When Are Mixed Equilibria Relevant? *

by Daniel Friedman and Shuchen Zhao, †

This Version: August 23, 2019

Abstract

Mixed Nash equilibria are a cornerstone of game theory, but their empirical relevance has always been controversial. We study in the laboratory two games whose unique NE is in completely mixed strategies; other treatments include the matching protocol (pairwise random vs population mean-matching), whether time is discrete or continuous, and whether players can specify mixtures or only pure strategies. Comparing point predictions, NE always does better than maximin and often does no worse than Logit QRE. NE predicts better than Center (50-50 mixes) under mean-matching, but otherwise not as well. By contrast, in a dominance solvable game, NE predicts better than alternatives in all treatments. Qualitative and quantitative dynamic models capture regularities across all treatments.

JEL Classification: C72, C73, C92.

Keywords: Nash equilibrium, minimax, mixed strategy, directional learning, laboratory experiment.

---

*We are grateful for support that contributed to this paper from National Science Foundation grants SES-1357867 and SES-0925039; for programming assistance from Morgan Grant and James Pettit; to seminar participants at the WZB (Berlin) MC brown bag workshop May 8 and the ESA World Meetings Vancouver July 6, 2019; and for helpful comments from Ciril Bosch, Steffen Huck, Johannes Leutgeb, Joerg Oechssler, and Julian Romero. Joint work with Ryan Oprea and Nat Wilcox circa 2012 helped us formulate some of the ideas and treatments presented below.

†Economics Department, University of California Santa Cruz, CA 95064; dan@ucsc.edu, szhao19@ucsc.edu

"There he goes," said Holmes, as we watched the [special train] carriage swing and rock over the point. "There are limits, you see, to our friend's intelligence. It would have been a *coup-de-maître* had he deduced what I would deduce and acted accordingly."
— Arthur Conan Doyle (1893)

1 Introduction

Generalized matching pennies games capture the essence of strategic situations (e.g., in hunting, warfare and sports) where the central task for each participant is to outguess opponents. For example, in the epigraph above, Sherlock Holmes gloats that his own level-3 strategy of exiting at Canterbury bested his archenemy Moriarty's level-2 strategy of engaging a special train, but Holmes recognizes that higher levels are possible. Since level-$(k+1)$ beats level-k for every positive k in generalized matching pennies, these games suffer from infinite regress, a Gordian knot that blocked progress in game theory for centuries. Von Neumann (1928) finally cut that knot with the idea of mixed strategy equilibrium.

Although mixed strategy equilibrium remains a cornerstone of game theory, it continues to provoke theoretical and empirical controversies. We will see in the next section that theorists have well reasoned doubts about the predictive power of mixed Nash equilibrium, and have proposed alternatives. Applied economists typically focus on pure strategy Nash equilibria when they exist, but turn to mixed NE in games (such as generalized matching pennies) with no pure NE. The empirical evidence supporting those equilibria, however, is itself mixed at best, as has been recognized for over 50 years.

The present paper is motivated by the following research questions. First, under what conditions (if any) does mixed Nash equilibrium do a good job of predicting behavior in generalized matching pennies games? Second, is there a better point prediction — perhaps maximin, as von Neumann proposed, or quantal response equilibrium? Third, can qualitative or quantitative dynamic models explain behavior when it departs from point predictions? Given the importance and ubiquity of strategic interaction with a matching pennies flavor, the answers to such questions have first order importance for applied social scientists and biologists as well as for theorists.

After reviewing some previous literature in Section 2 and some established theory in Section 3, we describe in Section 4 a fresh laboratory examination of games with unique equilibria in mixed strategies. Using a new graphical player screen display for 2x2 bimatrix games, we present two different matching pennies games plus (as a control) a dominance solvable game. Section 4 concludes with lists of testable hypotheses about competing point predictions, about treatment effects on mean choices and on dispersion, and about adaptive learning dynamics.

Section 5 collects results. Some of the point predictions do better than others in some circumstances, but overall none of them predicts very well. We find that the data are generally consistent with a qualitative directional learning model, and that a quantitative regret-based version of directional learning captures some important regularities.

A concluding discussion in Section 6 summarizes our findings and suggests potential implications for game theory and for applied research. Appendices include supplementary data analysis, and instructions to subjects.

2 Previous literature

Early game theory emphasized two-player zero-sum bimatrix games, where Nash equilibrium and maximin mixed strategies coincide, but recognized that these equilibrium mixes differ in asymmetric matching pennies games (e.g., Solan et al., 2013). Early theoretical work on fictitious play dynamics (Julia Robinson, 1951; George Brown, 1951) established convergence in the zero-sum case, but Shapley (1964) found some non-zero-sum games with a unique NE in mixed strategies to which such dynamics do not converge. Subsequent generations of theorists have not reached consensus on dynamic stability: Stahl (1988), Crawford (1985) and others showed that convergence to equilibrium generally fails for their favored dynamics in asymmetric matching pennies games, while Hofbauer and Hopkins (2005), among others, prove convergence for different sorts of dynamics.

These theoretical controversies, for point predictions as well as for dynamics, highlight the need for empirical work. This was recognized long ago, but so far the empirical results have been mixed at best. Rapoport & Orwant (1962) surveyed early laboratory experiments, and found that average play typically was closer to a uniform mix (e.g., 50-50) than to the NE or maximin mix. O'Neill (1987) found that average empirical mixtures were surprisingly close to the NE in a particular zero-sum 4x4 game, but his results were challenged by (James) Brown and Rosenthal (1990). The subsequent controversy (e.g., Walker and Wooders, 2001; Chiappori et al (2002); Palacios-Huerta, 2003) left many readers with the impression that professionals closely approximate equilibrium mixed strategies but the usual undergrad lab subjects cannot. A closer reading suggests that, outside their familiar environments, professionals are typically no more successful than the usual subjects (Wooders, 2010; Levitt et al, 2010), but that populations of the usual subjects can collectively, if not individually, successfully implement equilibrium mixtures (e.g., Friedman, 1996; Binmore et al, 2001).

We know of only one previous empirical paper comparing maximin to Nash mixtures. Ochs (1995) considers several treatments (including one that uses a set of 9 explicit mixtures) but finds that neither Nash equilibrium nor maximin tracks the observed changes in average play when game parameters change. Goeree, Holt and Palfrey (2003) find that quantal response equilibrium with one free parameter (for logit precision) also fails to track such changes, but adding a second parameter (for risk aversion) improves performance.

There is also an empirical literature on adaptive dynamics in matching pennies games. Mookherjee and Sopher (1994) find that belief learning (responsive to payoffs that would have been earned by strategies not employed) beats rote learning. Erev and Roth (1998) offer a three parameter model to rehabilitate rote learning. Camerer and Ho (1999)'s EWA model includes an extra parameter to hybridize belief learning (a la Friedman and Cheung, 1996) with rote learning; the authors show that it is able to fit a variety of games, including some matching pennies. Tang (1999) presents 3x3 bimatrix game data that favors the Selten (1991) anticipatory dynamics model over the Crawford (1985) model. Stephenson (2019) reports an experimental test of evolutionary models in coordinated attacker-defender games, which include own-population effects (Friedman, 1991) not considered in our generalized matching pennies games. His results are consistent with non-sign-preserving adaptive dynamic models.

In sum, despite important prior work by leading game theorists and experimentalists, all three motivating research questions remain open.

Name	8002	3117	IDDS
Bimatrix	$\begin{pmatrix} 800,0 & 0,200 \\ 0,200 & 200,0 \end{pmatrix}$	$\begin{pmatrix} 300,100 & 100,300 \\ 100,200 & 700,100 \end{pmatrix}$	$\begin{pmatrix} 200,500 & 0,600 \\ 400,300 & 200,100 \end{pmatrix}$
NE	(0.5, 0.2)	(0.33, 0.75)	(0, 1)
Maximin	(0.2, 0.5)	(0.75, 0.67)	(0, 1)

Table 1: Payoff bimatrices and equilibrium mixtures. The notation (a, b) refers to row mixture aTop$\oplus(1-a)$Bottom and column mixture bLeft$\oplus(1-b)$Right.

3 Theory

Point predictions. Table 1 shows the specific bimatrix games that we will study. The first two, named 8002 and 3117 after their row payoffs, are asymmetric matching pennies games. The Appendix includes the straightforward computation of the unique Nash equilibrium (NE) and maximin mixed strategies; these games were chosen in part to create separation between those mixtures. The third game, named IDDS because it is dominance solvable, is intended as a control; its unique NE is in pure strategies.

Figure 1 graphically displays the mixed extension of the 8002 game from the row player's perspective: at mixed strategy profile $(a, b) \in [0, 1]^2$ her payoff is

$$\begin{aligned} f_R(a,b) &= (a, 1-a)\begin{pmatrix} 800 & 0 \\ 0 & 200 \end{pmatrix}\begin{pmatrix} b \\ 1-b \end{pmatrix} = 800ab + 200(1-a)(1-b) \\ &= 1000ab - 200a - 200b + 200. \end{aligned} \tag{1}$$

These payoffs are displayed as colors in a "heat map;" the thermometer bar on the right side shows the corresponding numerical values. These range bi-linearly from 0 at the corners $(a, b) = (0, 1), (1, 0)$ to 200 at (0,0) and to 800 at (1,1). Superimposed on the heatmap are alternative point predictions of empirical average mixtures: Nash equilibrium (NE), Maximin (MM), Center, and the arc of Logit quantal response equilibria (QRE) as the precision parameter ranges from 0 (at Center) to ∞ (at NE).

At least since Nash (1951), game theorists have recognized two distinct interpretations of equilibrium in 2-player games. In the first interpretation, two highly rational individuals, fully aware of each other's circumstances, make choices (possibly mixtures) that they have no incentive to change. In the second, members of a large row player population match

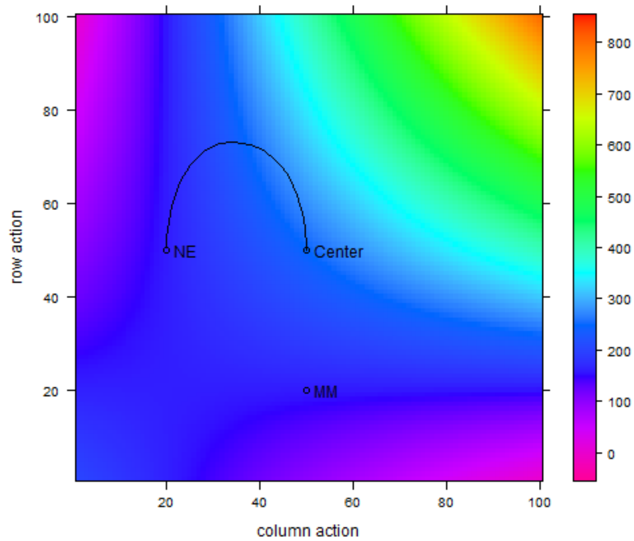

Figure 1: Heatmap for 8002 row player. The color at coordinates (x, y) indicates, via scaled thermometer at right, the row player's expected payoff at mixed strategy profile $100(a, 1-b)$. NE, MM, Center respectively mark the coordinates of Nash equilibrium, maximin and Center profiles. The arc connecting Center $(x, y) = (50, 50)$ to NE includes all Logit quantal response equilibrium (QRE) profiles.

anonymously with members of a large column player population, and the distribution of action profiles remains unchanged as individual players adapt. Binmore et al. (2001), among others, claim that the appropriate dynamic model of how players adapt their choices, and thus the stability of an equilibrium profile, may depend on whether the game is played individualistically or by populations. That claim motivates our mean-matching vs random-pairwise treatments, explained below.

Sign preserving dynamics. In games where each player has only two pure strategies, there is a broad class of adaptive dynamics that applies to both the individualistic and the population interpretations (Friedman, 1991, Weibull, 1997, Friedman and Fung, 1996). The idea is simply that players (individually or collectively) should increase the weight on the pure strategy with currently higher payoff.

