When Are Mixed Equilibria Relevant?*

by Daniel Friedman† and Shuchen Zhao‡

August 2, 2021

Abstract

Mixed strategy equilibria — Nash (NE) and maximin (MM) — are cornerstones of game theory, but their empirical relevance has always been questionable. We study in the laboratory two games, each with a unique NE and a unique (and distinct) MM in completely mixed strategies. Treatment variables include the matching protocol (pairwise random vs population mean matching), whether time is discrete or continuous, and whether players can specify explicit mixtures or only pure strategy realizations. NE mixes predict observed behavior relatively well in population mean matching treatments, and predict better than MM in all treatments. However, in most random pairwise treatments, uniform mixes predict better than NE. Regret-based and sign preserving dynamics capture regularities across all treatments.

JEL Classification: C72, C73, C92.

Keywords: Nash equilibrium, Maximin, Mixed strategy, Sign preserving dynamics, Laboratory experiment.

*We are grateful for support from National Science Foundation grants SES-1357867 and SES-0925039; for programming assistance from Morgan Grant and James Pettit; to seminar participants at WZB (Berlin) and Trento workshops in May 2019, and at ESA conferences in Vancouver and Los Angeles in July and October 2019; and for helpful comments from Ciril Bosch, Tim Cason, Vince Crawford, Ed Hopkins, Steffen Huck, Johannes Leutgeb, Luigi Mittone, Heinrich Nax, Joerg Oechssler, Julian Romero, Nirvikar Singh, Daniel Stephenson, Donald Wittman and John Wooders, and especially an anonymous referee of this journal. Earlier unpublished work with Shmuel Zamir and with Ryan Oprea and Nat Wilcox helped us formulate some of the ideas and treatments presented below. **Declarations of interest: none.**

†Economics Departments, University of Essex and University of California Santa Cruz; dan@ucsc.edu;

‡Economics Department, University of California Santa Cruz; szhao19@ucsc.edu. (Corresponding author)

1 Introduction

Asymmetric matching pennies games capture the essence of strategic situations (e.g., in hunting, warfare and sports) where success comes from outguessing opponents. To game theorists, such games are notable because they have no pure strategy equilibrium. In a typical asymmetric matching pennies game, there is one specific profile of mixed strategies that constitutes a Nash equilibrium (NE), and a different mixed strategy profile that constitutes a maximin (MM).

Does either equilibrium notion have predictive value? The literature reviewed below will note that different theoretical models offer sharply different answers to that question, and that the empirical evidence is also quite mixed.

The present paper pursues two more specific research questions. First, under what conditions (if any) does either equilibrium do a good job of predicting behavior in asymmetric matching pennies games? Second, are there alternative point predictions or dynamic predictions that can do better?

Those questions are important to applied social scientists and biologists as well as to theorists. Cyber-attacks and -defenses are naturally modeled as asymmetric matching pennies games, as are military conflicts and penalty kicks in soccer games. Standard biological models of predator-prey interactions can be regarded as dynamic versions of those games. Answers to our research questions should help guide work in such applications.

We address the questions by means of a laboratory experiment that deploys a variety of treatment combinations intended to reveal favorable circumstances for each competing prediction. The treatment variables include two different simple matching pennies games as well as a control game that has only a pure Nash equilibrium; two matching protocols that contrast population games (as in biology and some economic applications such as national cybersecurity policy) to standard pairwise matching; allowing players to specify explicit mixtures versus only pure strategy realizations; and simultaneous moves in discrete time versus asynchronous moves in continuous time.

The results are instructive. We find some circumstances under which Nash equilibrium predicts human subject behavior fairly well, but find no circumstance under which maximin

does better. There is a surprisingly wide range of circumstances under which uniform random mixing predicts better than Nash equilibrium. For the asymmetric matching pennies games we consider, logit quantal response equilibrium typically offers slight or no improvement over its edge cases, Nash equilibrium and uniform mixing. The upshot is that mixed equilibria are empirically relevant over a narrower range of circumstances than one might have supposed. On the other hand, we find that simple dynamic models, with qualitative directional adjustments or quantitative regret-based adjustments in players' strategies, capture important regularities in the short run as well as in the long run.

After reviewing some previous literature in Section 2 and some established theory in Section 3, we present our experimental design in Section 4. We implement our 2x2 bimatrix games using a new graphical screen display for players. Section 4 concludes with lists of testable hypotheses about the competing point predictions, about treatment effects on mean choices and on dispersion, and about adaptive dynamics. Section 5 presents the results, and a concluding discussion in Section 6 summarizes our findings and suggests implications for game theorists and applied researchers. Online appendices include supplementary analysis, and instructions to subjects.

2 Previous literature

> "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)

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 to Dover,[1] but Holmes recognizes that higher levels are possible. Since level-$(k+1)$ beats level-k for every positive k in asymmetric matching pennies, these games suffer from infinite regress.

[1] Aficionados of A.C. Doyle's story and of level k reasoning might regard random attacks by Moriarty as level 0, and regard Holmes taking the train to Dover as level 1.

That is not a problem in short stories such as Doyle (1893) or Poe (1844), whose hero is always exactly one level above the villain, but for proto-game theorists, infinite regress is a Gordian knot that blocked progress for centuries. Von Neumann (1928) finally cut that knot with the idea of mixed strategies, and introduced maximin (or minimax for loss functions) as an equilibrium concept.

Early game theory emphasized two-player zero-sum bimatrix games where Nash equilibrium (NE) and maximin (MM) mixtures coincide, but it recognized that NE and MM mixtures differ in asymmetric matching pennies games (e.g., Solan et al., 2013). Nash (1951) and his contemporaries also recognized that payoffs are flat in the neighborhood of completely mixed NE and worried about whether actual play might approach it. Fictitious play dynamics (Julia Robinson, 1951; George Brown, 1951) guarantee that time-average play converges to equilibrium in the zero-sum case, but Shapley (1964) devised a non-zero-sum game with a unique NE in mixed strategies for which fictitious play dynamics converge to a limit cycle and not to the NE.

Subsequent theory does not yield clear predictions on dynamic stability. Stahl (1988), Crawford (1985) and others showed that convergence to equilibrium in asymmetric matching pennies games generally fails for their favored dynamics. On the other hand, for such games Fudenberg and Kreps (1993) showed that stochastic fictitious play converges, and the noise inherent in the Erev and Roth (1998) reinforcement learning model produces a similar result. Hopkins (2002) showed that for both of those stochastic models, convergence is to QRE and thus to NE in the limit as noise amplitude vanishes. Convergence fails for replicator dynamics (Taylor and Jonker, 1978): the matching pennies NE is neutrally stable. Hofbauer and Hopkins (2005) found that for a broad class of dynamics, stable mixed equilibria are vanishingly rare in games with more than 2 pure strategies.

These heterogeneous theoretical results on point predictions and dynamics underline the need for empirical work, but so far the empirical results are also quite heterogeneous. Rapoport and Orwant (1962) surveyed early laboratory experiments, and found that average play typically was closer to a uniform mix (e.g., .50-.50) than to a NE or MM mix. O'Neill (1987) found that overall time-average play is surprisingly close to the NE mix in a particular zero-sum 4x4 game, but (James) Brown and Rosenthal (1990) noted that this does not imply that individual players employ NE strategies, and they indeed found substantial

departures from the specified iid mixes. Subsequent empirical contributions such as Walker and Wooders (2001), Chiappori et al (2002) and 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 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, implement equilibrium mixtures (e.g., Friedman, 1996; Binmore et al, 2001).

