency of each clustered strategy as a function of SizeBad (the parameter defined in (3.1), which, roughly speaking, captures the size of the basin of attraction of the cooperative equilibrium). The focus of this

[7]The small difference of 2 between the number of *cd* histories and the number of *dc* histories is due to a small discrepancy in the data presented for Duffy & Ochs (2009) in Dal Bó & Fréchette's aggregate database.

Figure 6.1: The Frequency of each Clustered Strategy as a Function of SizeBad

SIZEBADNew.png

The x-coordinate of each dot presents the average value of SizeBad in the relevant interval (within the 6 intervals of [0,0.1], [0.1,0.3], [0.3,0.5], [0.5,0.7], [0.7,0.9], and [0.9-1]), and the y-coordinate presents the average frequency of the clustered strategy in the interval.

analysis on SizeBad is motivated by the finding of Dal Bó & Fréchette (2018) that SizeBad predicts the cooperation rate very well.[8]Specifically, we divide the domain of SizeBad values to 6 intervals: [0,0.1], [0.1,0.3], [0.3,0.5], [0.5,0.7], [0.7,0.9], and [0.9-1]. The x-coordinate of each dot in Figure 6.1 presents the average value of SizeBad in the relevant interval, and the y - of each dot presents the average frequency of the clustered strategy in the interval.

Figure 6.1 shows that increasing SizeBad decreases the frequency of the more cooperative strategies (AC_c,TFT_C and GR_C), while it increases the frequency of the less cooperative strategies (AD_C and $STFT_C$). Specifically, when SizeBad is close to 0, 85% of the players play the three more cooperative strategies (the share of TFT_C is about 40%, and the share of GR_C and AC_c is about 20-25% each), while when SizeBad is close to 1, about 90% of the players play the two less cooperative strategies (the share of isAD_C about 70% and the share of $STFT_C$ is about 20%).

6.2 Mean Payoffs of the Clustered Strategies

Next, we study the average payoffs induced by following each of the five clustered strategies, and how they depend on the parameters of the game.

We begin by calculating for each subject in each treatment her (aggregate) payoff in all the rounds of all supergames in which the subject has participated (i.e., the sum of her payoffs in the all the rounds that she played). In order to allow comparison of various experimental treatments with different payoff parameters and different supergame lengths, we normalize the subjects' payoffs in each treatment by a linear transformation that changes the payoff of the subject with the highest payoff to 1, and the payoff of the subject with the lowest payoff to 0. Figure 6.2 illustrates the average normalized payoff of the players clustered to each

[8]Similar finding is presented in Embrey *et al.* (2018) for finitely-repeated prisoner's dilemma. Mengel *et al.* (2022) suggest that, in addition to SizeBad, the earlier match length realizations impact the cooperation rate; we leave for future reseach the interesting question of the impact on earlier match realizations on the frequencies and payoffs of the clustered strategeies.

strategy in each of the six intervals of SizeBad.

Figure 6.2: Normalized Payoffs as a Function of SizeBad

`normalized payoffs.png`

The x-coordinate of each dot presents the average value of SizeBad in the relevant interval (within the six intervals of [0,0.1], [0.1,0.3], [0.3,0.5], [0.5,0.7], [0.7,0.9], and [0.9-1]), and the y-coordinate presents the average normalized payoff of the players clustered to the relevant strategy in the interval. The data for AC_c is presented only in the first 3 intervals because a single player was classified to AC_c in the remaining 3 intervals.

Figure 6.2 shows that:

1. The two most cooperative clustered strategies AC_c and TFT_C yield high payoffs for low levels of SizeBad (up to 0.4). By contrast, when SizeBad is large (at-least 0.6), these clustered strategies either yield low payoffs (TFT_C) or their frequency decreases to zero (AC_c).

2. The two least cooperative strategies AD_C and $STFT_C$ yield low payoffs for low levels of SizeBad. By contrast, when SizeBad is large, they yield high payoffs.

3. The intermediate cooperative strategy GR_C is robust in the sense of yielding relatively high payoffs for all possible values of SizeBad (which only slightly decrease for higher values of SizeBad).

Observe that the frequencies and payoffs of the different strategies are positively correlated. That is, the strategic choices of the players react to the environment in the sense of increasing the frequency of the clustered strategies that yield higher payoffs.

6.3 Comparison with Romero & Rosokha (2023)

Romero & Rosokha 's(2023) main treatment have elicited memory-1 behavior strategies of 124 players who played a repeated prisoner's dilemma experiment with parameters of $\delta = 0.95$, $g = 2.57$, and $l = 1.86$, which induce a value of 10% for SizeBad. Table 11 compares our model's predictions (with SizeBad=10%, see Figure 6.1) with the result of the clustering analysis of the elicited strategies in the main treatment of Romero & Rosokha (2023, Figure 2).