To formalize, let the time t (strictly) mixed strategy profile be $(a(t), b(t))$ for a bimatrix game $M = (M^R; M^C)$. For example, in the 8002 game, $M^R = \begin{pmatrix} 800 & 0 \\ 0 & 200 \end{pmatrix}$. For play in continuous time, $(\dot{a}(t), \dot{b}(t))$ denotes time rate of change. The payoff difference between

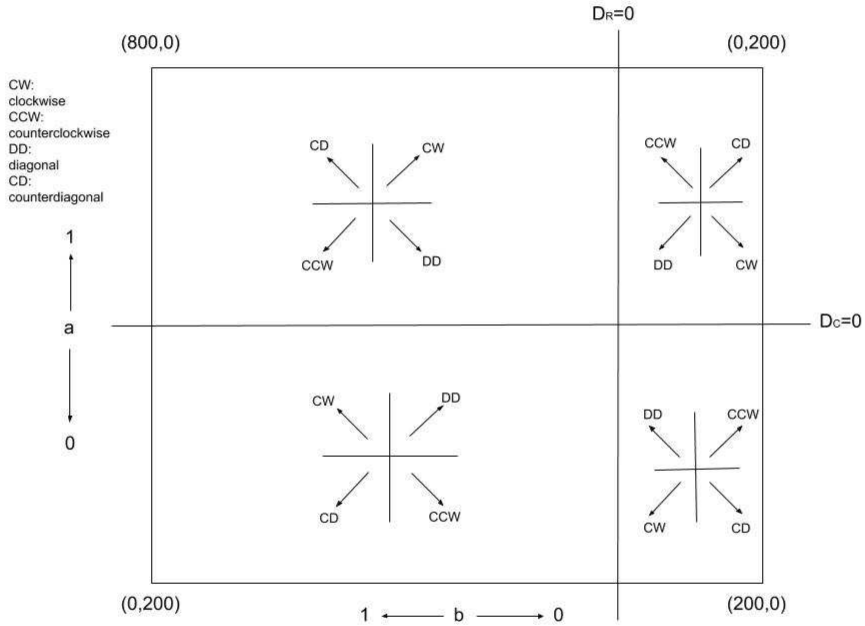

Figure 2: Classifying directional changes in 8002 Matching Pennies. The horizontal (column player's) axis is reversed to be consistent with the bimatrix form in Table 1.

pure strategies is denoted $D_R(t) = (1, -1)M^R \cdot (b(t), 1 - b(t))$ for the row player(s) and $D_C(t) = (1, -1)M^C \cdot (a(t), 1 - a(t))$ for the column player(s).

The dynamic process is *sign preserving* if, at all interior profiles $(a(t), b(t)) \in (0, 1)^2$, we have $\dot{a}(t)D_R(t) > 0$ unless $D_R(t) = 0$, and $\dot{b}(t)D_C(t) > 0$ unless $D_C(t) = 0$. That is, for both row and column players, the weight $a(t)$ or $b(t)$ on the first pure strategy strictly increases (resp. decreases) whenever it has a strictly higher (resp. lower) payoff than the alternative strategy. This is a minimal property of learning and evolution, satisfied by all standard adaptive dynamics including replicator and perturbed best response.

To see the implications, suppose that dynamics are continuous and sign preserving. Draw the isoclines $D_R = 0$ and $D_C = 0$, i.e., the lines for which, respectively, row players and column players are indifferent between their pure strategies. These isoclines divide the state space in $(a(t), b(t)) \in [0, 1]^2$ into regions, each with its own implied direction of change.

Figure 2 illustrates for the 8002 game. From equation (1), $D_R = f(1, b) - f(0, b) = 800b - (200 - 200b) = 1000b - 200$, so the isocline $D_R = 0$ is the vertical line $b = 0.2$.

Similar calculations show that $D_C = 200 - 400a$ so the isocline $D_C = 0$ is the horizontal line $a = 0.5$. These isoclines necessarily intersect at the NE point, and they chop the state space into four rectangles. For example, in the Northeast rectangle $a > 0.5, b < 0.2$, we have $D_R < 0, D_C < 0$, so sign preserving dynamics imply a trajectory with $\dot{a} < 0, \dot{b} < 0$, that is, moving clockwise towards the Southeast rectangle $a < 0.5, b < 0.2$. Similarly, in that Southeast rectangle, sign preserving dynamics imply $\dot{a} < 0, \dot{b} > 0$, moving clockwise towards the Southwest rectangle. Indeed, straightforward calculations show that sign preserving dynamics for the 3117 game as well as the 8002 game imply clockwise moves from each rectangle to the next.

Of course, human subject behavior is noisy, so sign preserving dynamics predicts only that clockwise (CW) will, at least in some treatments in matching pennies games, be the most common direction of change from one observation to the next. Figure 2 depicts the other three possible directions: counterclockwise (CCW), diagonal (DD) towards the Nash equilibrium mix, and counterdiagonal (CD) towards the nearest corner of the state space.

Directional learning model. We now construct a more quantitative model of adaptive dynamics called regret-based directional learning. Let $s_{it} \in [0, 1]$ denote player (or player population) i's mixture at time (subperiod or tick) t, and let $f_i(s_{it}, s_{-it})$ be the corresponding payoff. For example, for row players in the 8002 game, $f_i(s_{it}, s_{-it}) = f_R(a(t), b(t))$. Regret is defined as the normalized shortfall from maximal payoff,[1] $R_{it} = \frac{f_i(\hat{s}_{it}, s_{-it}) - f_i(s_{it}, s_{-it})}{\max_{0 \le x,y \le 1} f_i(x,y)} \ge 0$ for $\hat{s}_{it} \in argmax_x f_i(x, s_{-it})$. The model predicts the change in mixture $\triangle s_{it} = s_{i,t+1} - s_{it}$ as a sign-preserving linear function of regret,

$$\triangle\, s_{it} = \beta_1 R_{it} sign(\hat{s}_{it} - s_{it}) + \epsilon_{it}. \tag{2}$$

The sign function is $sign(x) = +1$ if $x > 0$; $= 0$ if $x = 0$; and $= -1$ if $x < 0$. When $argmax_x f_i(x, s_{-it})$ includes some values larger than s_{it} and other values smaller than s_{it}, then the convention is that $sign\{\hat{s}_{it} - s_{it}\} = 0$.

Alternative specifications we consider in the Appendix include best response learning

$$\triangle\, s_{it} = \beta_1 (\hat{s}_{it} - s_{it}) + \epsilon_{it} \tag{3}$$

[1] The normalization assumes that maximal payoff is positive, as is the case in the bimatrix games used in the present paper. In other cases, the normalization could be dropped, or else all payoffs shifted upward.

and pure directional learning

$$\triangle s_{it} = \beta_1 sign(\hat{s}_{it} - s_{it}) + \epsilon_{it} \tag{4}$$

4 Laboratory Implementation

4.1 Treatment variables

Our experiment has four treatments. The first is the payoff bimatrix: as noted earlier, we consider two generalized matching pennies games, denoted 8002 and 3117, as well as a dominance solvable game denoted IDDS. A second treatment is the action set. In condition P, subjects use radio buttons to select one of two **p**ure strategies, and the display highlights the current payoffs to both players. In condition M, subjects use a vertical slider to select a **m**ixture of the two strategies, as illustrated in Figure 3, and the heatmap display indicates the resulting payoffs.

The third treatment concerns time. In the standard discrete time (D) condition, subjects' choices are updated simultaneously at regular time intervals, here 6000 ms. In the continuous time (C) condition, subjects update choices asynchronously in real time, with an imperceptible latency of around 50 ms, and data are recorded every 500 ms. In both conditions, payoffs are accumulated over time, as illustrated in the lower right graph in Figure 3 in condition C. In condition D, the blue area representing payoffs consists of adjoining rectangles of width 6 seconds and height given by the payoff at the chosen profile.

The remaining treatment is the matching protocol. There are always two distinct populations: row players match only with column players and vice-versa. In the standard random pairwise (rp) protocol, each subject interacts directly with only one matched opponent, and subjects are randomly rematched at the beginning of each new period. In the mean matching (mm) protocol, each subject plays against the average choice of all subjects in the other population or, equivalently for bimatrix games, gets the mean payoff over matches with all subjects in the other population. In terms of notation introduced earlier, s_{-it} is the time-t action of a particular randomly assigned opponent in rp, while in mm it is the time-t mean action of all possible opponents.

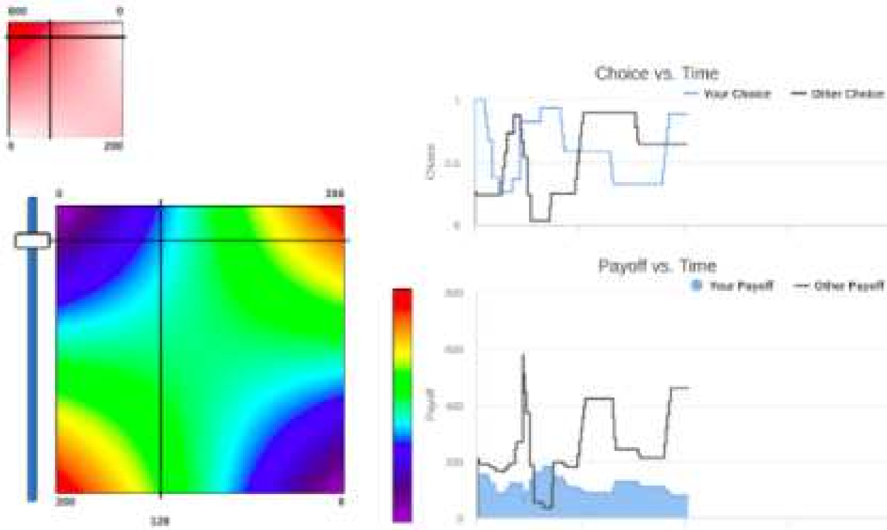

Figure 3: Main features of oTRw screen for MCrp 8002 game. The subjects uses slider at left to adjust her mixture (horizontal line); vertical line shows matched players' current mix. Heatmap color at intersection of these lines codes her current flow payoff; thermometer sets scale. Graph at lower right shows how her flow payoffs accumulate (blue area); black line is matched players' (average) flow payoff. Graph in upper right shows evolution of own and matched players' mixtures. Small red heatmap in upper left shows matched players' payoff function.

4.2 Design

The data analyzed below come from 8 sessions detailed in Table 2. The oTRw software for conducting the experiment is a hybrid of oTree (Chen et al, 2016) and LEEPS lab's Redwood suite, illustrated for the most distinctive treatments in Figure 3. Subjects are recruited from the LEEPS lab subject pool using a local implementation of ORSEE (Greiner, 2015). Each session lasts for around 90 minutes, with a 20-minute instruction/practice stage, 60-minute game play stage and 10-minute payment/closing stage. Average payment is about US $17.

4.3 Testable Hypotheses

Our design transforms the original research questions into the following testable hypotheses.

Date	Treatment	Block Size	# Subjects
4/4/2019	PCrp	6 x 90s periods	8
4/4/2019	MDrp	6 x 90s periods	10
4/5/2019	PDrp	6 x 90s periods	12
4/5/2019	MCrp	6 x 90s periods	8
4/11/2019	PDmm	4 x 150s periods	8
4/11/2019	MCmm	4 x 150s periods	10
4/16/2019	PCmm	4 x 150s periods	8
4/16/2019	MDmm	4 x 150s periods	12

Table 2: Experiment Design. P=pure, M=mixed strategy choice; C=continuous, D=discrete time; rp=random pairwise, mm=mean matching protocol. The following sequence of bimatrices is used in the 5 blocks of each session: 8002, 3117, IDDS, 8002, and 3117.