To our knowledge, only one previous empirical paper compared maximin mixtures to NE mixtures. Ochs (1995) considered several treatments (including one that uses a set of 9 explicit mixtures) in asymmetric matching pennies games, but finds that neither Nash equilibrium nor maximin tracks the observed changes in average play when game parameters change. Goeree, Holt and Palfrey (2003) found 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 empirical literature on adaptive dynamics in matching pennies games. Mookherjee and Sopher (1994) found that belief learning (responsive to payoffs that would have been earned by strategies not employed) beats reinforcement learning. On the other hand, Erev and Roth (1998) found that their three parameter reinforcement learning model compares favorably to other learning models in predicting behavior in 12 laboratory studies of games with unique interior mixed NE. Camerer and Ho (1999)'s EWA model includes an extra parameter to hybridize belief learning (a la Friedman and Cheung, 1996) with reinforcement learning; the authors showed that it is able to fit a variety of games, including some matching pennies games. Tang (1999) presented two 3x3 bimatrix games with different predicted dynamic stability properties; the data favor the Selten (1991) anticipatory dynamics model over the Crawford (1985) model. Stephenson (2019) reported an experimental test of evolutionary models in coordinated attacker-defender games that include own-population effects (Friedman, 1991) not considered in our asymmetric matching pennies games. His results are consistent with a class of adaptive dynamic models. Taken together, these papers suggest that there may be a useful empirical role for dynamic models of some sort, but it is not yet clear which sorts are best nor when they will predict convergence or non-convergence to a point prediction. In sum, despite important prior work by leading game theorists and experimentalists, our motivating research questions remain open.

3 Theoretical Considerations

Name	AMPa	AMPb	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)
MM	(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.

Point predictions. Table 1 shows the specific bimatrix games that we will study. The first two are asymmetric matching pennies (AMP) games. The Online Appendix includes the straightforward computation of the unique Nash equilibrium (NE) and maximin (MM) strategies; these games were chosen in part to create separation between those mixtures. The third game, IDDS, is dominance solvable. Its unique NE coincides with its MM and is in pure strategies, and therefore provides a useful contrast to the AMP games.

Figure 1 graphically displays the mixed extension of the AMPa 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), and Center (uniform mixture).

At least since Nash (1951), game theorists have recognized two different 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

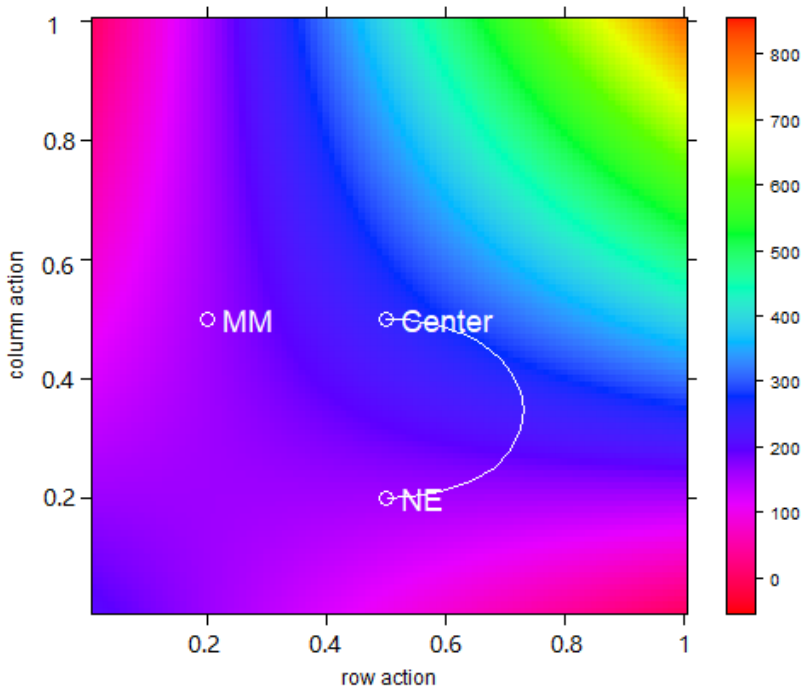

Figure 1: Heatmap for AMPa row player. The color at coordinates $(x, y) = (a, b)$ indicates, via scaled thermometer at right, the row player's expected payoff at mixed strategy profile (a, b). NE, MM, Center respectively mark the coordinates of Nash equilibrium, maximin and the Center profiles. The arc connecting Center $(x, y) = (0.5, 0.5)$ to NE includes all symmetric logit quantal response equilibrium (QRE) profiles.

anonymously with members of a large column player population, and an equilibrium distribution of action profiles remains unchanged as individual players adapt. Binmore et al. (2001), among others, claimed 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 by pairs of individuals or by populations. That claim motivates the mean-matching vs random-pairwise protocols presented in Section 4.

Of course, the adaptation process, and hence the stability of mixed equilibrium, may also depend on whether individual players can choose mixture weights explicitly (as for example, a political strategist deciding the fraction of positive vs negative ads to run) or can choose only the frequency of pure actions. The adaptation process also depends on whether players' choices are made simultaneously in discrete time (as in most laboratory experiments) or

asynchronously in continuous time (as in Cyber-attacks and -defenses, for example). Such considerations motivate the action set and time treatment variables presented in Section 4.

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; Friedman and Fung, 1996; Weibull, 1997). The idea is simply that players (individually or collectively) should increase the weight on the pure strategy with the currently higher payoff.

To formalize such qualitative "directional" dynamics, let the time t mixed strategy profile be $(a(t), b(t))$ for a bimatrix game $M = (M^R; M^C)$. For example, in the AMPa game, $M^R = \begin{pmatrix} 800 & 0 \\ 0 & 200 \end{pmatrix}$. For smooth play in continuous time, $(\dot{a}(t), \dot{b}(t))$ denotes time rate of change. The payoff difference between pure strategies is $D_R(t) = (1, -1)M^R \cdot (b(t), 1 - b(t))$ for row player(s) and is $D_C(t) = (1, -1)M^C \cdot (a(t), 1 - a(t))$ for column player(s).

A 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. Such directional change is a fundamental property of learning and evolution, satisfied (at least approximately)[2] by all standard versions of adaptive dynamics. 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 four regions, each with its own implied direction of change.

Figure 2 illustrates for the AMPa 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 horizontal line $b = 0.2$. Similar calculations show that $D_C = 200 - 400a$ so the isocline $D_C = 0$ is the vertical 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 Southeast rectangle $a > 0.5, b < 0.2$, we have $D_R < 0, D_C < 0$, so sign preserving dynamics imply a trajectory with $\dot{a} < 0, \dot{b} < 0$.

[2]The reason for this qualification is that a lower payoff strategy's share when sufficiently small may be increased by a noise term (e.g., in stochastic best response dynamics), as noted e.g. by Stephenson (2019).