Table 11 shows a good fit between our predictions and the distribution of elicited strategies in 11. In particular, the behavior of 39% of the players in their experiment who have been clustered in their analysis to clusters 1 and 3 is similar to TFT_c (cluster 1 is closer to the pure counterpart tit-for-tat strategy, while cluster 3 introduces more randomness). The players in their cluster 2 present similar behavior to GR_c and players in their cluster 5 behave similar toAD_c. Finally, the players in their smaller clusters of 6, 7, and 8 behave in a way similar to $STFT_c$ (with cluster 6 somewhat closer to the similar strategy of suspicious-grim, cluster 8 close to the pure counterpart strategy of $STFT$, and cluster 7 introducing a bit more randomness). The only, relatively small, inconsistency between our model's predictions and their result is that their data does not include a cluster corresponding to AC_c, and instead have cluster 4, which includes random behavior that is somewhat difficult to interpret.

Romero & Rosokha (2023, Figure 3) present data about the performance of different strategies in their setup. They show that the players clustered to the more cooperative strategies of TFT and to GR in their analysis achieved the highest average payoffs, which is consistent with our result for low values of SizeBad close to 0.

7 Conclusion

We apply a new approach of k-mean cluster analysis to revisit the question of evaluating the strategies employed in laboratory experiments of the infinitely-repeated prisoner's dilemma. Our analysis reveals that subjects can be grouped into five distinct clusters, with each cluster displaying behavior that is relatively similar to a pure strategy, but with some systematic deviations that are important in understanding and predicting players' behavior:

1. 30% of the players almost always defect, with a slightly lower probability of defection if they cooperated in the previous round.

2. 25% of the players play tit-for-tat, except that occasionally they play their own action in the previous round instead of the opponent's action.

3. 20% of the players play suspicious-tit-for-tat, with the same systematic deviation as for tit-for-tat.

4. 20% of the players play grim, except that they occasionally (1) defect in the first round, and (2) cooperate after a unilateral opponent's defection.

5. 5% of the players almost always cooperate, with a slightly lower probability of cooperation if either player defected in the previous round.

The above frequencies describe the average values in the entire database; higher values of SizeBad increase the frequencies of the less cooperative strategies (always defect and suspicious-tit-for-tat), and decrease the frequencies of the more cooperative strategies.

Our model improves the predictive power compared to both existing analyses (Dal Bó & Fréchette, 2018; Backhaus & Breitmoser, 2021). Furthermore, our analysis demonstrates a large potential to apply k-means cluster analysis to study experimental data, and, in particular, to estimate players' strategies. In some setups, this can help gain additional insights and, possibly, allow to verify and approve the predictions of the commonly-applied methods of SFEM (Strategy Frequency Estimation Method, developed in Dal Bó & Fréchette, 2011, and later applied, among others, in Bigoni *et al.*, 2015; Aoyagi *et al.*, 2019; Vespa, 2020), and of explicit strategy elicitation (e.g., Romero & Rosokha, 2018, 2023)

Table 4: Frequency of Cooperation and Size of Each Cluster ($k = 1, ..., 7$)

k	% Players	Cooperation Frequency after each memory-1 history $(\sigma_0, \sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd})$	Mean Cooperation Frequency	Cluster Name (Closest memory-1 pure strategy)
1	100%	(44%, 96%, 31%, 35%, 5%)	41%	-
2	30%	(7%, 27%, 17%, 10%, 2%)	5%	AD_c
	70%	(62%, 97%, 33%, 48%, 7%)	55%	GR_c
3	29%	(7%, 21%, 16%, 10%, 2%)	5%	AD_c
	36%	(60%, 96%, 28%, 25%, 6%)	48%	GR_c
	35%	(65%, 97%, 37%, 77%, 9%)	61%	TFT_c
4	29%	(7%, 18%, 16%, 9%, 2%)	5%	AD_c
	25%	(79%, 96%, 25%, 20%, 7%)	59%	GR_c
	24%	(85%, 98%, 35%, 78%, 8%)	71%	TFT_c
	22%	(23%, 94%, 39%, 52%, 7%)	31%	$STFT_c$
5	29%	(7%, 18%, 16%, 9%, 2%)	5%	AD_c
	21%	(74%, 96%, 26%, 16%, 7%)	56%	GR_c
	25%	(85%, 98%, 29%, 75%, 6%)	68%	TFT_c
	20%	(22%, 95%, 39%, 52%, 7%)	31%	$STFT_c$
	5%	(80%, 94%, 61%, 56%, 61%)	82%	AC_c
6	28%	(7%, 12%, 16%, 8%, 2%)	4%	AD_c
	16%	(86%, 96%, 28%, 19%, 7%)	65%	GR_c
	24%	(87%, 98%, 28%, 74%, 6%)	69%	TFT_c
	14%	(26%, 95%, 38%, 76%, 9%)	37%	$STFT_c$
	5%	(87%, 94%, 62%, 55%, 60%)	84%	AC_c
	13%	(28%, 91%, 32%, 21%, 6%)	25%	SGR_c
7	26%	(6%, 11%, 16%, 4%, 2%)	4%	AD_c
	16%	(87%, 96%, 28%, 19%, 7%)	65%	GR_c
	24%	(87%, 98%, 27%, 74%, 6%)	69%	TFT_c
	12%	(26%, 97%, 43%, 75%, 8%)	37%	$STFT_c$
	4%	(88%, 95%, 644%, 56%, 57%)	84%	AC_c
	13%	(27%, 92%, 31%, 20%, 5%)	25%	SGR_c
	5%	(23%, 40%, 20%, 60%, 12%)	25%	–