H1: The time-average observed profile will closely approximate: (H1a) Nash equilibrium, or (H1b) Maximin, or (H1c) Center (.5, .5) , or (H1d) logit quantal response equilibrium for some positive precision parameter.

H2: The time-average observed profile will be closer to Nash equilibrium: (H2a) under mean matching than under random pairwise matching protocol; (H2b) with mixed strategies allowed than with only pure strategies; and (H2c) in continuous time interaction than in discrete time. For other versions of H2, replace the Nash equilibrium point prediction by an alternative such as Maximin.

H3: There will be less dispersion around the time average observed profile: (H3a) under mean matching than under random pairwise matching protocol; (H3b) with mixed strategies allowed than with only pure strategies; and (H3c) in continuous than in discrete time.

We operationalize dispersion as the geometric mean interquartile range. That is, for d_R = 75th percentile - 25th percentile of Row mixes in the sample, and d_C similarly defined for Column mixes, dispersion is defined as $d_G = d_R^{0.5} d_C^{0.5}$. Alternative dispersion measures explored in the Appendix include the harmonic mean, $d_H = (d_R^{-1}/2 + d_C^{-1}/2)^{-1} = \frac{2d_R d_C}{d_R + d_C}$ and the arithmetic mean $\frac{d_R + d_C}{2}$. All measures explored give qualitatively similar results.

H4: In terms of qualitative dynamics, the most frequently observed direction of change will be clockwise (CW) in all treatments in generalized matching pennies games. In pure strategy discrete time (PD) treatments, diagonal (DD) will also be frequently observed.

H5: In the quantitative learning model (2), the coefficient estimate β_1 will be significantly positive in all treatments. That estimate will be more positive for mean matching than for random pairwise matching, and also more positive for pure than for mixed strategy treatments.

5 Results

To gain perspective before reporting hypothesis tests, we examine a few examples raw data. Each panel in Figure 4 displays the time path of action profiles for one instance of each treatment combination in the 8002 bimatrix game. Panel a shows the pure strategy choices of a pair of players in discrete time, the treatment combination most common in previous lab studies. In this instance, the players always best respond to the previous period profile, resulting in a clockwise tour of the four corners of state space. Thus average play is close to the Center, and dispersion is maximal. In Panel b, time is continuous and the (pure) strategy profile is recorded twice per second. The short vertical segments of the time path indicate episodes where both players stayed with their previous strategies, but again the most common change is a clockwise move to the next corner of the state space. In a handful of episodes, both players switch strategies in the same half-second interval and so make a diagonal (DD) move. In panel c, the players can choose explicit mixtures in discrete time, and there is far less dispersion, but it is unclear whether average play is closer to the Center or to NE in this instance. The player pair in Panel d appears to usually move clockwise, sometimes wandering around NE and sometimes wandering away. The next four panels of Figure 4 come from mean-matching (population game) sessions. They all have less dispersion than their random-pairwise counterparts. In particular, the mixed strategy discrete time profile path shown in Panel g is usually in the vicinity of NE.

In testing point predictions, it is appropriate to focus on settled behavior, so subsequent analysis drops first period in each block, and first 18 seconds (or 3 subperiods) of each period. For the remaining part of each remaining period in a given matching, we collapse the time path to its time average profile (m_R, m_C), and look at the distribution over all instances for

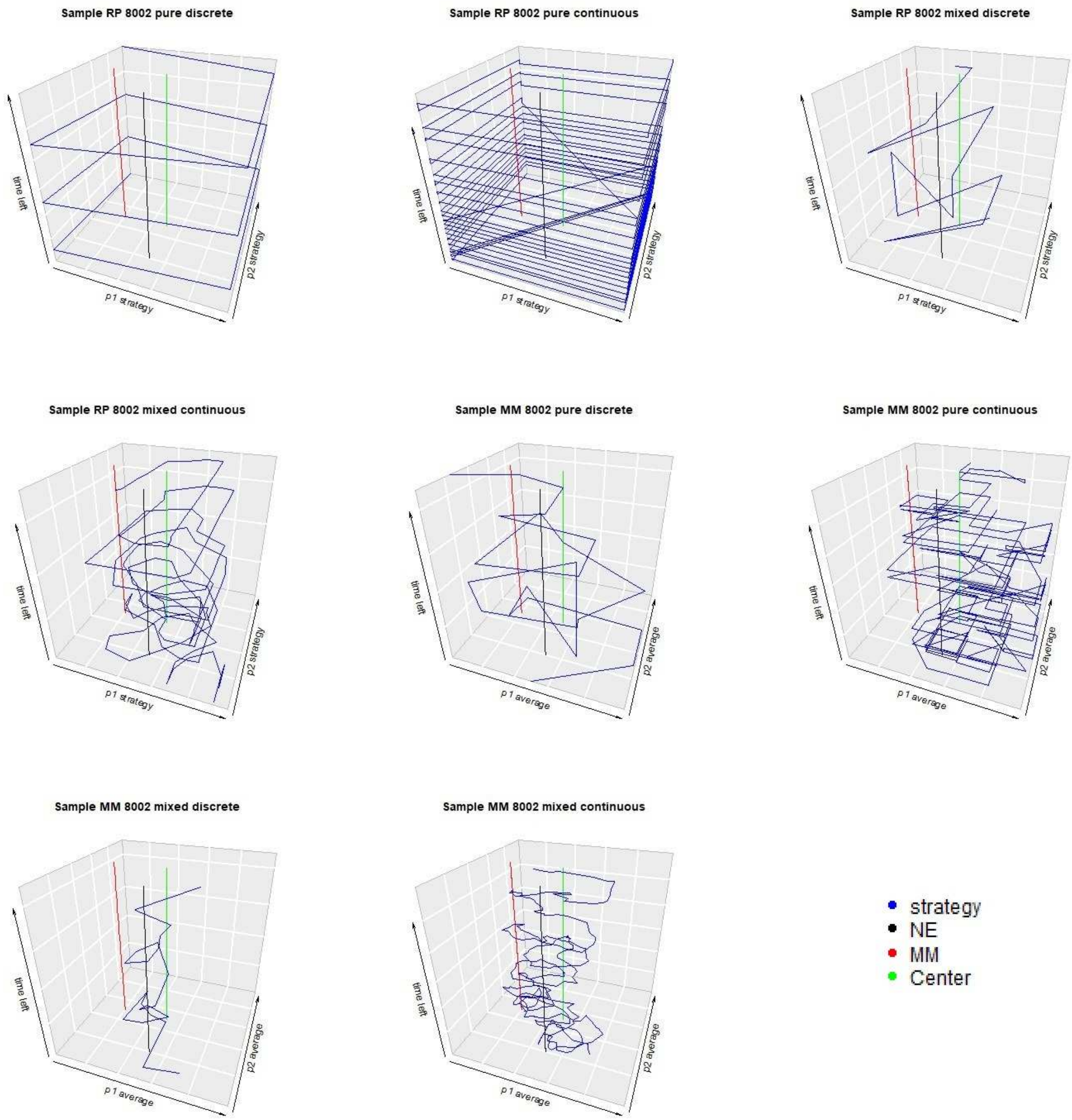

Figure 4: Sample time paths of action profiles. The horizontal plane is the state space, the (a, b) square. The vertical axis is time remaining, so the time paths begin at the top and spiral downward, reaching the bottom plane at the end of the period. Point predictions are time-invariant and therefore appear as vertical lines.

a given treatment combination.

Figure 5 displays the mean and standard deviation of these time averages. For example, the light green box (for the mixed strategy discrete time mean matching data from the 3117 bimatrix) in Panel a easily contains the column NE mix of 0.75 between the mean minus the standard deviation (about 0.55) and the mean plus the standard deviation (about 0.90), while the row NE mix of 0.33 is also within a standard deviation of the row time average mean (roughly 0.2 to 0.55). Comparisons across panels a and b and the various other treatment combinations suggest that maximin is seldom the best predictor of central tendency, while NE predicts well in some cases. More often Center is best, especially in treatment combinations that have large boxes, indicating heterogeneity across instances. By contrast, panel c shows that the IDDS data under all treatments cluster much closer to the NE=MM point (0,1) than to the Center point (.5, .5). Another version of the Figure using median and interquartile instead of mean and standard deviation range appears in the Appendix. It gives qualitatively similar impressions.

5.1 Point Predictions

Table 3 summarizes tests of Hypotheses 1 and Table 4 summarizes tests of Hypotheses 2 and 3; robustness checks using both mean and median profiles can be found in the Appendix. Excluding the initial periods and seconds as noted earlier, for each instance (matching and period) τ, we compute the time average profile (a_τ, b_τ) and its Euclidean distance $[(a_\tau - a_p)^2 + (b_\tau - b_p)^2)]^{0.5}$ from a given point prediction (a_p, b_p). For example, the first line of the Table 3 shows that for mean matching instances in the 3117 game (pooling over Continuous and Discrete time, and over Pure and Mixing action sets), the mean distance in the action space between the time-average profile and the NE prediction is just 0.157, and according to the two sample t-test, this is significantly less than 0.224, the mean distance between those same time average profiles and the Center point. The same line in the table shows that the mean distance to the maximin prediction, 0.398, is significantly larger.

Thus the first lines of Panels A and B in Table 3 support Hypothesis H1a, that NE is the best point prediction, for mean-matching treatments in generalized matching pennies games. Consistent with Hypothesis H1c, Center is best in all other treatments in these games, with

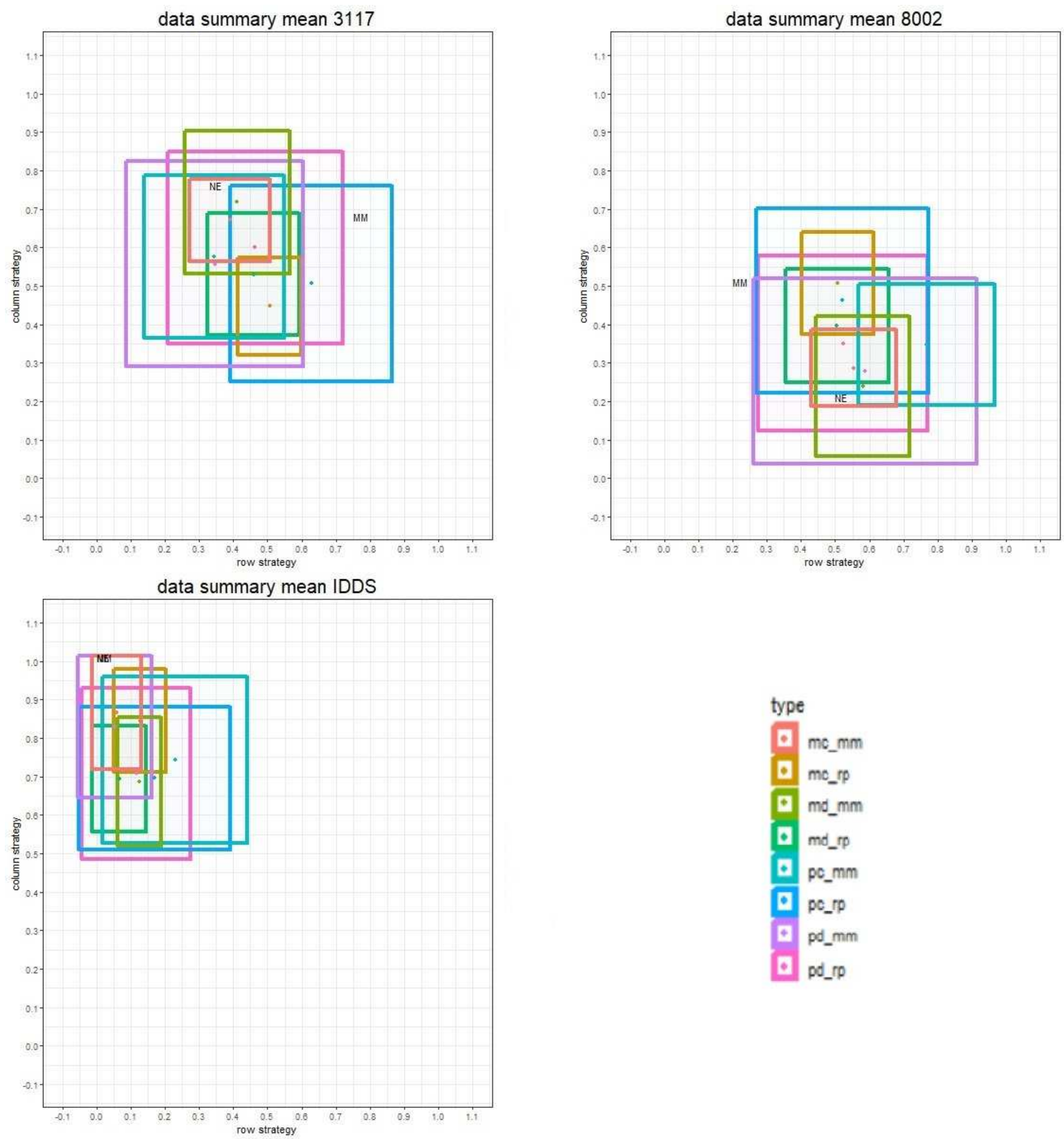

Figure 5: Data summary of population average in 3 games colored by treatments. Dots are mean data. Rectangles are shaped by mean +- standard deviation.