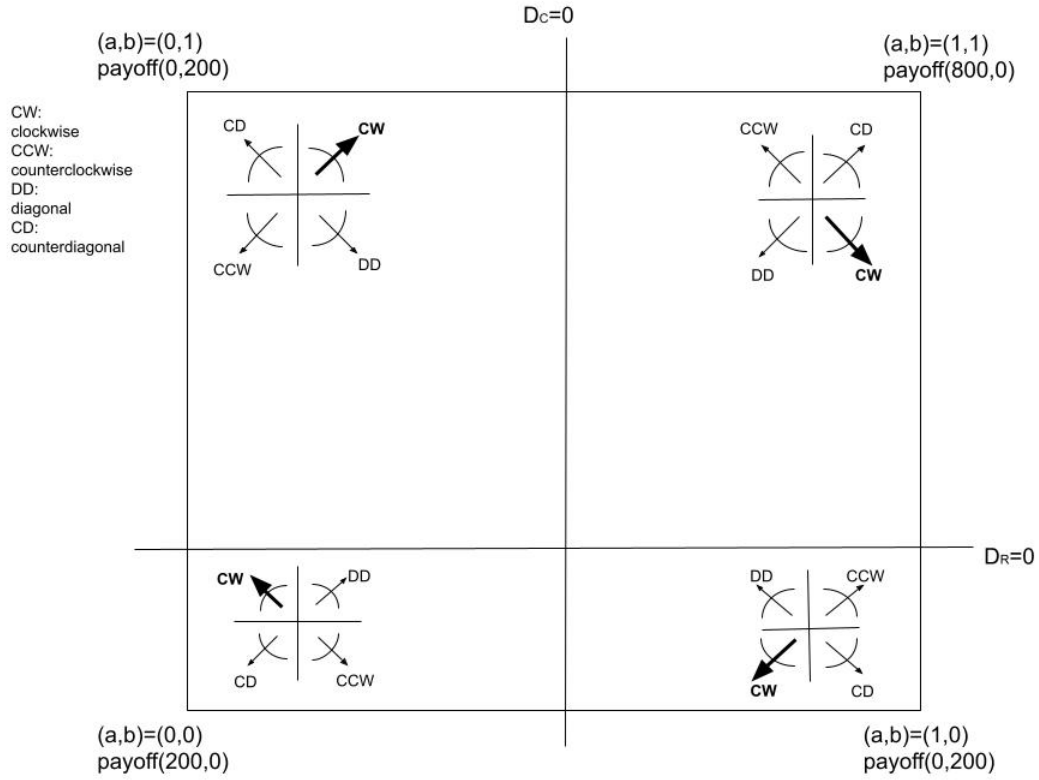

Figure 2: Classifying directional changes in AMPa games. The horizontal (column player's) axis is reversed to be consistent with the payoff matrices in Table 1. Arcs and arrows in each rectangle indicate the four possible directions defined in the text.

Since the trajectory can not exit the square, this implies that $(a(t), b(t))$ moves clockwise towards the Southwest rectangle $a < 0.5, b < 0.2$. Similarly, in that Southwest rectangle, sign preserving dynamics imply $\dot{a} < 0, \dot{b} > 0$, i.e., moving clockwise towards the Northwest rectangle. Indeed, straightforward calculations show that sign preserving dynamics imply clockwise moves from each rectangle to the next for both AMP games.

Of course, human subject behavior is noisy, so the prediction from this theory is only that clockwise (CW) will be the most common direction of change. Figure 2 depicts the other three possible directions: counterclockwise (CCW), diagonal (DD) towards the Nash equilibrium mix, and counterdiagonal (CD) towards the corner (pure strategy profile) contained in that rectangle. These directions are all defined by the signs of $\dot{a}$ and $\dot{b}$ or, in empirical practice, by the signs of first differences in successive observations of $a(t)$ and $b(t)$.

In do-it-yourself randomization treatments, pairs of individual subjects must satisfy $a(t), b(t) \in \{0, 1\}$, so sign preserving dynamics will jump clockwise from one corner of the square to the next. If jumps are always clockwise and the time between jumps is roughly constant, the average strategy profile will approximate the Center, $(0.5, 0.5)$. On the other hand, for protocols that either allow explicit mixing or that involve population matching, the trajectory typically lies in the interior of the state space. Clockwise trajectories may converge to NE via damped cycles, or may cycle endlessly with constant average amplitude, or may spiral away from the intersection (at NE) of the zero isoclines. These possibilities are all consistent with sign preserving dynamics.

Regret-based dynamics. To obtain sharper predictions, we now construct a quantitative model of directional dynamics based on a notion of regret. Let $s_{it} \in [0, 1]$ denote player (or player population) i's mixture at time t, and let $f_i(s_{it}, s_{-it})$ be the corresponding payoff. For example, for a row player i in the AMPa game, $f_i(s_{it}, s_{-it}) = f_R(a(t), b(t))$. Regret is defined as the normalized shortfall from maximal payoff,[3] $R_{it} = \frac{f_i(\hat{s}_{it}, s_{-it}) - f_i(s_{it}, s_{-it})}{\max_{0 \leq x,y \leq 1} f_i(x,y)} \geq 0$ 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,

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

The signum (sign) function is $\mathrm{sgn}(y) = +1$ if $y > 0$; $= 0$ if $y = 0$; and $= -1$ if $y < 0$.

As long as $\beta_1 > 0$, equation (2) implies sign preserving dynamics in which the adaptation speed is proportional to regret, that is, to the potential advantage. Cruder variants (whose estimation is reported in the Appendix) include best response learning

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

and pure directional learning

$$\triangle\, s_{it} = \beta_1 \mathrm{sgn}(\hat{s}_{it} - s_{it}) + \epsilon_{it}\ . \tag{4}$$

Sets of static predictions. In this paper we focus on static point predictions and on adaptive dynamics, but in passing we will also refer to two leading static predictions that

[3]The normalization assumes that maximal payoff is positive, as is the case in the bimatrix games used in the present paper. Otherwise the denominator could be replaced by $\max_{0 \leq x,y \leq 1} f_i(x, y) - \min_{0 \leq x,y \leq 1} f_i(x, y)$.

involve sets of points. The first is symmetric logit quantal response equilibrium (QRE), due to McKelvey and Palfrey (1995) and Chen et al. (1997). QRE is defined by a parametrized implicit equation reproduced in the Online Appendix; for each value of the parameter $\lambda > 0$ we obtain a unique point prediction. As the parameter λ ranges from 0 to ∞, we get a one-dimensional arc from the state space Center to the NE. Rank-dependent choice equilibrium (RDCE), due to Goeree et al (2019) is the set of quantal response equilibria for all admissible functional forms, not just logit. RDCE uses the $D_R = 0$ and $D_C = 0$ loci as in Figure 2, together with the $a = 0.5$ and $b = 0.5$ loci, to define regions where the best response in each population is the more prevalent strategy. Thus RDCE non-parametrically defines a two-dimensional subset of the state space which contains all quantal response equilibria; see Online Appendix for calculations. The RDCE prediction is that time-average play will converge asymptotically to some particular point in that subset.