Table 5: The Five Clusters in the Repeated Prisoner's Dilemma

	% Players for $k = 5$	Cooperation Frequency after each memory-1 history $(\sigma_0, \sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd})$	Mean Coop. Frequency
AD_c	29% $(28-29\% \ \forall 2 \leq k \leq 7)$	$(7\%, 18\%, 16\%, 8\%, 2\%)$ Coop. $< 18\% \ \forall$ history & $\forall 4 \leq k \leq 7$	5%
$STFT_c$	20% $(12-22\% \ \forall 4 \leq k \leq 7)$	$(22\%, 95\%, 39\%, 52\%, 7\%)$ essentially the same $\forall 4 \leq k \leq 7$	31%
GR_c	21% $(16-25\% \ \forall 4 \leq k \leq 7)$	$(74\%, 96\%, 26\%, 16\%, 7\%)$ σ_0 increases to 86% at $k = 6$; all else is essentially the same $\forall 4 \leq k \leq 7$	56%
TFT_c	25% $(24-25\% \ \forall 4 \leq k \leq 7)$	$(85\%, 98\%, 29\%, 76\%, 6\%)$ σ_{cd} increases to 35% at $k = 6$; all else is essentially the same $\forall 4 \leq k \leq 7$	68%
AC_c	5% $(4-5\% \ \forall 5 \leq k \leq 7)$	$(80\%, 94\%, 61\%, 56\%, 61\%)$ essentially the same $\forall 5 \leq k \leq 7$	82%

The numbers in the first line in each cell describe the main specification of $k = 5$ clusters, while the remaining lines describe how these values are affected when changing k between 4 and 7.

Table 6: Comparing the Clusters with Dal Bó & Fréchette (2018)

	% Players	Cooperation Frequency after each memory-1 history $(\sigma_0, \sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd})$	Cluster Name
Dal Bó & Fréchette (2018)	29%	$(4\%, 4\%, 4\%, 4\%, 4\%)$	AD
	20%	$(22\%, 78\%, 22\%, 78\%, 22\%)$	$STFT$
	22%	$(91\%, 91\%, 9\%, 9\%, 9\%)$	GR
	26%	$(90\%, 90\%, 10\%, 90\%, 10\%)$	TFT
	3%	$(96\%, 96\%, 96\%, 96\%, 96\%)$	AC
Our Clusters and the systematic deviations from their pure counterparts	29%	$(7\%, 18\%, 16\%, 8\%, 2\%)$	AD_c
	A player who cooperated in previous round is somewhat more likely to cooperate again		
	20%	$(22\%, 95\%, 39\%, 52\%, 7\%)$	$STFT_c$
	Weaker response to the opponent's action when the actions differ		
	21%	$(74\%, 96\%, 26\%, 16\%, 7\%)$	GR_c
	Players are somewhat less likely to (1) cooperate in the first round, and (2) to defect following a unilateral opponent's defection		
	25%	$(85\%, 98\%, 29\%, 75\%, 6\%)$	TFT_c
	Weaker response to the opponent's action when the actions differ		
	5%	$(80\%, 94\%, 61\%, 56\%, 61\%)$	AC_c
	Tendency to defect if either player defected in the previous round		

Table 7: Log-likelihood and Completeness Comparison with Dal Bó & Fréchette (2018)

Analyzed Data	Clustering to 5 Pure Strategies of Dal Bó & Fréchette (2018)	Our Clustering (k-means with k=5)
All data	46,892 (77.8%)	41,668 (85.3%)
1^{st} half of each session	26,049 (75.1%)	23,789 (81.5%)
2^{nd} half of each session	16,852 (80.6%)	15,385 (84.3%)

The number describes the log-likelihood, and the percentage point is the completeness measure.

Table 8: The 3 Clusters induced by Backhaus & Breitmoser (2021)

	% Players	Cooperation Frequency after each memory-1 history $(\sigma_0, \sigma_{cc}, \sigma_{cd}, \sigma_{dc}, \sigma_{dd})$	Cluster Name
Backhaus & Breitmoser (2021)	24%	$(2\%, 2\%, 2\%, 2\%, 2\%)$	AD_B
	40%	(89%, 98%, 37%, 37%, 6%)	CSG_B
	36%	(30%, 89%, 39%, 39%, 8%)	SSG_B

Table 9: Log-likelihood Comparison with Backhaus & Breitmoser (2021)

Analyzed data	Clustering to the 3 Strategies of Backhaus & Breitmoser (2021)	Our clustering (k-means)		
		$k = 3$	$k = 4$	$k = 5$
All data	43,065 (83%)	46,015 (79%)	42,341 (84.3%)	41,668 (85.3%)
1^{st} half	24,723 (78.9%)	25,599 (76.4%)	24,310 (80.0%)	23,789 (81.5%)
2^{nd} half	16,033 (82.7%)	18,181 (77.2%)	15,773 (83.3%)	15,385 (84.3%)

The number describes the log-likelihood, and the percentage point is the completeness measure.