	Distance to NE		Distance to Center		Distance to MM
			Panel A: 3117 games		
mm	0.157	$<_{**}$	0.224	$<_{***}$	0.398
rp	0.313	$>_{***}$	0.124	$<_{***}$	0.295
Mixed	0.242	$>_{***}$	0.135	$<_{***}$	0.339
Pure	0.268	$>_{***}$	0.188	$<_{***}$	0.328
Continuous	0.301	$>_{***}$	0.181	$<_{***}$	0.330
Discrete	0.208	$>_{***}$	0.142	$<_{***}$	0.337
			Panel B: 8002 games		
mm	0.150	$<_{***}$	0.255	$<_{***}$	0.463
rp	0.247	$>_{***}$	0.115	$<_{***}$	0.321
Mixed	0.211	$>_{*}$	0.159	$<_{***}$	0.352
Pure	0.210	$>$	0.176	$<_{***}$	0.397
Continuous	0.263	$>_{***}$	0.156	$<_{***}$	0.375
Discrete	0.159	$<$	0.179	$<_{***}$	0.374
			Panel C: IDDS games		
mm	0.219	$<_{***}$	0.521	$>_{***}$	0.219
rp	0.283	$<_{***}$	0.479	$>_{***}$	0.283
Mixed	0.241	$<_{***}$	0.517	$>_{***}$	0.241
Pure	0.277	$<_{***}$	0.472	$>_{***}$	0.277
Continuous	0.244	$<_{***}$	0.496	$>_{***}$	0.244
Discrete	0.274	$<_{***}$	0.494	$>_{***}$	0.274

Table 3: Mean distance to predictions of time average profiles. Subscripted asterisks indicate p-values of .10, .05 and .01 for t tests of equality between adjacent columns.

the possible exception of discrete 8002, where NE is insignificantly better. In Panels A and B, there is no support for the maximin hypothesis H1b. Panel C confirms that the pure strategy NE (which coincides here with maximin) is a better point prediction than Center in all treatments for the dominance solvable game.

Testing Hypothesis H1d is potentially more complicated, since there is a whole arc of QRE that connect NE to Center, not just a single point prediction. However, as shown in the Appendix and Figure 1, that arc usually bends away from the mean profiles, and the closest point on the arc is typically very close to either NE or Center in all treatments. We therefore conclude that our data do not support Hypothesis H1d.

For the same empirical time average profiles, Table 4 shows the results of regressing treatment dummies and their interactions on prediction error and on dispersion. Hypothesis H2 asserts that relevant predictions are more accurate for certain treatments. The first column of Table 4 supports Hypothesis H2a, that NE is more accurate under mean matching than under random pairwise matching. Conclusions regarding H2b and H2c are more nuanced due to significant interactions: the direct effect of pure strategies and of mean matching both reduce NE prediction error while continuous time increases prediction error, but these are largely offset by the interactions of mean matching with continuous and pure. The upshot is that NE predicts especially well in mixed mean-matching treatments, confirming the impression from the previous Table. The entries in the second column confirm that maximin prediction errors are large in all treatments. Many treatments and interactions have opposite signs in the first and third columns, suggesting that they shift the observed behavior away from NE and towards Center, or the reverse.

Hypothesis 3 concerns dispersion. The last column of Table 4 reports the geometric mean of row dispersion (IQR) and column dispersion (IQR) as defined in the previous section. The second line of the Table supports H3b, that dispersion is less with mixed strategies. The significantly negative coefficient in line 8 offers limited support for H3a: mean matching reduces dispersion in pure strategy treatments but perhaps not in general. The Table does not support H3c; dispersion seems to be about the same for Continuous as for Discrete time treatments.

5.2 Qualitative Dynamics

The large constant term in the last column of Table 4 suggests that behavior typically does not settle down to a behavioral equilibrium. Does that mean that players wander aimlessly, or is there some regularity such as clockwise cycles?

To investigate, recall how Figure 2 classified profile moves $\Delta s_t = (\Delta s_{Rt}, \Delta s_{Ct}) = (s_{Rt+1} - s_{Rt}, s_{Ct+1} - s_{Ct}) \neq 0$ as clockwise (CW), diagonal (DD), counterclockwise (CCW), or counter diagonal (CD). Figure 6 shows how the classifications change over time in random pairwise matching sessions. For example, in the top panels we see that in the 15 six-second Discrete subperiods (14 moves since the cyclical behavior is determined by two consecutive

	Distance to NE	Distance to Maximin	Distance to Center	Dispersion
Treatment DVs:				
continuous	0.10±0.026***	-0.04±0.024	0.02±0.018	-0.09±0.079
pure	-0.06±0.026**	-0.06±0.024**	0.05±0.018***	0.50±0.079***
mm	-0.18±0.028***	0.02±0.026	0.11±0.020***	0.03±0.086
8002	-0.06±0.026**	-0.01±0.024	0.04±0.018**	0.06±0.079
continuous_pure	0.06±0.027**	0.01±0.025	0.04±0.019*	-0.07±0.083
continuous_mm	-0.10±0.028***	0.06±0.026**	0.02±0.020	-0.00±0.086
continuous_8002	0.01±0.027	0.01±0.025	-0.06±0.019***	0.06±0.083
pure_mm	0.15±0.028***	0.11±0.026***	-0.04±0.020**	-0.19±0.086**
pure_8002	-0.03±0.027	0.05±0.025**	-0.04±0.019*	-0.00±0.083
mm_8002	0.06±0.028**	0.04±0.026	0.04±0.020**	-0.08±0.086
Constant	0.28±0.021***	0.34±0.019***	0.08±0.015***	0.48±0.063***
Observations	128	128	128	128
R-squared	0.630	0.538	0.611	0.493

Table 4: Coefficients (and standard errors) for regressions of distance to predictions and of dispersion on treatment dummy variables and interactions. Nominal significance levels 1, 5, and 10% denoted ***,** ,*.

subperiods), there is a preponderance of CW moves (in red), a fair number of DD moves, no CD moves (impossible in Pure treatments), rather few CCW moves, and perhaps 10 - 30% Stay ($\Delta s_t = 0$). Indeed, in all treatments CW is more common than other moves, as predicted by sign preserving dynamics. It is no surprise that Stay is far more common and DD is relatively rare in Continuous treatments, since the time interval there is just half a second. CCW is especially rare in Pure Continuous sessions. CD is rare even in mixed treatments. DD is not uncommon in discrete time treatments, where it may indicate anticipatory behavior in the sense of Selten (1991). There seem to be no strong trends in behavior within periods, nor major differences between 8002 and 3117 games. There is, however, considerable heterogeneity across matched pairs, as can be seen from the by-pair breakdown in Appendix Figure 10.

Figure 7 presents similar evidence for mean matching sessions, where Δs_t represents population profile moves rather than individual pair moves. Not surprisingly, with mean matching we see fewer Stay and more CW moves in most treatments. DD becomes more common while CD and CCW remain rare.

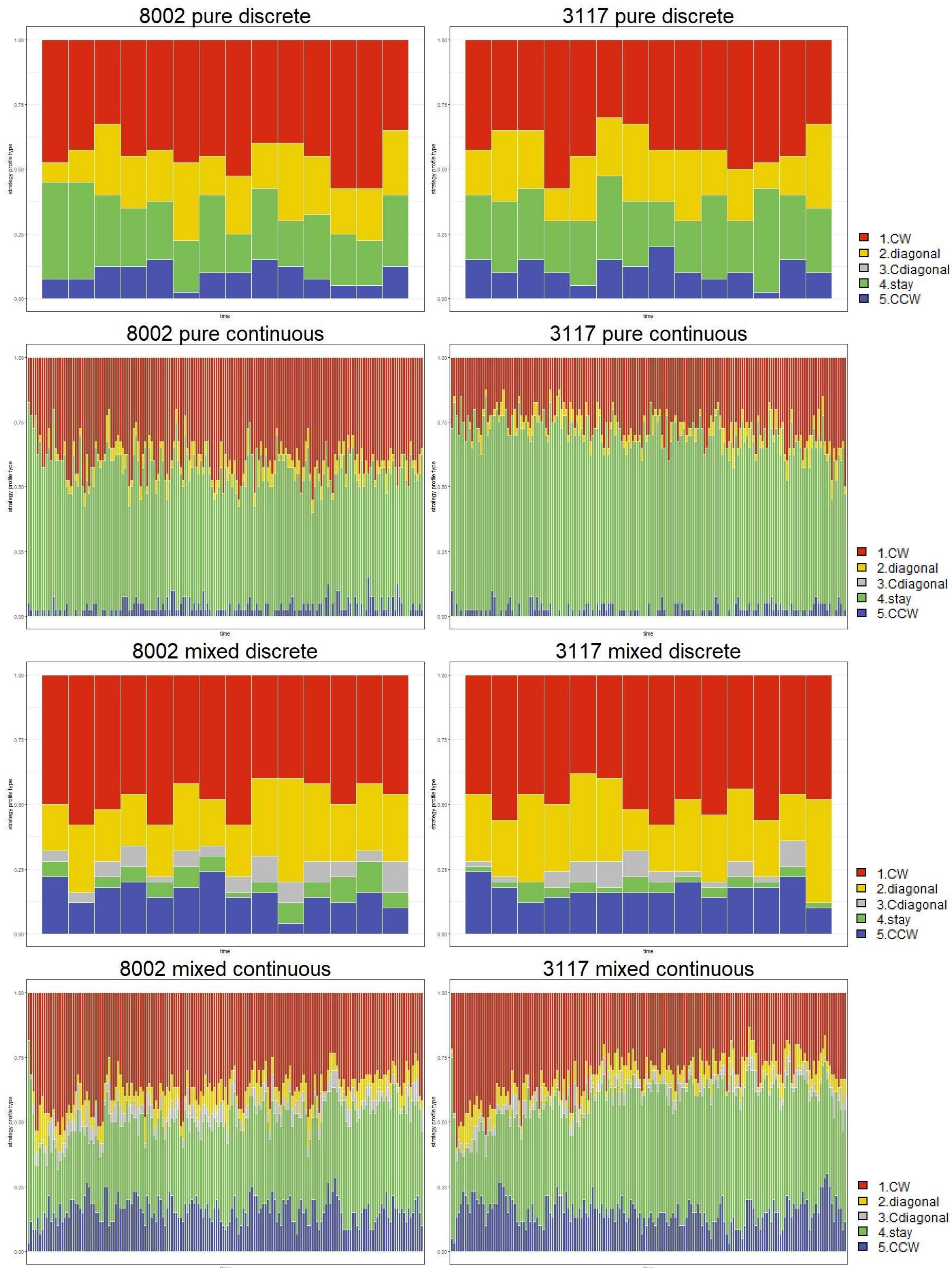

Figure 6: Relative frequency over time of directional moves in random pairwise sessions.