4 Laboratory Implementation

4.1 Treatment variables

Our experiment has four treatment variables. The first is the payoff bimatrix: as noted earlier, we consider two asymmetric matching pennies games, denoted AMPa and AMPb, as well as a dominance solvable game denoted IDDS.

The second treatment variable is the action set. In condition P (pure action set), subjects use radio buttons to select one of two bimatrix rows. The display highlights the cell with payoffs for that choice, given the column chosen by the matched player.[4] In condition M (mixed action set), subjects use a vertical slider to adjust an explicit mixture of the two strategies, as illustrated in Figure 3. The heatmap display indicates the payoff resulting from the player's chosen mix and the matched players' mix.

The third treatment variable concerns time. In our standard discrete time (D) condition, subjects' choices are updated simultaneously at regular time intervals, here 6000 ms (6 seconds). In the continuous time (C) condition, subjects update choices asynchronously in

[4]To maintain consistency from subjects' perspective, every player's screen shows her own choice as between rows, even for subjects labelled in this paper as column players.

real time, with an imperceptible latency of around 50 ms, and data are recorded every 500 ms (twice every second). Previous literature (e.g., Oprea et al., 2011 and Friedman et al., 2015) argues that continuous time is more realistic in many applications in sports, e-commerce and elsewhere, and that it can facilitate cooperation and speed convergence. To our knowledge, there is no previous report of the impact of continuous time treatments on games such as asymmetric matching pennies where cooperation is infeasible.

Payoffs are flows accumulated over time in condition C, as illustrated in the lower right graph in Figure 3. In condition D, the blue area representing payoffs is of adjoining rectangles of width 6 seconds and of height given by the payoff at the chosen profile.

The remaining treatment variable is the matching protocol. As noted earlier, the dynamic adjustment process and the stability of an equilibrium profile may depend on whether the game is played by pairs of individuals or played at the population level, and each interpretation has real-world applications. In our experiment, 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.

4.2 Design

The experiment used oTRw software, a hybrid of oTree (Chen et al, 2016) and LEEPS lab's Redwood suite, illustrated for the most distinctive treatments in Figure 3. Inexperienced subjects were recruited from the LEEPS lab subject pool using a local implementation of ORSEE (Greiner, 2015). Each session took around 90 minutes, with a 20-minute instruction/practice stage, 60-minute game play stage and 10-minute payment/closing stage. Payments ranged from US \$14 to \$24, and averaged about \$17 per subject.

The data analyzed below come from the 8 sessions specified in Table 2. Each session is

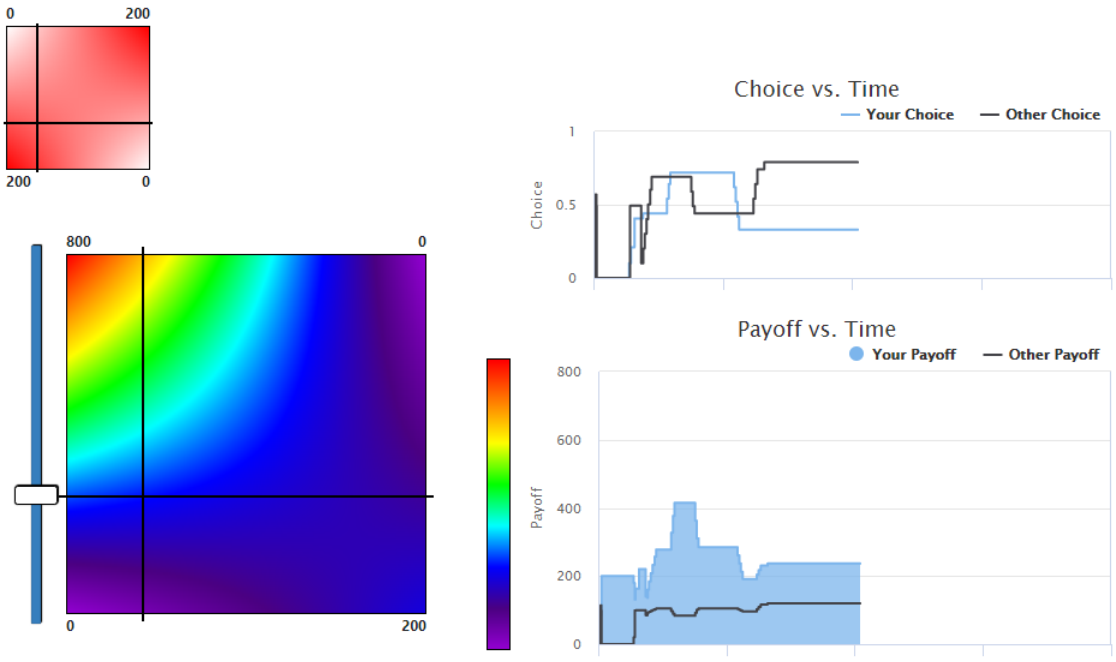

Figure 3: Main features of oTRw screen for MCrp AMPa game. The subject uses vertical slider at left to adjust her mixture (horizontal line); vertical line shows matched player's current mix. Heatmap color at intersection of these lines codes her current flow payoff; thermometer sets scale. Graph at lower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

equally divided into 5 blocks, each with a given payoff bimatrix; the sequence across blocks is AMPa, AMPb, IDDS, AMPa, and AMPb. The other three treatment variables — action set, matching and time — are constant within each session but vary across sessions according to the classic full factorial design.

In each block, subjects play several periods as detailed in Table 2: in rp sessions the periods last 90 seconds, with random rematching to start each new period, while they last 150 seconds in mm sessions. In D sessions, each period is divided into numerous 6-second subperiods and subjects update choices simultaneously, while in C sessions the subjects update freely in real time.

The 2x2x2 factorial design is very efficient for our purposes. For each of our three binary

Action Set	Time	Matching Protocol	Block Size	# Subjects
P	C	rp	6 × 90s periods	8
M	D	rp	6 × 90s periods	10
P	D	rp	6 × 90s periods	12
M	C	rp	6 × 90s periods	8
P	D	mm	4 × 150s periods	8
M	C	mm	4 × 150s periods	10
P	C	mm	4 × 150s periods	8
M	D	mm	4 × 150s periods	12
Game Sequence of 5 Blocks in Each Session				
Block 1	Block 2	Block 3	Block 4	Block 5
AMPa	AMPb	IDDS	AMPa	AMPb

Table 2: Experiment Design. P=pure, M=mixed action set; C=continuous, D=discrete time; rp=random pairwise, mm=mean matching protocol. The "Block Size" column reports the number × the length of periods in each block.

treatment variables, half the data (4 sessions, each with two 2 blocks of 4 or 6 periods) come from one level, e.g, C(ontinuous) time, and the other half of the data (the other 4 sessions) come from the other level, e.g., D(iscrete) time; likewise for P vs M action sets and for rp vs mm matching protocols. The design therefore enables broad-based tests (using the entire data set and thus a variety of environments) of the direct effect of each treatment variable. It should be acknowledged, however, that our design is underpowered for some other purposes, e.g., for estimating higher-order effect sizes such as how much the NE forecast deteriorates or improves when we move from M to P action sets in a D × rp environment. Due to the relatively small number of sessions, we apply the subcluster wild bootstrap method (MacKinnon and Webb, 2018; Roodman et al., 2019) when reporting statistical significance; this method is far more conservative than simply clustering errors at the session level.