Table 10: Cooperation Frequency and # of Occurrences of each Memory-1 History

Clustered strategy	Memory-1 History				
	1st round	cc	cd	dc	dd
AD_c	7% (12,481)	18% (258)	16% (957)	9% (3,653)	2% (21,445)
$STFT_c$	22% (8,355)	95% (4,822)	39% (2,603)	52% (2,908)	7% (13,222)
GR_c	74% (7,879)	96% (10,858)	26% (2,500)	16% (2,165)	7% (8,201)
TFT_c	85% (9,055)	98% (15,197)	29% (4,158)	75% (1,814)	6% (7,328)
AC_c	80% (1,320)	94% (3,260)	61% (787)	56% (467)	61% (409)
Total	44% (39,090)	96% (34,395)	31% (11,005)	35% (11,007)[7]	5% (50,305)
All except AD_c	62% (26,609)	97% (34,137)	33% (10,048)	48% (7,354)	7% (29,160)

Table 11: Comparison with Romero & Rosokha (2023, Figure 2)

Our clustered strategy	Frequency in our model for $SizeBad = 10\%$	Romero & Rosokha's (2023)main treatment	
		Frequency	Clusters
TFT_c	35%	39%	1+3
GR_c	25%	19%	2
AD_c	10-15%	12%	5
$STFT_c$	10-15%	17%	6+7+8
AC_c	10-15%	-	
	-	13%	4

References

Aoyagi, Masaki, & Fréchette, Guillaume. 2009. Collusion as public monitoring becomes noisy: Experimental evidence. *Journal of Economic theory*, **144**(3), 1135–1165.

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

Backhaus, Teresa, & Breitmoser, Yves. 2021. *Inequity aversion and limited foresight in the repeated prisoner's dilemma*. Tech. rept. Center for Mathematical Economics Working Papers.

Bigoni, Maria, Casari, Marco, Skrzypacz, Andrzej, & Spagnolo, Giancarlo. 2015. Time horizon and cooperation in continuous time. *Econometrica*, **83**(2), 587–616.

Bigoni, Maria, Casari, Marco, Salvanti, Andrea, Skrzypacz, Andrzej, & Spagnolo, Giancarlo. 2022. It's Payback time: new insights on cooperation in the repeated prisoners' dilemma.

Blonski, Matthias, & Spagnolo, Giancarlo. 2015. Prisoners' other dilemma. *International Journal of Game Theory*, **44**(1), 61–81.

Blonski, Matthias, Ockenfels, Peter, & Spagnolo, Giancarlo. 2011. Equilibrium selection in the repeated prisoner's dilemma: Axiomatic approach and experimental evidence. *American Economic Journal: Microeconomics*, **3**(3), 164–92.

Breitmoser, Yves. 2015. Cooperation, but no reciprocity: individual strategies in the repeated prisoner's dilemma. *American Economic Review*, **105**(9), 2882–2910.

Bruttel, Lisa, & Kamecke, Ulrich. 2012. Infinity in the lab. How do people play repeated games? *Theory and Decision*, **72**, 205–219.

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

Camera, Gabriele, Casari, Marco, & Bigoni, Maria. 2012. Cooperative strategies in anonymous economies: An experiment. *Games and Economic Behavior*, **75**(2), 570–586.

Cooper, David J, & Kagel, John H. 2022. Using team discussions to understand behavior in indefinitely repeated prisoner's dilemma games. *American Economic Journal: Microeconomics*.

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

Dal Bó, Pedro, & Fréchette, Guillaume R. 2011. The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, **101**(1), 411–29.

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

Dal Bó, Pedro, & Fréchette, Guillaume R. 2019. Strategy Choice in the Infinitely Repeated Prisoner's Dilemma. *American Economic Review*, **109**(11), 3929–52.

Dreber, Anna, Rand, David G, Fudenberg, Drew, & Nowak, Martin A. 2008. Winners don't punish. *Nature*, **452**(7185), 348–351.

Dreber, Anna, Fudenberg, Drew, & Rand, David G. 2014. Who cooperates in repeated games: The role of altruism, inequity aversion, and demographics. *Journal of Economic Behavior & Organization*, **98**, 41–55.

Duffy, John, & Ochs, Jack. 2009. Cooperative behavior and the frequency of social interaction. *Games and Economic Behavior*, **66**(2), 785–812.

Duffy, John, Hopkins, Ed, & Kornienko, Tatiana. 2021. *Facing the Grim Truth: Repeated Prisoner's Dilemma Against Robot Opponents*. Tech. rept. Working Paper.

Ellison, Glenn. 1994. Cooperation in the prisoner's dilemma with anonymous random matching. *The Review of Economic Studies*, **61**(3), 567–588.

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

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

Friedman, Daniel, & Oprea, Ryan. 2012. A continuous dilemma. *American Economic Review*, **102**(1), 337–363.

Fudenberg, Drew, Rand, David G, & Dreber, Anna. 2012. Slow to anger and fast to forgive: Cooperation in an uncertain world. *American Economic Review*, **102**(2), 720–49.

Fudenberg, Drew, Kleinberg, Jon, Liang, Annie, & Mullainathan, Sendhil. 2022. Measuring the completeness of economic models. *Journal of Political Economy*, **130**(4), 956–990.

Fujiwara-Greve, Takako, & Okuno-Fujiwara, Masahiro. 2009. Voluntarily separable repeated prisoner's dilemma. *The Review of Economic Studies*, **76**(3), 993–1021.