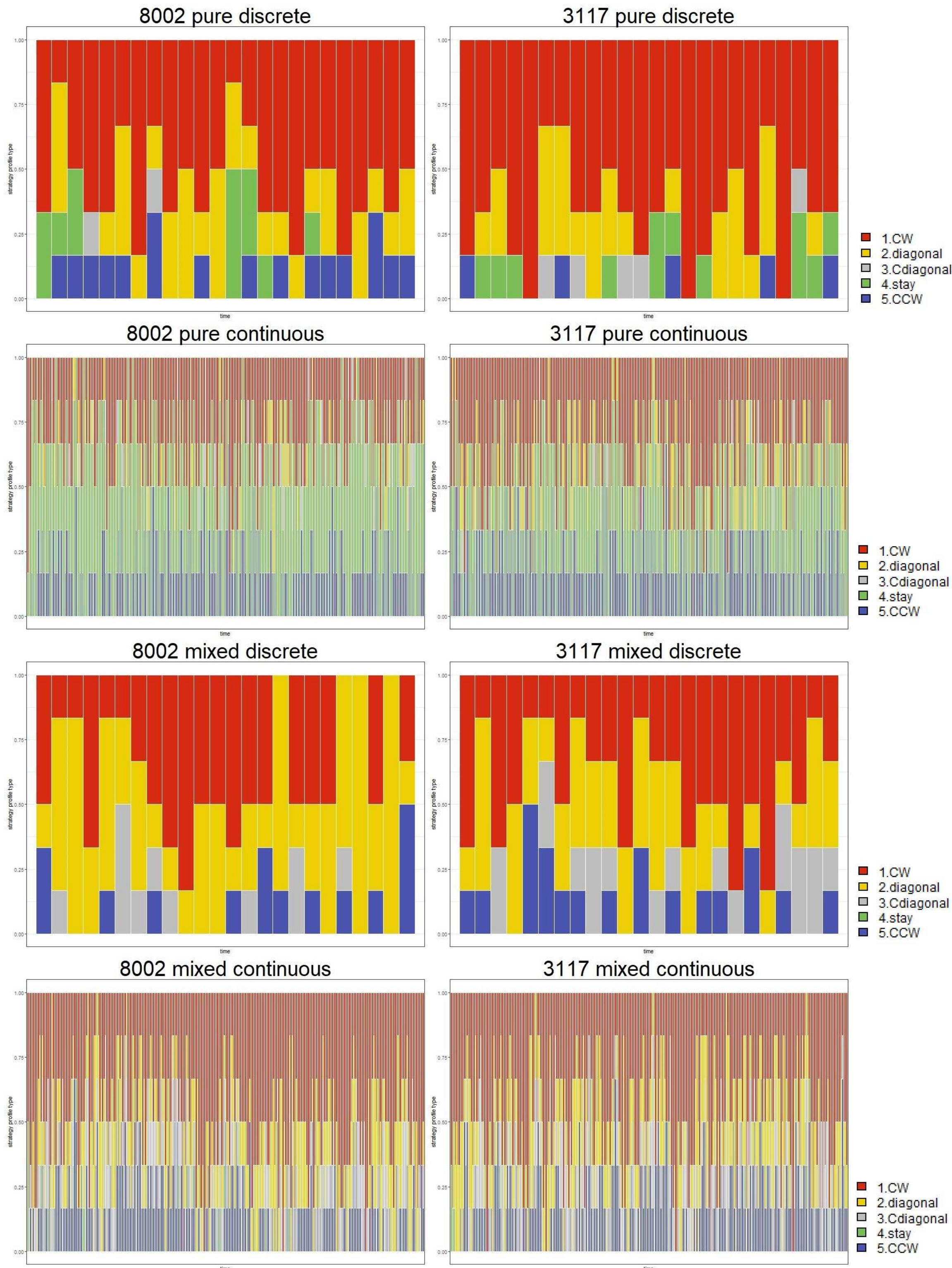

Figure 7: Relative frequency over time of directional moves in mean matching sessions.

	Row-Continuous	Row-Discrete	Col-Continuous	Col-Discrete
β_1	0.12±0.021***	1.14±0.099***	0.26±0.028***	1.06±0.111***
pure	0.47±0.067***	-0.14±0.116	0.55±0.078***	-0.02±0.131
mm	0.24±0.050***	-0.16±0.218	-0.04±0.051	0.33±0.267
8002	-0.03±0.025	-0.22±0.134	-0.17±0.032***	-0.62±0.120***
IDDS	0.04±0.057	0.07±0.244	-0.24±0.031***	0.28±0.319
pure_mm	-0.03±0.153	0.75±0.292**	-0.18±0.151	0.35±0.345
pure_8002	0.10±0.090	0.22±0.155	-0.14±0.087	0.17±0.147
pure_IDDS	-0.21±0.123*	0.98±0.335***	-0.20±0.123	1.12±0.406***
mm_8002	-0.03±0.066	0.00±0.316	0.06±0.061	0.17±0.303
mm_IDDS	-0.13±0.099	-0.12±0.368	0.06±0.055	-0.35±0.450
pure_mm_8002	-0.34±0.195*	0.24±0.390	-0.10±0.164	-0.75±0.384*
pure_mm_IDDS	0.53±0.314*	-0.36±0.472	0.36±0.217*	-0.26±0.626
Observations	79,145	4,995	79,145	4,995
R-squared	0.213	0.337	0.251	0.253
Number of Pairs	415	345	415	345

Table 5: Directional Learning Model (5) coefficient estimates (± standard error) for Row and Col(umn) player actions in Continuous and Discrete time. Least squares with pair and tick fixed effects. Nominal significance levels 1, 5, and 10% denoted ***,** ,* .

The data shown here (and in the Appendix, e.g., Tables 8 and 9) thus support Hypothesis H4. Overall, CW moves are indeed the most prevalent, representing up to half of total observations. DD often ranks second, and other directional moves are relatively rare. Move types have similar distributions in the two generalized matching pennies bimatrices and (if we ignore Stay) in continuous time and discrete time. The distributions also seem roughly similar in pure and mixed strategy conditions and in mean matching and random pairwise.

5.3 Fitted Dynamic Model

To test the more quantitative dynamic hypothesis H5, we fit the regret-based learning model (2) allowing for fixed effects and for treatment-specific response to regret,

$$\triangle\, s_{it} = (\beta_1 + \sum_k \beta_k D_k) R_{it} sign\{\hat{s}_{it} - s_{it}\} + b_i + c_t + \epsilon_{it}. \tag{5}$$

Table 5 collects the results. The first row clearly supports H5: the baseline response to regret β_1 is very significantly positive. Remaining rows show that this support is not reversed by any treatment or interaction considered. H5 also predicts that response is stronger in pure strategy treatments (since moves there must be to corners, not just incremental), and the continuous time data clearly support this prediction, but the impact is insignificant in discrete time. The remaining part of H5 predicts stronger response in mean matching than in random pairwise matching. This prediction is supported for Row players in continuous time sessions, but elsewhere the impact is insignificant.

Other entries in the Table mostly seem reasonable upon reflection. Since continuous time data are sampled twelve times as frequently as discrete time data, it is natural for the continuous time coefficients to be much smaller in absolute value. Column players adjust more slowly in 8002, perhaps because of the greater asymmetry in that game than in the 3117 baseline. Adjustment is faster in IDDS in discrete time with pure strategies, perhaps due to the strategic clarity of that treatment combination. The Appendix reports regressions for related specifications (3) and (4), with results generally consistent with those of Table 5.

To complement the hypothesis tests, we ran simulations of equation 5 using the coefficient estimates reported in Table 5 with error terms set to zero. Figure 8 (cf the 3D version, Figure 11 in the Appendix) shows that, according to the fitted models, players (or populations) move in clockwise cycles that very gradually contract towards a limit cycle surrounding the Nash equilibrium. Thus the data suggest that, practically speaking, there will never be convergence to any point prediction, but rather that (a) cycles will persist for a very long time, and (b) Nash equilibrium is a crude approximation of the long-run time-average profile.

6 Discussion

The hypothesis test results suggest answers to the broad research questions concerning generalized matching pennies, i.e., concerning strategic situations with equilibrium only in mixed strategies. First, we find that mixed Nash equilibrium is a reasonably good predictor of behavior in population games. That is, when players interact with entire groups of other players, not just a single player, our results for matching pennies games suggest that the

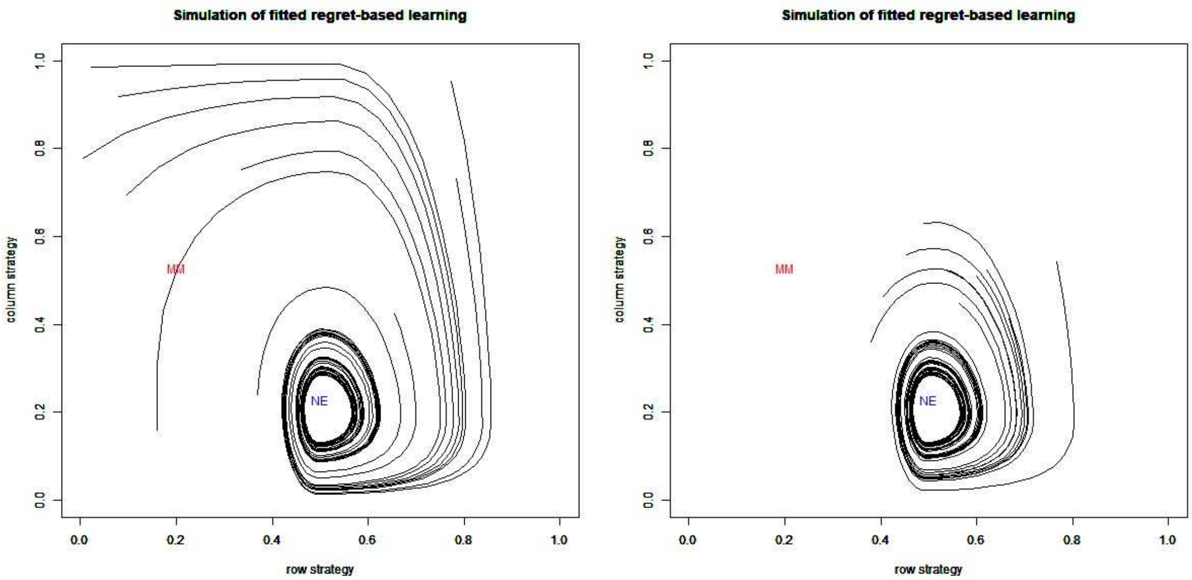

Figure 8: Long time horizon simulation of 5 with parameters fitted for continuous time mixed strategy 8002 games; panel A is for random pairwise matching and B is for mean matching.

players tend to sort themselves out so that overall realized strategies approximate the Nash equilibrium mixture.

Second, popular alternative point predictions, such as maximin or Quantal Response Equilibrium (for a fixed precision parameter) did not improve on Nash equilibrium in any of our treatments. However, the atheoretic prediction Center (all actions equally likely) predicts time-average behavior better than Nash equilibrium under most treatments involving pairwise matches, especially in discrete time and with pure strategy choices only.

Another negative result deserves emphasis. None of the point predictions does well in predicting behavior at a given moment of time in pairwise matchings, due to persistent dispersion around the time average behavior.