4.3 Testable Hypotheses

Our design transforms the original research questions into a series of testable hypotheses. Hypotheses H1 - H3 concern the competing point predictions (NE, MM, Center), while H4

and H5 concern dynamic predictions. The first research question comes down to comparing average observed profiles to predicted profiles.

H1-NE. The time-average observed profile (a) will not differ significantly from Nash equilibrium, and (b) will be better approximated by NE than by alternative point predictions.
H1-MM. The time-average observed profile (a) will not differ significantly from maximin, and (b) will be better approximated by maximin than by alternative point predictions.
H1-Center. The time-average observed profile (a) will not differ significantly from Center, and (b) will be better approximated by Center than by alternative point predictions.

The first research question also asks about the impact of environmental conditions, which for us are the matching protocol, the action set treatment variable and the time treatment variable. H2 makes directional conjectures on conditions more conducive to convergence. The conjectured directions are based only on folk wisdom and fragmentary existing evidence; indeed, a major goal of our experiment is to help fill gaps in existing knowledge of how such environmental conditions affect the outcome of games with only mixed equilibria.

H2-NE. The time-average observed profile will be closer to Nash equilibrium (NE) (a) under mean matching (mm) than under random pairwise (rp) matching protocol; (b) with explicit mixes (M) than with only pure actions (P); and (c) in continuous time interaction (C) than in discrete time (D).
H2-MM and **H2-Center** replace NE by the alternative point predictions MM and Center.

Again for the sake of specificity in the absence of clear theory or empirical evidence, H3 makes directional conjectures regarding convergence to *any* particular point, predicted or otherwise. We measure (lack of) convergence empirically in terms of *dispersion*, operationalized as the geometric mean of row and column standard deviations. That is, for $d_R =$ standard deviation of Row mixes $a(t)$ in the observed sample and d_C similarly defined for Column mixes $b(t)$, dispersion is defined as $d_G = \sqrt{d_R d_C}$.

H3: There will be less dispersion around the observed time average profile (a) under mean matching (mm) than under random pairwise (rp) matching protocol; (b) with explicit mixes (M) than with only pure actions (P); and (c) in continuous time interaction (C) than in discrete time (D).

Hypothesis H4 comes directly from the discussion of sign-preserving dynamics in Section

3, while H5 is the hypothesis of regret-based adjustment together with our conjectures on conducive treatments.

H4: In all treatments in asymmetric matching pennies games, the most frequently observed direction of change will be clockwise (CW).

H5: Estimates of the rate coefficient β_1 in learning model (2) (a) will be significantly positive in all treatments. Estimates will be more positive (b) for mm than for rp matching, (c) for D than for C time interaction, and (d) for P than for M action sets.

5 Results

To gain perspective before reporting the hypothesis test results, we examine some exemplary raw data. Each panel in Figure 4 displays the time path of action profiles for one instance of each treatment in the AMPa bimatrix game. An instance is defined as a matched pair of players in one period for any rp treatment, or as a single period of play in any mm (population) treatment.

Panel a shows the pure action 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 action profile is recorded twice per second. The short vertical segments of the time path indicate episodes where both players stay 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 usually moves 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.

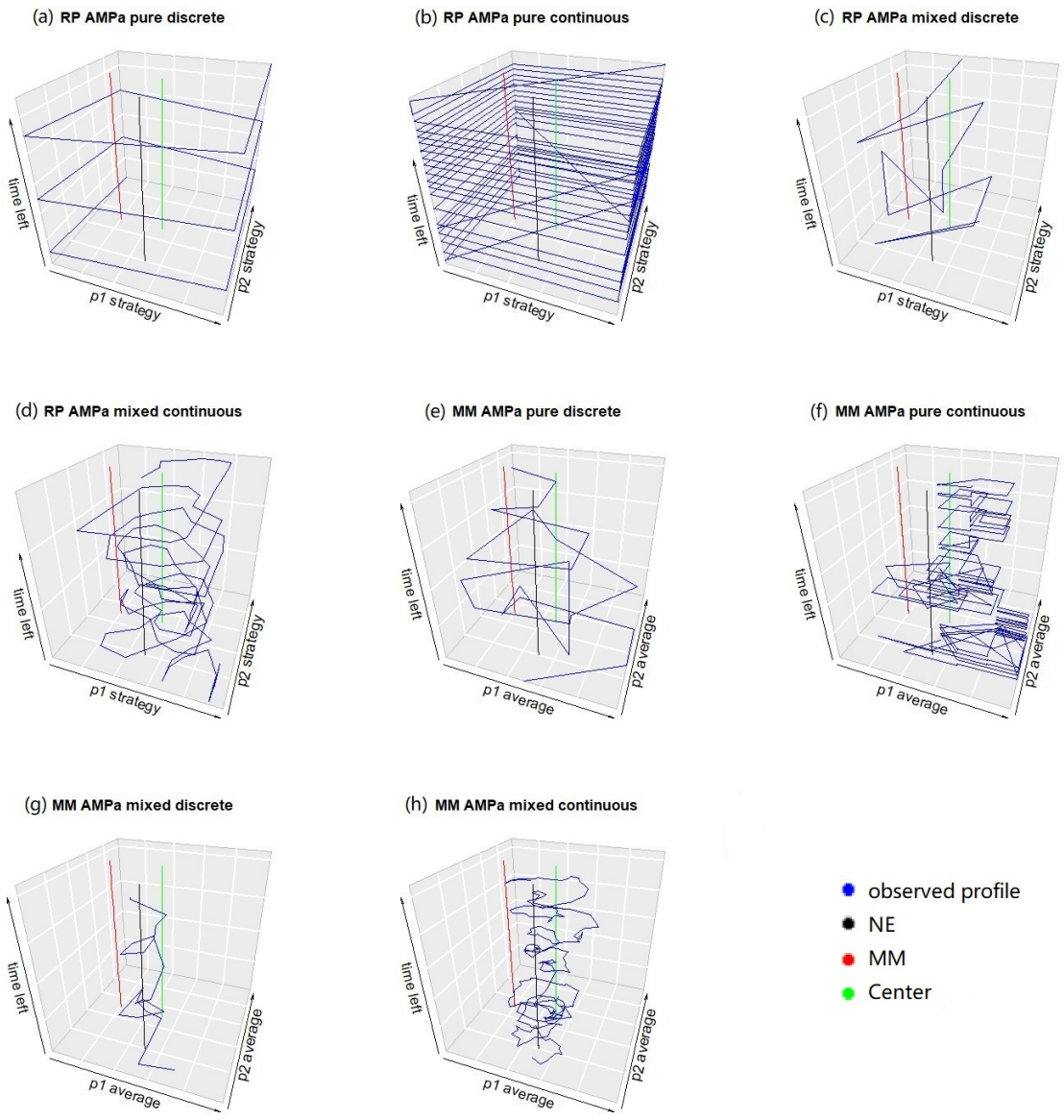

Figure 4: Sample time paths of strategy profiles. The horizontal plane is the state space, the (a, b) square. The vertical axis is time remaining, so the time paths begin at the top and spiral downward, and reach the bottom plane at the end of the period. Point predictions are time-invariant and therefore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

5.1 Point Predictions

To test point predictions, we focus on behavior after players have had some exposure to the current payoff matrix and after the transient effects of the random initialization wear off. Therefore subsequent analysis drops the first period in each block and the first 18 seconds (or 3 subperiods) of each period; with untrimmed data, results are similar but a bit noisier.