Gill, David, & Rosokha, Yaroslav. 2020. Beliefs, learning, and personality in the indefinitely repeated prisoner's dilemma. *Available at SSRN 3652318*.

Harsanyi, John, & Selten, Reinhard. 1988. *A General Theory of Equilibrium Selection in Games*. Tech. rept. The MIT Press.

Heller, Yuval, & Mohlin, Erik. 2018. Observations on cooperation. *The Review of Economic Studies*, **85**(4), 2253–2282.

Honhon, Dorothée, & Hyndman, Kyle. 2020. Flexibility and reputation in repeated prisoner's dilemma games. *Management Science*, **66**(11), 4998–5014.

Kagel, John H, & Schley, Dan R. 2013. How economic rewards affect cooperation reconsidered. *Economics Letters*, **121**(1), 124–127.

Kandori, Michihiro. 1992. Social norms and community enforcement. *The Review of Economic Studies*, **59**(1), 63–80.

Kasberger, Bernhard, Martin, Simon, Normann, Hans-Theo, & Werner, Tobias. 2023. Algorithmic Cooperation. *Available at SSRN 4389647*.

Larose, Daniel T, & Larose, Chantal D. 2014. *Discovering knowledge in data: an introduction to data mining*. Vol. 4. John Wiley & Sons.

Mengel, Friederike. 2018. Risk and Temptation: A Meta-study on Prisoner's Dilemma Games. *The Economic Journal*, **128**(616), 3182–3209.

Mengel, Friederike, Orlandi, Ludovica, & Weidenholzer, Simon. 2022. Match length realization and cooperation in indefinitely repeated games. *Journal of Economic Theory*, **200**, 105416.

Mullainathan, Sendhil, & Spiess, Jann. 2017. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, **31**(2), 87–106.

Proto, Eugenio, Rustichini, Aldo, & Sofianos, Andis. 2019. Intelligence, personality, and gains from cooperation in repeated interactions. *Journal of Political Economy*, **127**(3), 1351–1390.

Rand, David G, Fudenberg, Drew, & Dreber, Anna. 2015. It's the thought that counts: The role of intentions in noisy repeated games. *Journal of Economic Behavior & Organization*, **116**, 481–499.

Romero, Julian, & Rosokha, Yaroslav. 2018. Constructing strategies in the indefinitely repeated prisoner's dilemma game. *European Economic Review*, **104**, 185–219.

Romero, Julian, & Rosokha, Yaroslav. 2023. Mixed Strategies in the Indefinitely Repeated Prisoner's Dilemma. *mimeo*.

Sherstyuk, Katerina, Tarui, Nori, & Saijo, Tatsuyoshi. 2013. Payment schemes in infinite-horizon experimental games. *Experimental Economics*, **16**(1), 125–153.

Umargono, Edy, Suseno, Jatmiko Endro, & Gunawan, SK Vincensius. 2020. K-means clustering optimization using the elbow method and early centroid determination based on mean and median formula. *Pages 121–129 of: The 2nd International Seminar on Science and Technology (ISSTEC 2019)*. Atlantis Press.

Umuhoza, Eric, Ntirushwamaboko, Dominique, Awuah, Jane, & Birir, Beatrice. 2020. Using unsupervised machine learning techniques for behavioral-based credit card users segmentation in africa. *SAIEE Africa Research Journal*, **111**(3), 95–101.

Vassilvitskii, Sergei, & Arthur, David. 2006. k-means++: The advantages of careful seeding. *Pages 1027–1035 of: Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*.

VESPA, EMANUEL. 2020. An experimental investigation of cooperation in the dynamic common pool game. *International Economic Review*, **61**(1), 417–440.

YUAN, CHUNHUI, & YANG, HAITAO. 2019. Research on K-value selection method of K-means clustering algorithm. *J*, **2**(2), 226–235.

Online Appendices

A Classification to pure Memory-1 Strategies

In this section we briefly present an analysis that classifies each subject to the pure strategy that best predicts her behavior. Specifically, we classify each of the 1,734 subjects to one of the $32 = 2^5$ pure strategy that predicts correctly the subject's play in the highest share of rounds played the subject played. If multiple pure strategies shared the same highest prediction rate, we have broken the tie by dividing the player's weight equally between these strategies.[9] Figure A.1 shows the distribution of the pure strategies that had the highest prediction rate. Observe that 4 out of the 5 most frequent pure strategies in this prediction-based classification are the pure strategies closest to the 4 largest clustered strategies in our main text analysis (namely, GR, TFT, AD, $STFT$), and the fifth one (SGR) presents similar behavior to $STFT$. The two pure strategies only differ in their play following history dc (in which the player was the sole defector in the previous round). Our clustered strategy $STFT_c$ lies between the two pure strategies of $STFT$ and SGR, slightly closer to the former (as the mean frequency of cooperation after observing dc is 52%, slightly above half).

Figure A.2 shows the distribution of these highest memory-1 pure prediction rates for the different subjects. Observe that the behavior of 15% of the subjects is exactly predicted in all of their supergames by a pure memory-1 pure strategy. Moreover, the behavior of 63% of the subjects is predicted by a memory-1 pure strategy with a rate of at-least 90%. This is in line with Dal Bó & Fréchette's (2018) findings about the good experimental fit of pure memory-1 strategies. Having said that, out main text analysis suggests that some systematic deviations from the pure strategies seem to play a substantial role in understanding the subjects' behavior.