The corresponding positive result is that there is order beneath the dispersion. Although it is easier to see in some treatments (e.g., mixed strategies or population means in continuous time) than in others, there is a clear tendency for play to cycle in generalized matching pennies games. Typically one player (or player population) has a stronger incentive to switch strategies, and doing so gives the same incentive to the other player population, creating (with our sign conventions) clockwise cycles. In discrete time treatments we saw some evidence that players tried to anticipate and exploit these regularities, but they nevertheless persist, especially in continuous time and in population games.

Possible future lab experiments involve mixed strategy elicitation and convergence improvements. In our experiments, subjects explicitly choose their mixed strategies and are paid by the expected payoff given the strategy profiles, which removes the need to randomize actions dynamically. Conversely, Romero and Rosokha (2018) elicit subjects' history-dependent actions in repeated prisoners' dilemma. That elicitation allows a closer look at subjects' repeated game strategies and could be applied to matching pennies games. Another direction is to seek new treatments that facilitate convergence to Nash equilibrium or other point predictions. In pilot sessions, we tried adding indicators showing best and worst possible payoffs and slowing adjustment speed in continuous time, but found little impact.

We hope that our work encourages game theorists to take adaptive dynamics more seriously, and to model how they respond to different sorts of treatments such as those considered in this paper. Even more, we hope that our results encourage applied researchers to work in a more nuanced fashion with mixed strategy equilibrium. Biologists since Lotka and Volterra (Lotka, 1925) have recognized that dynamics are crucial to understanding generalized matching pennies interactions such as between predators and prey. Social scientists may benefit from similar thinking. For example, 'hot spot' dispatch of law enforcement resources (e.g., Lazzati and Menichini, 2016) is a generalized matching pennies population game, and our work suggests how adaptive dynamics could supplement equilibrium analysis.

7 Appendix

Computation of NE, Maximin and QRE for 8002 games

In this section we use 8002 games as an example to show the computation of NE, Maximin and QRE curve in Table 1 and Figure 1.

To calculate Nash equilibrium, recall from sign preserving dynamics that we calculated $D_R(t)$ and $D_C(t)$, which show the payoff difference between pure strategies for row and column players, respectively. By definition, the unique mixed Nash equilibrium can be solved by

$$D_R(t) = 1000b - 200 = 0 \tag{6}$$

$$D_C(t) = 200 - 400a = 0 \tag{7}$$

As a result, $(a_{NE}, b_{NE}) = (0.5, 0.2)$.

To calculate Maximin for row players, recall $f_R(a, b)$ from equation (1).

$$f_R(a, b) = 1000ab - 200a - 200b + 200$$

The Maximin problem for row players is the following:

$$\max_a min\{f_R(a, 1), f_R(a, 0)\} \tag{8}$$

For these linear functions of a, the max must occur where $f_R(a, 1) = f_R(a, 0)$, yielding $a_{MM} = 0.2$. Similarly, we can construct the optimization problem for column players and get $b_{MM} = 0.5$.

To calculate QRE curve for both players, the logit payoff response function is 8002 games is as follows.

$$a = \frac{\exp(\lambda f_R(1, b))}{\exp(\lambda f_R(1, b)) + \exp(\lambda f_R(0, b))} \tag{9}$$

$$b = \frac{\exp(\lambda f_C(a, 1))}{\exp(\lambda f_C(a, 1)) + \exp(\lambda f_C(a, 0))} \tag{10}$$

When $\lambda \to 0$, we have $(a, b) = (.5, .5)$. When $\lambda \to \infty$, we have $(a, b) = (a_{NE}, b_{NE}) = (.5, .2)$. The arc curve between two extreme cases in shown in Figure 1

Point prediction details

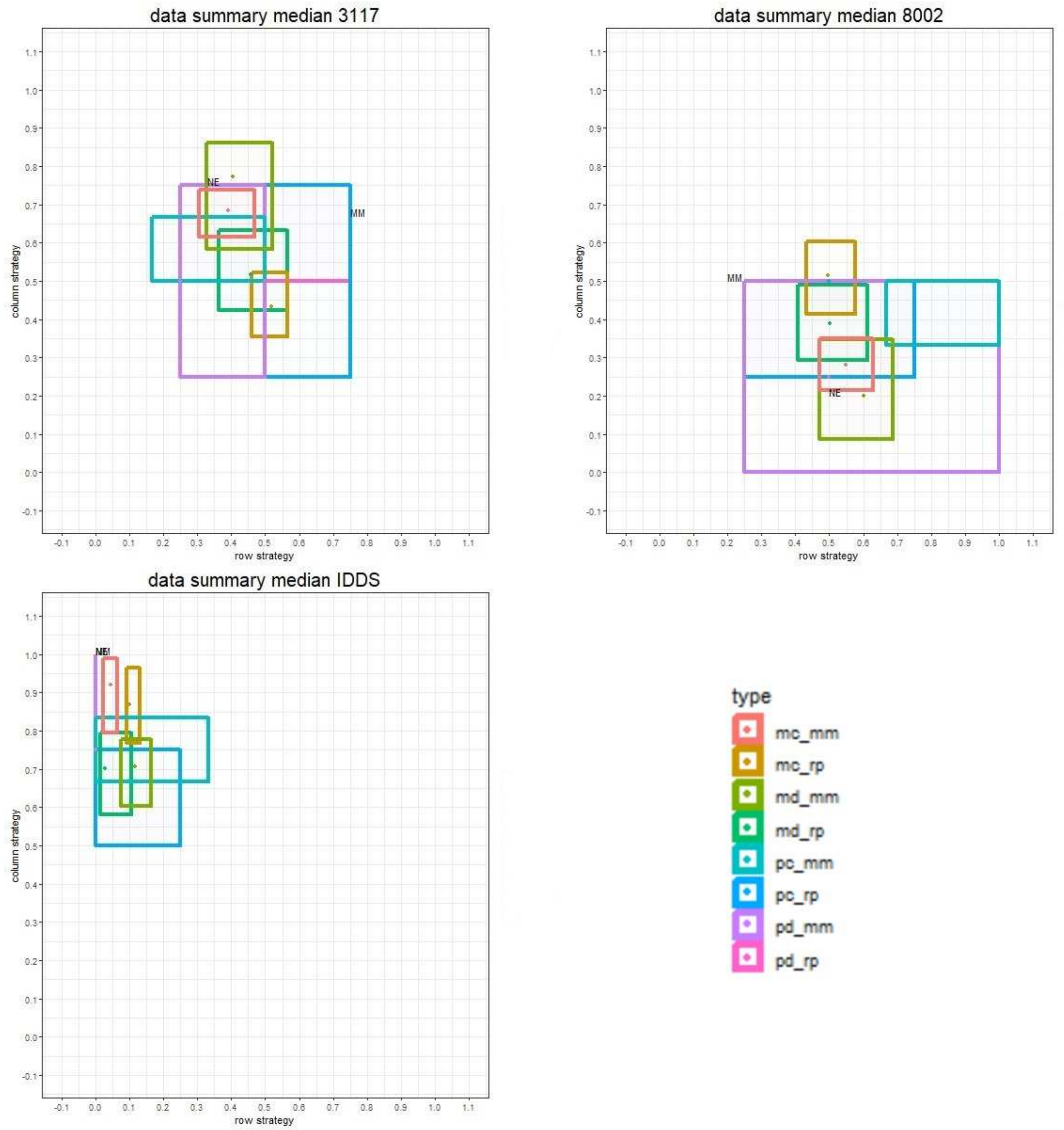

Figure 9: Data summary of population average in 3 games colored by treatments. Dots are median data. Rectangles are bounded by 1st and 3rd quantiles.

Mean of by-period Mean Summary Table

Treatments	row median	column median	To NE	p-value	To Center	p-value	To MM	Harmonic Disp	Geometric Disp
				Panel A: 3117 games					
mm	0.371	0.632	0.157	0.011	0.224	0.000	0.398	0.297	0.601
rp	0.523	0.519	0.313	0.000	0.124	0.000	0.295	0.331	0.669
p-value	0.000	0.000	0.000	-	0.000	-	0.000	0.451	0.442
Mixed	0.461	0.558	0.242	0.000	0.135	0.000	0.339	0.214	0.443
Pure	0.472	0.565	0.268	0.006	0.188	0.000	0.328	0.422	0.844
p-value	0.705	0.813	0.433	-	0.009	-	0.674	0.000	0.000
Continuous	0.498	0.529	0.301	0.000	0.181	0.000	0.330	0.287	0.583
Discrete	0.434	0.593	0.208	0.001	0.142	0.000	0.337	0.349	0.704
p-value	0.027	0.034	0.005	-	0.052	-	0.763	0.143	0.147
				Panel B: 8002 games					
mm	0.605	0.289	0.150	0.000	0.255	0.000	0.463	0.306	0.616
rp	0.501	0.440	0.247	0.000	0.115	0.000	0.321	0.378	0.761
p-value	0.000	0.000	0.000	-	0.000	-	0.000	0.086	0.080
Mixed	0.512	0.396	0.211	0.057	0.159	0.000	0.352	0.248	0.507
Pure	0.568	0.371	0.210	0.121	0.176	0.000	0.397	0.453	0.906
p-value	0.014	0.380	0.953	-	0.460	-	0.069	0.000	0.000
Continuous	0.550	0.429	0.263	0.000	0.156	0.000	0.375	0.335	0.674
Discrete	0.531	0.337	0.159	0.259	0.179	0.000	0.374	0.367	0.738
p-value	0.429	0.001	0.000	-	0.314	-	0.988	0.400	0.393
				Panel C: IDDS games					
mm	0.083	0.804	0.219	0.000	0.521	0.000	0.219	0.055	0.075
rp	0.106	0.752	0.283	0.000	0.479	0.000	0.283	0.057	0.106
p-value	0.409	0.178	0.119	-	0.238	-	0.119	0.974	0.625
Mixed	0.078	0.782	0.241	0.000	0.517	0.000	0.241	0.050	0.126
Pure	0.116	0.761	0.277	0.000	0.472	0.000	0.277	0.071	0.062
p-value	0.187	0.577	0.351	-	0.167	-	0.351	0.775	0.375
Continuous	0.122	0.805	0.244	0.000	0.496	0.000	0.244	0.064	0.112
Discrete	0.073	0.738	0.274	0.000	0.494	0.000	0.274	0.046	0.077
p-value	0.084	0.072	0.429	-	0.950	-	0.429	0.675	0.634

Table 6: Mean of the mean observations of pairs with mean Distance to predictions. Harmonic and geometric distances are calculated by IQR of both players. p-value in column 5 and 7 shows p-value for t test of by-period mean data for given treatments between distance to predictions.