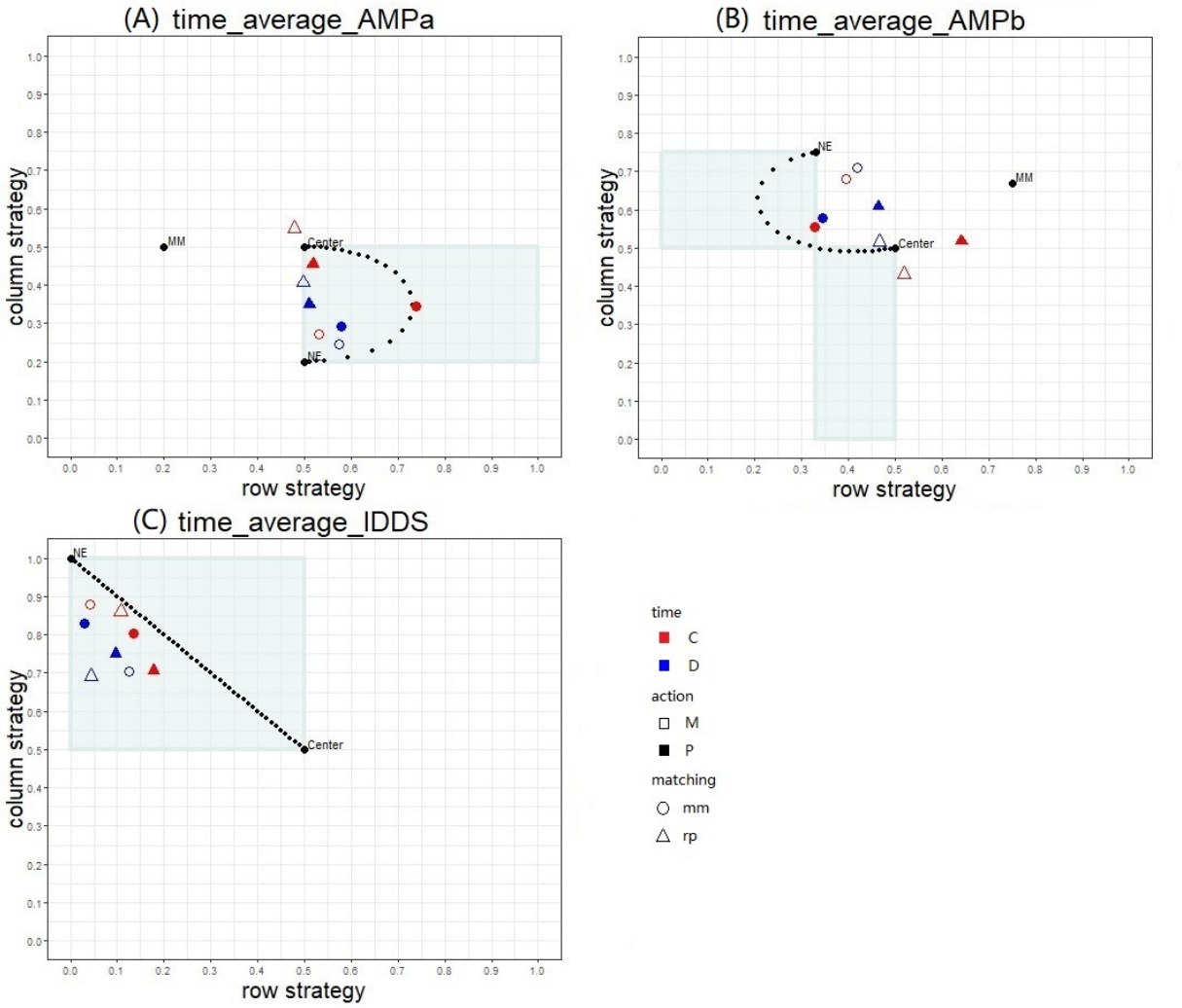

Figure 5: Scatter plots for all three games. Each colored symbol shows the average mixture for a treatment, e.g., the blue open triangle indicates observed play averaged over all instances of Discrete time × Mixed action × random pairwise matching (DMrp). Shaded rectangles represent the rank-dependent choice equilibrium (RDCE) sets. The dotted arc represents the set of symmetric logit QRE.

Panel A of Figure 5 shows that time average action vectors for most treatments in AMPa lie inside the RDCE region but near the left edge. The main exception is Pure Continuous mean matching, for which a point near the middle of the (symmetric logit) QRE arc is the best prediction. For the other seven treatments, QRE offers only slight improvement or no improvement over its extreme points, Center and NE. Panel B shows that only one of the eight treatments for AMPb lies in the RDCE region, and for only one other treatment does a non-trivial QRE offer a better prediction than Center or NE. Panel C shows that in IDDS, the average action vector in every treatment lies closer to NE than to Center. Here the QRE locus is a straight line, a diagonal of the square RDCE region. Interestingly, all the data lie below that line; perhaps in this dominance solvable game, the the first iteration from Row players is stronger than the second iteration from Column players. In any case, the Figure shows that, overall, QRE and RDCE have very limited predictive power for our data.

Figure 5 also offers evidence for testable hypotheses H1 and H2. In Panel A most of the circles (mean matching treatments) lie closer to NE than to Center, while the opposite is true for triangles (random pairwise treatments). In Panel B, only the two open circles (for CMmm and PMmm) are noticeably closer to NE; average actions for all random pairwise treatments are again closer to Center. In neither panel does any treatment have average behavior closer to MM than to an alternative point prediction. See Appendix Tables 6 and 7 for test statistics supporting these apparent regularities. To summarize,

Result 1. In our asymmetric matching pennies games, maximin is not an accurate prediction of average play under any conditions we examine. Nash equilibrium is the more accurate prediction in some population game mean matching treatments, especially those allowing explicit mixed actions. Center is consistently the best of our point predictions in our random pairwise treatments.