B Robustness to Removing Early Supergames

The analyses in the exiting literature often either ignore the data in the first supergames (see, e.g., Blonski *et al.*, 2011), or give more focus to behavior in the second halves of the experimental treatments (see, e.g., Backhaus & Breitmoser,

[9] 55% of the subjects had a single strategy with the highest prediction rate. The remaining 45% of the subjects had ties between multiple strategies that share the highest prediction rate: 25% of them had a tie between 2 strategies, 16% a tie between 3 strategies, and 4 players had a tie between 4 or more strategies.

Figure A.1: The distribution of memory-1 strategies that best fit behavior

str domine2.png

Figure A.2: The distribution of memory-1 strategies that best fit behavior

fitting str.png

2021). This is done because subjects' behavior might be more noisy in the early supergames before the players had the time to fully understand the experimental environment, and to learn how to play. By contrast, our main-text analysis relies on the whole data, including the first supergames. In this appendix we show the robustness of our main results to removing up to 50% of the early supergames played in each treatment.

Specifically, we have redone our main text analysis (namely, the k-means classification with 5 clusters), while removing 10%-50% of the first supergames in each treatment . Figure B.1 shows the results of this analysis. Specifically, it shows for each clustered strategy (1) the share of players who were classified to this strategy, and (2) the average play of the players clustered in this strategy after each memory-1 history when the analysis is restricted the last 50%, 60%, 70%, 80 and 90% of the supergames in each treatment (rounding up, e.g., if there were 28 supergames in a treatment, we included 26 when restricting to the last 90% of the supergames), and compares it with the main-text analysis of the entire experiential data.

Figure B.1 demonstrates the robustness of our results to removing the initial supergames. The five clustered strategies remain similar in all cases, both in terms of their frequencies as well as in the average behavior in each clustered strategy. Gradually removing up to 50% of the initial supergames has the following, relatively limited, two effects:

1. The share of players playing either the most cooperative or the least cooperative strategies (namely, AD_c and AC_c) somewhat increases by 3-5 percentage points. In addition, both of these strategies move somewhat closer to their pure counterpart: the probability of cooperating somewhat decreases in AD_c, and somewhat increases in AC_c.

2. The share of agents following either $STFT_c$ or GR_C somewhat decrease by 3-6 percentage points.

C Memory-2 Histories

Our main-text analysis focuses entirely on memory-1 strategies. In this appendix we demonstrate the robustness of our analysis to considering the impact of memory-2 histories (i.e., to allow players' behavior to depend on the observed play in the previous two rounds of the supergame, rather than only on the most recent round).

Each of the four non-empty memory-1 histories (namely, cc, cd, dc, dd, where the first letter describe the player's action and the second letter the opponent's action

Figure B.1: The Five Clusters of the Repeated Prisoner's Dilemma

`Last k.png`

The figure shows the frequency of players clustered to each strategy (the top number) and the average behavior of these players after each memory-1 history (as a vector with 5 components showing the probability of cooperation in the first round, after both players cooperating, after the player being the sole cooperator, after the player being the sole defector, and after both players defecting, respectively).

Table 12: Frequency of Cooperation after each Memory-1 and Memory-2 History

Memory-1 History	Memory-2 History	Cooperation Frequency	Difference with Memory-1 Freq.	Number of histories
$\emptyset$ (round 1)		44%	-	39,090
	$(\emptyset, cc)$	94%	-2%	7,107
	(cc, cc)	97%	1%	25,029
	(cd, cc)	90%	-6%	1,044
	(dc, cc)	86%	-10%	1,045
	(dd, cc)	69%	-27%	170
cc	Sum	96%	-	34,395
	$(\emptyset, cd)$	28%	-3%	4,818
	(cc, cd)	27%	-4%	1,003
	(cd, cd)	32%	1%	1,551
	(dc, cd)	44%	14%	1,854
	(dd, cd)	30%	-1%	1,779
cd	Sum	31%	-	11,005
	$(\emptyset, dc)$	31%	-4%	4,821
	(cc, dc)	40%	5%	1,001
	(cd, dc)	59%	24%	1,854
	(dc, dc)	19%	-16%	1,552
	(dd, dc)	31%	-4%	1,779
dc	Sum	35%	-	11,007
	$(\emptyset, dd)$	8%	3%	8,854
	(cc, dd)	33%	28%	116
	(cd, dd)	8%	3%	3,762
	(dc, dd)	12%	7%	3,762
	(dd, dd)	3%	-2%	33,811
dd	Sum	5%	-	50,305

in the previous round) corresponds to five memory-2 histories. For example, The memory-1 history cc is split into five memory-2 histories: $(\emptyset, cc)$ (where the players are now in round 2 of the supergame and cc was played in round 1), (cc, cc) (where cc was played in the previous two rounds), (cd, cc) (when cd was played in the penultimate round and cc in the previous round), (dd, cc), and (dd, cc).