Median of by-period Mean Summary Table

Treatments	row median	column median	To NE	p-value	To Center	p-value	To MM	Harmonic Disp	Geometric Disp
				Panel A: 3117 games					
mm	0.376	0.664	0.134	0.007	0.216	0.000	0.391	0.255	0.516
rp	0.516	0.529	0.302	0.000	0.107	0.000	0.314	0.261	0.540
p-value	0.000	0.000	0.000	-	0.000	-	0.000	0.579	0.530
Mixed	0.465	0.544	0.241	0.002	0.114	0.000	0.346	0.219	0.452
Pure	0.500	0.582	0.240	0.009	0.176	0.000	0.317	0.500	1.000
p-value	0.692	0.702	0.524	-	0.007	-	0.653	0.000	0.000
Continuous	0.529	0.544	0.341	0.001	0.198	0.000	0.352	0.219	0.456
Discrete	0.450	0.604	0.212	0.023	0.128	0.000	0.326	0.276	0.558
p-value	0.017	0.092	0.011	-	0.045	-	0.909	0.055	0.066
				Panel B: 8002 games					
mm	0.584	0.284	0.107	0.000	0.243	0.000	0.443	0.253	0.511
rp	0.500	0.432	0.238	0.000	0.110	0.000	0.320	0.440	0.881
p-value	0.000	0.000	0.000	-	0.000	-	0.000	0.111	0.101
Mixed	0.510	0.411	0.218	0.102	0.162	0.000	0.352	0.243	0.488
Pure	0.542	0.359	0.198	0.207	0.167	0.000	0.371	0.500	1.000
p-value	0.041	0.702	0.732	-	0.577	-	0.161	0.000	0.000
Continuous	0.521	0.431	0.265	0.002	0.138	0.000	0.352	0.276	0.573
Discrete	0.527	0.333	0.154	0.449	0.173	0.000	0.371	0.376	0.753
p-value	0.914	0.003	0.000	-	0.219	-	0.541	0.386	0.356
				Panel C: IDDS games					
mm	0.055	0.807	0.227	0.001	0.536	0.001	0.227	0.013	0.000
rp	0.084	0.737	0.278	0.000	0.487	0.000	0.278	0.002	0.000
p-value	0.360	0.155	0.116	-	0.224	-	0.116	0.423	0.983
Mixed	0.075	0.747	0.262	0.000	0.516	0.000	0.262	0.029	0.096
Pure	0.078	0.775	0.244	0.004	0.490	0.004	0.244	0.000	0.000
p-value	0.509	0.611	0.402	-	0.270	-	0.402	0.055	0.001
Continuous	0.102	0.799	0.230	0.002	0.500	0.002	0.230	0.007	0.000
Discrete	0.061	0.738	0.272	0.000	0.490	0.000	0.272	0.017	0.000
p-value	0.147	0.147	0.515	-	0.669	-	0.515	0.821	0.916

Table 7: Median of the mean observations of pairs with median Distance to predictions. Harmonic and geometric distances are calculated by IQR of both players. p-value in column 5 and 7 shows p-value for Wilcoxon signed-rank test of by-period mean data for given treatments between distance to predictions.

Dynamics details

Type	(1) CW	(2) Diagonal	(3) Stay	(4) CCW	(5) CD
continuous	-2.94***	-4.06***		-2.61***	-2.84***
	(0.319)	(0.381)		(0.322)	(0.385)
pure	-2.16***	-2.21***		-2.40***	-22.28***
	(0.307)	(0.348)		(0.329)	(0.370)
8002	-0.54	-0.60		-0.61	-0.20
	(0.462)	(0.496)		(0.468)	(0.534)
continuous_pure	1.49***	1.24**		0.09	2.27***
	(0.384)	(0.523)		(0.425)	(0.459)
continuous_8002	0.89*	0.93		0.84	0.76
	(0.506)	(0.584)		(0.511)	(0.609)
pure_8002	0.66	0.60		0.50	0.32
	(0.497)	(0.552)		(0.531)	(0.593)
continuous_pure_8002	-0.39	-0.54		-0.15	-0.57
	(0.587)	(0.753)		(0.646)	(0.708)
second_half	-0.16**	-0.06		-0.17**	-0.16
	(0.066)	(0.083)		(0.080)	(0.103)
block_2	-0.58***	-0.57***		-0.62***	-0.75***
	(0.150)	(0.218)		(0.156)	(0.249)
Constant	3.04***	2.37***		1.99***	0.76**
	(0.301)	(0.347)		(0.307)	(0.359)
Observations	37,920	37,920	37,920	37,920	37,920

Table 8: Multinomial logistic regressions of move type fractions on dummy variables. Dependent variables are dummy variables of classified types of dynamics given the observations. "Stay" type is used as the baseline. Independent variables are treatment dummies. Significance level: *** 0.01 ** 0.05 * 0.1.

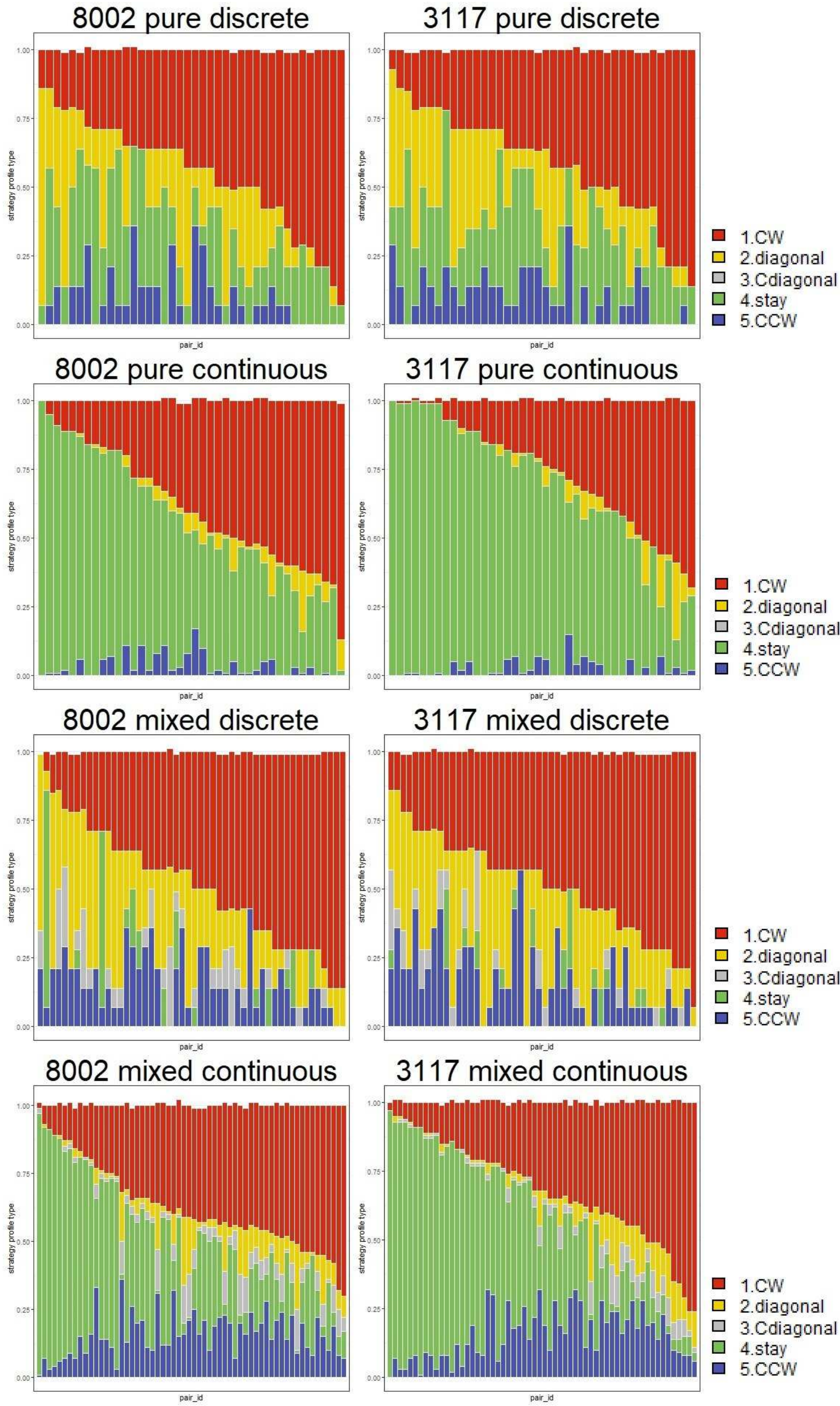

Figure 10: Types of cyclical behavior of each pair. From top to bottom: PD, PC, MD, MC. From left to right: 8002 games, 3117 games.

	(1)	(2)	(3)	(4)
Type	CW	CCW	diagonal	stay
continuous	-0.16***	-0.01	-0.21***	0.40***
	(0.036)	(0.018)	(0.024)	(0.041)
pure	-0.07*	-0.06***	-0.05*	0.22***
	(0.040)	(0.020)	(0.027)	(0.046)
8002	-0.01	-0.01	-0.02	0.02
	(0.038)	(0.019)	(0.025)	(0.043)
continuous_pure	0.01	-0.07***	0.03	0.02
	(0.056)	(0.027)	(0.037)	(0.063)
continuous_8002	0.06	0.02	0.03	-0.10*
	(0.051)	(0.025)	(0.034)	(0.058)
pure_8002	0.04	-0.00	0.01	-0.03
	(0.057)	(0.028)	(0.038)	(0.065)
continuous_pure_8002	0.03	0.01	-0.01	-0.03
	(0.079)	(0.039)	(0.052)	(0.090)
Constant	0.49***	0.17***	0.26***	0.04
	(0.027)	(0.013)	(0.018)	(0.030)
Observations	380	380	380	380
R-squared	0.125	0.234	0.359	0.475

Table 9: Regression of move type fractions at pair level. Dependent variables and numbers between 0 and 1 and show fraction of time pairs play each type of classified cycles. Independent variables are treatment dummies. Significance level: *** 0.01 ** 0.05 * 0.1.

Directional learning details

	BR learning			
	(1)	(2)	(3)	(4)
	Row	Row	Col	Col
	Dependent: $\triangle s_{it} = s_{i,t+1} - s_{it}$			
β_1	0.05***	0.45***	0.05***	0.40***
	(0.007)	(0.042)	(0.006)	(0.036)
pure	0.36***	0.15***	0.39***	0.18***
	(0.043)	(0.053)	(0.040)	(0.050)
mm	-0.01*	-0.21***	-0.02***	-0.15***
	(0.009)	(0.060)	(0.008)	(0.055)
8002	-0.01	-0.01	-0.01	-0.04
	(0.009)	(0.061)	(0.008)	(0.046)
IDDS	0.03	0.15	-0.04***	-0.03
	(0.027)	(0.119)	(0.008)	(0.100)
pure_mm	-0.14**	0.35***	-0.27***	0.11
	(0.056)	(0.097)	(0.050)	(0.080)
pure_8002	0.07	0.09	0.06	0.05
	(0.059)	(0.077)	(0.053)	(0.075)
pure_IDDS	-0.23***	0.27	-0.32***	0.11
	(0.067)	(0.166)	(0.046)	(0.136)
mm_8002	0.01	-0.06	0.00	0.08
	(0.011)	(0.089)	(0.011)	(0.075)
mm_IDDS	0.07	0.08	0.03***	0.10
	(0.043)	(0.160)	(0.010)	(0.122)
pure_mm_8002	-0.16**	-0.09	-0.03	-0.08
	(0.072)	(0.127)	(0.067)	(0.123)
pure_mm_IDDS	0.39***	-0.15	0.35***	0.07
	(0.148)	(0.209)	(0.066)	(0.187)
Observations	79,145	4,995	79,145	4,995
R-squared	0.263	0.371	0.262	0.274
Number of Pairs	415	345	415	345

Table 10: BR learning regression table. Column (1)(3) use continuous time data and column (2)(4) use discrete time data. Significance level: *** 0.01 ** 0.05 * 0.1.

Pure directional learning				
	(1)	(2)	(3)	(4)
	Row	Row	Col	Col
Dependent: $\triangle s_{it} = s_{i,t+1} - s_{it}$				
β_1	0.03***	0.20***	0.03***	0.16***
	(0.004)	(0.025)	(0.004)	(0.026)
pure	0.39***	0.40***	0.42***	0.43***
	(0.042)	(0.041)	(0.040)	(0.044)
mm	-0.01	-0.10***	-0.01***	-0.05
	(0.005)	(0.034)	(0.005)	(0.038)
8002	-0.01	0.03	-0.00	-0.00
	(0.005)	(0.043)	(0.005)	(0.033)
IDDS	-0.00	-0.01	-0.03***	-0.13*
	(0.007)	(0.059)	(0.005)	(0.077)
pure_mm	-0.15***	0.23***	-0.28***	0.00
	(0.055)	(0.083)	(0.050)	(0.069)
pure_8002	0.07	0.05	0.06	0.01
	(0.058)	(0.063)	(0.052)	(0.068)
pure_IDDS	-0.20***	0.43***	-0.33***	0.21*
	(0.062)	(0.129)	(0.045)	(0.120)
mm_8002	0.00	-0.09*	0.00	0.02
	(0.006)	(0.053)	(0.006)	(0.047)
mm_IDDS	-0.01	0.03	0.02***	0.14
	(0.008)	(0.071)	(0.006)	(0.099)
pure_mm_8002	-0.16**	-0.07	-0.03	-0.03
	(0.071)	(0.105)	(0.067)	(0.109)
pure_mm_IDDS	0.47***	-0.10	0.36***	0.04
	(0.142)	(0.152)	(0.065)	(0.173)
Observations	79,145	4,995	79,145	4,995
R-squared	0.263	0.339	0.262	0.240
Number of Pairs	415	345	415	345

Table 11: Pure directional learning regression table. Column (1)(3) use continuous time data and column (2)(4) use discrete time data. Significance level: *** 0.01 ** 0.05 * 0.1.