Table 3 tests hypothesis H2 and H3; see Appendix Table 8 for an extended version. Here and for the rest of this section we drop the IDDS data to focus on AMP games. In the first column of Table 3, the significantly negative entry for mm confirms hypothesis H2-NE(a) that NE is more accurate under mean matching than under random pairwise matching. The (barely significant at 5% level) positive entry for continuous contradicts H2-NE(c) — it seems that NE predicts time average behavior better in discrete than in continuous time. In the second column of the table, the large constant coefficient, together with the small other

	Distance to NE	Distance to MM	Distance to Center	Dispersion
continuous	0.12 (0.044)**	0.03 (0.210)	0.03 (0.079)*	-0.09 (0.056)*
pure	-0.00 (0.951)	-0.01 (0.730)	0.02 (0.206)	0.22 (0.001)***
mm	-0.17 (0.014)**	0.06 (0.109)	0.03 (0.159)	0.04 (0.150)
AMPa	-0.05 (0.156)	0.01 (0.555)	-0.01 (0.792)	0.03 (0.112)
Constant	0.29 (0.000)***	0.35 (0.000)***	0.20 (0.000)***	0.25 (0.000)***
#Instances	428	428	428	428
R-squared	0.185	0.034	0.015	0.591

Table 3: Coefficient estimates (and p values) for OLS regressions with subcluster wild bootstrap (MacKinnon and Webb, 2018). The bootstrap applies 999 replications with wild weights drawn from the Webb six-point distribution and clustering at the session level. Dependent variable in last column is geometric mean dispersion, and in other columns is Euclidean distance from mean observed profile to predicted profile. Independent variables are treatment indicators. #Instances is the number of pairs in rp and periods in mm in all sessions. Significance levels 1, 5, and 10% respectively denoted ***,** ,*.

coefficients, confirms that maximin prediction errors are large in all treatments. Entries in the third column indicate few consistent treatment effects for Center predictions. To summarize the main conclusions regarding H2, we have

Result 2. Time average-observed profiles are closer to NE under mean matching than under random pairwise matching, and in discrete than in continuous time. We find no consistent evidence to support maximin or Center treatment effects.

Hypothesis H3 concerns dispersion, defined in the previous section as the geometric mean of row and column standard deviations. The last column of Table 3 reports a regression with dispersion as the dependent variable. The negative coefficient in first line of that column (almost significant at the 5% level) supports H3(c), that dispersion is less, i.e., convergence is better, in continuous time than in discrete time. The significantly positive coefficient in second line supports H3(b), that dispersion is less with mixed action sets. To summarize,

Result 3. In our asymmetric matching pennies games, the data support hypotheses H3(b) and H3(c), but neither support nor reject H3(a). That is, convergence to a point (a "behavioral equilibrium") is stronger with explicit mixed action sets and in continuous time than with the more usual conditions of only pure action choices in discrete time.

5.2 Qualitative Dynamics

The large constant terms in Table 3 suggest 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?

	CW	CCW	DD	CD	Stay	#Instances
Panel A: AMPa games						
mm	0.443	0.133	0.224	0.062	0.139	24
rp	0.415	0.118	0.142	0.033	0.292	190
Mixed	0.411	0.152	0.166	0.061	0.210	122
Pure	0.428	0.077	0.132	0.003	0.361	92
Continuous	0.375	0.116	0.062	0.033	0.413	112
Discrete	0.466	0.123	0.249	0.039	0.124	102
Panel B: AMPb games						
mm	0.483	0.131	0.188	0.080	0.117	24
rp	0.369	0.118	0.139	0.025	0.349	190
Mixed	0.390	0.157	0.160	0.052	0.242	122
Pure	0.371	0.071	0.124	0.003	0.431	92
Continuous	0.303	0.109	0.057	0.025	0.506	112
Discrete	0.469	0.132	0.240	0.037	0.122	102

Table 4: Average fraction of directional changes by game and treatment.

To investigate, recall how Figure 2 classifies the direction of profile changes $\Delta s_t = (s_{Rt+1} - s_{Rt}, s_{Ct+1} - s_{Ct}) \neq 0$ as clockwise (CW), diagonal (DD), counterclockwise (CCW), or counter diagonal (CD). Table 4 begins with the fractions of directional changes (or Stay, i.e., no change) in each instance τ, and reports their averages across all instances of each treatment. To avoid repetitive asterisks, the Table entries do not indicate statistical significance, but in every case, we strongly reject the null hypothesis of no difference across row entries in favor of research hypothesis H4, that the CW fraction is greater.[5] Not surpris-

[5]More specifically, in each row a conservative test rejects the null hypothesis at the 1% level or better that the first column entry, for CW, is equal to each subsequent entry for an alternative direction. The test statistic, subcluster wild bootstrap (MacKinnon and Webb, 2018), uses 999 bootstrap replications, with wild weights drawn from the Webb six-point distribution and clustering at the session level.

ingly and not inconsistent with H4, Stay is relatively common in continuous time, and also with Pure action sets and with random pairwise matching. Diagonal moves (DD) are more common in the Discrete treatment than elsewhere, but even there are very significantly less common than clockwise moves (CW).

Result 4. Consistent with hypothesis H4, by far the most prevalent directional change in both asymmetric matching pennies games is clockwise (CW), where both player populations increase use of the higher payoff action. Diagonal (DD) moves are usually the next most common, especially in discrete time and with population mean matching.

The Online Appendix reports supplementary analysis supporting Result 4, including linear regressions (Table 9) and two bar plots (Figures 7 and 8) that show how the direction fractions change (slightly) over time.

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 using indicator variables D_k to capture treatment-specific response to regret,

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

Table 5 collects the results. The positive entries in the first row offer some support for H5(a); the estimated baseline (mixed action, random pairwise matching, AMPb bimatrix) response to regret β_1 is significantly positive for Row in discrete time setting and (for the one-sided test)[6] for Column in continuous time, but slightly misses one-sided 10% significance in the other baseline cases. Wald-tests on sums of coefficients reported in the Table confirm that, for Row in discrete time, response to regret is significantly positive at the two-sided 5% level for all treatments except AMPa-Mmm where it is still significant at the 10% level. The same is true for one-sided significance for Column players in continuous time. For the other two cases, Row-Continuous and Column-Discrete, the point estimates are all positive and one-sided significance is mostly in the vicinity of 10%; see Online Appendix Table 10.

[6]Reported p-values and significance our tables are all two-sided, but Hypothesis 5 is naturally one-sided — regret theory *requires* positive β, in contrast to some of our more conjectural hypotheses.

	Row-Continuous	Row-Discrete	Col-Continuous	Col-Discrete
β_1	0.10 (0.254)	1.17 (0.041)**	0.24 (0.072)*	0.96 (0.224)
pure	0.49 (0.173)	-0.16 (0.157)	0.58 (0.056)*	0.15 (0.685)
mm	0.40 (0.167)	-0.27 (0.154)	0.03 (0.864)	0.70 (0.339)
AMPa	-0.00 (0.986)	-0.29 (0.109)	-0.14 (0.128)	-0.50 (0.317)
pure_mm	-0.22 (0.415)	0.86 (0.053)**	-0.26 (0.090)*	-0.21 (0.631)
pure_AMPa	0.06 (0.743)	0.31 (0.127)	-0.17 (0.159)	-0.02 (0.950)
mm_AMPa	-0.32 (0.335)	0.22 (0.174)	-0.03 (0.883)	-0.36 (0.568)
Observations	63,648	4,272	63,648	4,272
R-squared	0.214	0.337	0.269	0.259
#Instances	332	276	332	276

Table 5: Directional Learning Model (5) coefficient estimates (and p values) for Row and Col(umn) player actions in Continuous and Discrete time. Observations show number of time ticks in each regression. Estimation is OLS with pair fixed effects and subcluster wild bootstrap (MacKinnon and Webb, 2018). The bootstrap uses 999 replications, with wild weights drawn from the Webb six-point distribution and clustering at the session level. Significance levels 1, 5, and 10% denoted ***,** ,* .