Table 12 shows the average frequency of cooperation after each memory-1 and memory-2 history. The table shows that the memory-2 histories have a relatively small impact on the aggregate average behavior.[10] Specifically, the median difference between the frequency of cooperation of each memory-2 history and its memory-1 counterpart is 4 percentage points, and for 14 out of the 20 memory-2 histories this

[10] An additional analysis that we did focusing only on the second half of the super games in each treatment yielded similar results.

difference is at most 7 percentage points. The six memory-2 histories for which the cooperation frequency differs from the memory-1 counterpart history by more than 7 percentage points are as follows:

1. The frequencies of cooperation after the memory-2 histories of (dc, cd) and (cd, dc) are higher than after their counterpart memory-1 histories. It might suggest that a "reciprocal" memory-2 history in which each player was the dole defector once, may encourage both players to cooperate (in line of the strategy "payback", recently suggested by Bigoni *et al.*, 2022).

2. The frequencies of cooperation after the memory-2 histories of (dc, cc) and (dc, dc) are lower than after their counterpart memory-1 histories. It might suggest that a player who was the sole defector in the penultimate round, and her opponent cooperated in the last round, might infer that her opponent is likely to be an "unconditionally cooperator", which, in turn, might encourage the player to exploit it by defecting.

3. The frequency of cooperation after the memory-2 history (dd, cc) is substantially lower than after the memory-1 history cc, and, similarly, the frequency of cooperation after the memory-2 history (cc, dd) is substantially higher than after the memory-1 history dd. Both of these two memory-2 histories are very rare - they happen each in less than 0.15% of the choices in the data.

Interestingly, the above impacts of memory-2 histories are different than those reported in Romero & Rosokha (2018), where strategies were elicited by an interface allowing of pure strategies with long memories. The main finding they report is of strategies in which players start defecting after a sequence of at least two periods of mutual cooperation. We do not see evidence for this in our data (players have very high probaiblity of cooperation of 97% after observing two consecutive rounds with mutual cooperation). Moreover, none of the three impacts of memory-2 histories discussed above is reported in Romero & Rosokha (2018).

D Frequencies and Mean Normalized Payoffs

Tables 13-14 describe the frequency and the mean normalized payoff (as defined in Section 6.2) of each clustered strategy in each treatment (Table 13 sorts the treatments according to SizeBad, while Table 14 sort them alphabetically, as in Table 3.

Table 13: Frequencies and Normalized Payoffs in Each Treatment (Sorted by SizeBad)

Treat. ID	Size Bad	#Pla -ers	Frequency					Mean Normalized Payoff				
			AD_c	$STFT_c$	GR_c	TFT_c	AC_c	AD_c	$STFT_c$	GR_c	TFT_c	AC_c
AF1	1%	38	0.03	0.21	0.29	0.08	0.39	0.16	0.56	0.54	0.25	0.55
Fea30	5%	48	0.08	0.23	0.44	0.10	0.15	0.10	0.71	0.63	0.43	0.73
DF21	10%	36	0.17		0.33	0.39	0.11	0.38		0.59	0.59	0.79
DO28	11%	102	0.07	0.37	0.26	0.23	0.07	0.18	0.48	0.55	0.38	0.36
Sea32	11%	56	0.13	0.21	0.39	0.21	0.05	0.44	0.62	0.51	0.44	0.62
FY29	13%	50	0.12	0.34	0.40	0.04	0.10	0.22	0.47	0.57	0.30	0.57
DF14	16%	44		0.57	0.36	0.02	0.05		0.57	0.65	0.00	0.39
DF20	16%	164	0.11	0.26	0.43	0.09	0.12	0.26	0.52	0.57	0.38	0.63
BOS3	19%	20	0.35	0.35	0.15	0.15		0.22	0.78	0.20	0.41	
KS31	20%	114	0.20	0.29	0.38	0.11	0.02	0.21	0.46	0.43	0.33	0.50
DF23	22%	168	0.29	0.14	0.21	0.35	0.02	0.27	0.37	0.41	0.38	0.50
BK11	23%	36	0.22	0.28	0.17	0.33		0.53	0.68	0.65	0.45	
Df19	27%	38	0.08	0.34	0.39	0.16	0.03	0.16	0.49	0.52	0.26	1.00
BOS4	29%	20	0.35	0.25	0.10	0.25	0.05	0.79	0.72	0.62	0.69	0.00
D13	31%	42	31%	0.29	0.24	0.21	0.19	0.07	0.44	0.50	0.47	0.56
BOS5	33%	20	0.55	0.15	0.10	0.15	0.05	0.48	0.42	0.20	0.56	0.00
Dea26	33%	22	0.18	0.27	0.36	0.18		0.71	0.54	0.66	0.46	
BOS8	35%	20	0.75	0.15	0.05	0.05		0.60	0.25	0.00	0.49	
D12	35%	60	35%	0.33	0.18	0.30	0.15	0.03	0.46	0.53	0.43	0.40
BOS2	36%	20	0.15	0.40	0.30	0.10	0.05	0.67	0.59	0.61	0.18	0.34
DF15	38%	46	0.24	0.24	0.20	0.33		0.62	0.66	0.42	0.54	
DF22	38%	140	0.30	0.23	0.26	0.17	0.04	0.28	0.44	0.49	0.29	0.66
BOS9	58%	20	0.90			0.10		0.76			0.50	
BOS6	67%	40	0.50	0.10	0.15	0.25		0.63	0.48	0.31	0.48	
Dea27	67%	28	0.29	0.43	0.14	0.14		0.73	0.62	0.13	0.57	
DF17	72%	50	0.52	0.06	0.06	0.34	0.02	0.44	0.30	0.10	0.27	0.05
BOS10	80%	20	1.00					0.78				
DF16	81%	44	0.32	0.05	0.16	0.48		0.65	0.42	0.40	0.57	
DF24	81%	114	0.42	0.08	0.14	0.35	0.01	0.50	0.53	0.43	0.45	0.00
BOS7	100%	20	0.65	0.20	0.05	0.10		0.70	0.48	0.00	0.46	
DF18	100%	44	0.73	0.02	0.02	0.20	0.02	0.57	0.35	0.05	0.60	0.51
DF25	100%	50	0.74	0.02	0.06	0.18		0.61	0.47	0.26	0.53	
Total		1734	29%	20%	21%	25%	5%	0.48	0.43	0.52	0.5	0.56