Directional learning with 5 lagged regret terms

	(1)	(2)	(3)	(4)
	Row	Row	Col	Col
	Dependent: $\triangle s_{it} = s_{i,t+1} - s_{it}$			
β_1	1.16***	0.62***	0.78***	0.50***
	(0.046)	(0.032)	(0.048)	(0.029)
β_1 L1	0.08*	0.01	-0.04	-0.02
	(0.042)	(0.022)	(0.036)	(0.020)
β_1 L2	0.08*	-0.18***	0.07*	-0.12***
	(0.046)	(0.026)	(0.043)	(0.017)
β_1 L3	0.17***	-0.08***	0.10**	-0.01
	(0.046)	(0.027)	(0.042)	(0.020)
β_1 L4	0.13***	0.01	-0.01	-0.01
	(0.043)	(0.016)	(0.041)	(0.015)
β_1 L5	0.24***	-0.06***	0.09**	-0.04***
	(0.044)	(0.019)	(0.036)	(0.013)
Observations	3,270	77,070	3,270	77,070
R-squared	0.284	0.220	0.187	0.231
Number of Pairs	345	415	345	415

Table 12: Directional learning regression table with lagged terms. Column (1)(2) are for row players learning and column (3)(4) are for column players learning. Column (1)(3) use continuous time data and column (2)(4) use discrete time data. Significance level: *** 0.01 ** 0.05 * 0.1.

Directional learning with 1 lagged regret term				
	(1)	(2)	(3)	(4)
	Row	Row	Col	Col
Dependent: $\triangle s_{it} = s_{i,t+1} - s_{it}$				
β_1	1.12***	0.61***	0.74***	0.51***
	(0.038)	(0.032)	(0.037)	(0.030)
β_1 L1	0.01	-0.12***	-0.08***	-0.11***
	(0.033)	(0.023)	(0.025)	(0.024)
Observations	4,650	78,730	4,650	78,730
R-squared	0.308	0.190	0.199	0.218
Number of Pairs	345	415	345	415

Table 13: Directional learning regression table with lagged terms. Column (1)(2) are for row players learning and column (3)(4) are for column players learning. Column (1)(3) use continuous time data and column (2)(4) use discrete time data. Significance level: *** 0.01 ** 0.05 * 0.1.

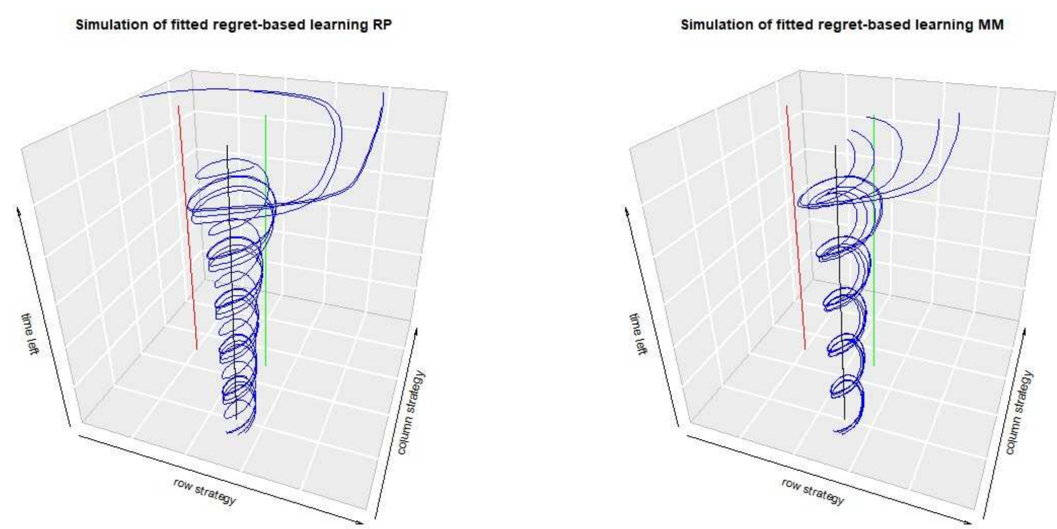

Figure 11: Simulation result in 8002 games mixed strategy treatments. Simulation under random pairwise matching is on the left and simulation under mean matching is on the right.

Directional learning with discrete regret				
	(1)	(2)	(3)	(4)
	Row	Row	Col	Col
Dependent: $\triangle s_{it} = s_{i,t+1} - s_{it}$				
β_1	0.02***	0.19***	0.03***	0.14***
	(0.004)	(0.020)	(0.004)	(0.022)
pure	0.15***	0.10***	0.19***	0.15***
	(0.019)	(0.027)	(0.020)	(0.028)
mm	-0.00	-0.09***	-0.01***	-0.02
	(0.004)	(0.030)	(0.004)	(0.034)
8002	-0.01	0.00	-0.01**	-0.03
	(0.004)	(0.031)	(0.005)	(0.025)
IDDS	0.00	0.02	-0.03***	0.01
	(0.007)	(0.052)	(0.004)	(0.062)
pure_mm	0.05	0.26***	-0.11***	0.08
	(0.035)	(0.063)	(0.029)	(0.051)
pure_8002	0.05**	0.03	-0.04*	-0.06*
	(0.027)	(0.039)	(0.023)	(0.034)
pure_IDDS	-0.07**	0.20***	-0.13***	0.11
	(0.029)	(0.078)	(0.025)	(0.076)
mm_8002	0.00	-0.06	0.01	0.04
	(0.006)	(0.043)	(0.006)	(0.040)
mm_IDDS	-0.02**	0.01	0.02***	-0.01
	(0.008)	(0.064)	(0.005)	(0.087)
pure_mm_8002	-0.15***	-0.00	0.02	-0.08
	(0.043)	(0.075)	(0.034)	(0.063)
pure_mm_IDDS	0.12*	-0.15	0.14***	-0.03
	(0.074)	(0.102)	(0.040)	(0.114)
Observations	79,145	4,995	79,145	4,995
R-squared	0.231	0.337	0.259	0.247
Number of session_round_pair_id	415	345	415	345

Table 14: Pure directional learning regression table. Column (1)(2) are for row players learning and column (3)(4) are for column players learning. Column (1)(3) use continuous time data and column (2)(4) use discrete time data. Significance level: *** 0.01 ** 0.05 * 0.1.

8 Bibliography

References

[1] Ken Binmore, Joe Swierzbinski, and Chris Proulx. Does minimax work? an experimental study. *Economic Journal*, pages 445–464, 2001.

[2] George W Brown. Iterative solution of games by fictitious play. *Activity analysis of production and allocation*, 13(1):374–376, 1951.

[3] James N Brown and Robert W Rosenthal. Testing the minimax hypothesis: a re-examination of o'neill's game experiment. *Econometrica (1986-1998)*, 58(5):1065, 1990.

[4] Colin Camerer and Teck Hua Ho. Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4):827–874, 1999.

[5] Daniel L. Chen, Martin Schonger, and Chris Wickens. otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9(1):88–97, 2016.

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

[7] P-A Chiappori, Steven Levitt, and Timothy Groseclose. Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*, 92(4):1138–1151, 2002.

[8] Vincent P Crawford. Learning behavior and mixed-strategy nash equilibria. *Journal of Economic Behavior & Organization*, 6(1):69–78, 1985.

[9] Arthur Conan Doyle. *The final problem.* Johan Brown, 2018 (originally 1893).

[10] Ido Erev and Alvin E Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, pages 848–881, 1998.

[11] Daniel Friedman. Evolutionary games in economics. *Econometrica*, 59(3):637–666, 1991.

[12] Daniel Friedman. Equilibrium in evolutionary games: Some experimental results. *Economic Journal*, 106(434):1–25, 1996.

[13] Daniel Friedman and KC Fung. International trade and the internal organization of firms: An evolutionary approach. *Journal of International Economics*, 41(1-2):113–137, 1996.

[14] Jacob K Goeree, Charles A Holt, and Thomas R Palfrey. Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, 45(1):97–113, 2003.

[15] Ben Greiner. Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125, 2015.

[16] Josef Hofbauer and Ed Hopkins. Learning in perturbed asymmetric games. *Games and Economic Behavior*, 52(1):133–152, 2005.

[17] Natalia Lazzati and Amilcar A Menichini. Hot spot policing: A study of place-based strategies for crime prevention. *Southern Economic Journal*, 82(3):893–913, 2016.

[18] Steven D Levitt, John A List, and David H Reiley. What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments. *Econometrica*, 78(4):1413–1434, 2010.

[19] Alfred J Lotka. Elements of physical biology. *Science Progress in the Twentieth Century (1919-1933)*, 21(82):341–343, 1926.

[20] Dilip Mookherjee and Barry Sopher. Learning behavior in an experimental matching pennies game. *Games and Economic Behavior*, 7(1):62–91, 1994.

[21] John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[22] J v Neumann. Zur theorie der gesellschaftsspiele. *Mathematische annalen*, 100(1):295–320, 1928.

[23] Jack Ochs. Games with unique, mixed strategy equilibria: An experimental study. *Games and Economic Behavior*, 10(1):202–217, 1995.

[24] Barry O'Neill. Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the national academy of sciences*, 84(7):2106–2109, 1987.

[25] Ignacio Palacios-Huerta. Professionals play minimax. *The Review of Economic Studies*, 70(2):395–415, 2003.

[26] Anatol Rapoport and Carol Orwant. Experimental games: A review. *Behavioral Science*, 7(1):1–37, 1962.

[27] Julia Robinson. An iterative method of solving a game. *Annals of mathematics*, pages 296–301, 1951.

[28] Julian Romero and Yaroslav Rosokha. Constructing strategies in the indefinitely repeated prisoners dilemma game. *European Economic Review*, 104:185–219, 2018.

[29] Reinhard Selten. Anticipatory learning in two-person games. In *Game equilibrium models I*, pages 98–154. Springer, 1991.

[30] Lloyd Shapley. Some topics in two-person games. *Advances in game theory*, 52:1–29, 1964.

[31] Eilon Solan, Michael Maschler, and Shmuel Zamir. *Game Theory*. Cambridge U Press, 2013.

[32] Dale O Stahl II. On the instability of mixed-strategy nash equilibria. *Journal of Economic Behavior & Organization*, 9(1):59–69, 1988.

[33] Daniel Stephenson. Coordination and evolutionary dynamics: When are evolutionary models reliable? *Games and Economic Behavior*, 113:381–395, 2019.

[34] Fang-Fang Tang. Anticipatory learning in two-person games: some experimental results. *Journal of economic behavior & organization*, 44(2):221–232, 2001.

[35] Mark Walker and John Wooders. Minimax play at wimbledon. *American Economic Review*, 91(5):1521–1538, 2001.

[36] Jörgen W Weibull. *Evolutionary game theory*. MIT press, 1997.

[37] John Wooders. Does experience teach? professionals and minimax play in the lab. *Econometrica*, 78(3):1143–1154, 2010.