Table 14: Frequencies and Mean Normalized Payoffs in Each Treatment (Sorted Alphabetically)

Treat. ID	Size Bad	#Pla -ers	Frequency					Mean Normalized Payoff				
			AD_c	$STFT_c$	GR_c	TFT_c	AC_c	AD_c	$STFT_c$	GR_c	TFT_c	AC_c
AF1	1%	38	0.03	0.21	0.29	0.08	0.39	0.16	0.56	0.54	0.25	0.55
BOS2	36%	20	0.15	0.40	0.30	0.10	0.05	0.67	0.59	0.61	0.18	0.34
BOS3	19%	20	0.35	0.35	0.15	0.15		0.22	0.78	0.20	0.41	
BOS4	29%	20	0.35	0.25	0.10	0.25	0.05	0.79	0.72	0.62	0.69	0.00
BOS5	33%	20	0.55	0.15	0.10	0.15	0.05	0.48	0.42	0.20	0.56	0.00
BOS6	67%	40	0.50	0.10	0.15	0.25		0.63	0.48	0.31	0.48	
BOS7	100%	20	0.65	0.20	0.05	0.10		0.70	0.48	0.00	0.46	
BOS8	35%	20	0.75	0.15	0.05	0.05		0.60	0.25	0.00	0.49	
BOS9	58%	20	0.90			0.10		0.76			0.50	
BOS10	80%	20	1.00					0.78				
BK11	23%	36	0.22	0.28	0.17	0.33		0.53	0.68	0.65	0.45	
D12	35%	60	35%	0.33	0.18	0.30	0.15	0.03	0.46	0.53	0.43	0.40
D13	31%	42	31%	0.29	0.24	0.21	0.19	0.07	0.44	0.50	0.47	0.56
DF14	16%	44		0.57	0.36	0.02	0.05		0.57	0.65	0.00	0.39
DF15	38%	46	0.24	0.24	0.20	0.33		0.62	0.66	0.42	0.54	
DF16	81%	44	0.32	0.05	0.16	0.48		0.65	0.42	0.40	0.57	
DF17	72%	50	0.52	0.06	0.06	0.34	0.02	0.44	0.30	0.10	0.27	0.05
DF18	100%	44	0.73	0.02	0.02	0.20	0.02	0.57	0.35	0.05	0.60	0.51
Df19	27%	38	0.08	0.34	0.39	0.16	0.03	0.16	0.49	0.52	0.26	1.00
DF20	16%	164	0.11	0.26	0.43	0.09	0.12	0.26	0.52	0.57	0.38	0.63
DF21	10%	36	0.17		0.33	0.39	0.11	0.38		0.59	0.59	0.79
DF22	38%	140	0.30	0.23	0.26	0.17	0.04	0.28	0.44	0.49	0.29	0.66
DF23	22%	168	0.29	0.14	0.21	0.35	0.02	0.27	0.37	0.41	0.38	0.50
DF24	81%	114	0.42	0.08	0.14	0.35	0.01	0.50	0.53	0.43	0.45	0.00
DF25	100%	50	0.74	0.02	0.06	0.18		0.61	0.47	0.26	0.53	
Dea26	33%	22	0.18	0.27	0.36	0.18		0.71	0.54	0.66	0.46	
Dea27	67%	28	0.29	0.43	0.14	0.14		0.73	0.62	0.13	0.57	
DO28	11%	102	0.07	0.37	0.26	0.23	0.07	0.18	0.48	0.55	0.38	0.36
FY29	13%	50	0.12	0.34	0.40	0.04	0.10	0.22	0.47	0.57	0.30	0.57
Fea30	5%	48	0.08	0.23	0.44	0.10	0.15	0.10	0.71	0.63	0.43	0.73
KS31	20%	114	0.20	0.29	0.38	0.11	0.02	0.21	0.46	0.43	0.33	0.50
Sea32	11%	56	0.13	0.21	0.39	0.21	0.05	0.44	0.62	0.51	0.44	0.62
Total		1734	29%	20%	21%	25%	5%	0.48	0.43	0.52	0.5	0.56