The 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 may adjust more slowly in AMPa because of the greater asymmetry in that game than in the AMPb baseline. Pure action sets seem to speed adjustment in continuous time where changes are asynchronous but may slow it in discrete time where changes are synchronized. Online Appendix Tables 11 and 12 report regressions for simplified regret models (3) and (4); results are generally consistent with those of Table 5.

Result 5. Consistent with hypothesis H5 and equation (5), both row and column players appear to respond positively to regret in all treatments.

5.4 Simulations and Long-run Dynamics

Following up on hypothesis tests, we ran simulations of equation (5) with AMPa games using the coefficient estimates reported in Table 5 with error terms set to zero. In every case, players' (or populations') strategy profile moves in clockwise cycles that converge to a limit cycle surrounding the Nash equilibrium. Figure 6 shows two examples using baseline parameter fits.

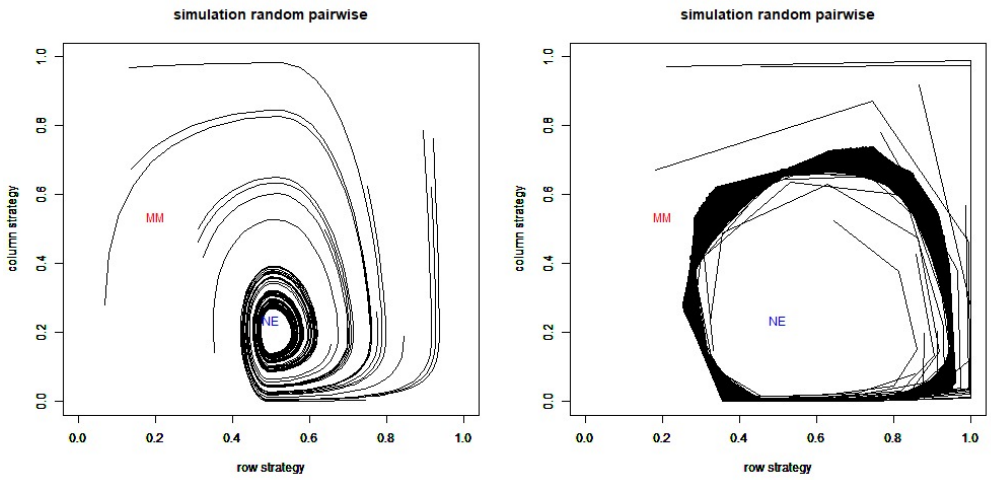

Figure 6: Simulations of (5) using baseline estimates from Table 5. Left panel uses small β_1's (0.1 for Row, 0.24 for column) from continuous time regressions and right panel uses large β_1s (1.17/0.96) from discrete time regressions. Simulations track the trajectory of mixed strategy profiles and are each run 500 periods from 20 different initial profiles.

Running simulations with a range of β_1 values (see Figures 9-13 in the Online Appendix) demonstrates that weaker response to regret (lower $\beta_1 > 0$) produces limit cycles that are, of course, slower but also that more tightly circle around the NE. That is, NE is a better prediction of simulated behavior when β_1 is small, in that the time average is closer to NE and also there is less dispersion around the time average.

This regularity is in some ways surprising. Circumstances that favor stronger response to regret (larger β_1) degrade the predictive power of NE — abstracting from behavioral noise, they lead to limit cycles that are closer to the edges of the square, and time average behavior closer to the center but with larger dispersion around that time average. Conversely, with weaker response to regret, the simulated state stays closer to NE, so the time average is even

closer and dispersion around that time average is small. We conjecture that the same may be true of actual human behavior.

Conjecture 1. Much of the impact of environmental treatments on the predictive power of Nash equilibrium is captured in the responsiveness β_1 to regret; weaker positive response is associated with greater predictive power.

We have only 16 estimates of row and column β_1's (for 2 payoff matrices $\times 2^3$ environments), and the conjecture arose only after looking at those estimates. Therefore we can not yet properly test the Conjecture, and can only say that it is not inconsistent with our data.

6 Discussion

Our results speak to the broad questions behind the testable hypotheses, and suggest practical advice for applied researchers.

First, we find that mixed Nash equilibrium is a reasonably good predictor of behavior in population games. For example, consider an application to the market for an intermediate good where upstream firms deal with many downstream firms, and conversely. When the market model has no pure NE, our results suggest that mixed Nash equilibrium may be a good candidate for describing time-average behavior.

Second, for most other circumstances in asymmetric matching pennies games, we find that uniform random behavior outpredicts mixed Nash equilibrium. In an otherwise similar application where a given upstream firm deals mainly with a single downstream firm and conversely, the mixed NE may not be the best predictor of time average behavior.

Third, for our asymmetric matching pennies games, we find no circumstance in which maximin predicted behavior better than either Nash equilibrium or uniform mixes. This may simplify work for applied economists.

Applied economists interested in short run dynamics should find our experiment particularly instructive. We find persistent dispersion and systematic directional change in our data. A regret-based directional learning model has significant explanatory power, and its coefficients were fairly consistent across treatments. Applied economists may be able to de-

ploy models that similarly aim to capture directional adaptations that are sensitive to regret (i.e., to the magnitude of potential gains).

Our results also suggest the sort of long run behavior one might expect in games with no pure Nash equilibrium. Simulations of regret-based dynamics suggest that behavior never settles down to any sort of mixed equilibrium profile. Instead, abstracting from noise, simulated action profiles converge to a limit cycle around Nash equilibrium. When the cycle amplitude is small, Nash equilibrium is not a bad predictor, but when the amplitude is large, the Center of the state space is a better predictor of the time average, albeit with large dispersion. The cycle amplitude is systematically related to sensitivity to regret, which in turn is related to the environment. We conjecture that environmental circumstances influence convergence to Nash equilibrium via their impact on responsiveness to regret: perhaps surprisingly, weaker positive responsiveness is associated with better convergence!

Our main message for theorists is that mixed equilibrium (Nash or otherwise) is relevant over a narrower range of circumstances than might have been supposed. We find rapid approximate convergence to pure Nash equilibrium in a (control) dominance solvable game, and to a lesser extent we find convergence to mixed NE in matching pennies population games. However, in traditional pairwise matching (especially with traditional pure strategy choice) for matching pennies, we find persistent dispersion that centers more on the uniform random mix than on an equilibrium mix.

We see dynamic regularities that need better theoretical explanation. In our asymmetric matching pennies games, typically one player (or player population) has a stronger incentive to switch strategies, and doing so strengthens that incentive in the other player or population, creating (with our sign conventions) clockwise (CW) cycles.[7]

Our experiment also suggests possible followup work in the laboratory. Our subjects explicitly choose their mixes in M treatments and receive the expected payoff, and so do not need to randomize actions dynamically. In contrast, Romero and Rosokha (2018) elicit subjects' history-dependent pure actions in a repeated prisoners' dilemma. It might be

[7]One might regard the next most common move, diagonal (DD), as indicating attempts to exploit those cycles, i.e., to outguess opponents. Our data, however, seem more consistent with the less interesting interpretation that DD moves are mainly due to relatively infrequent sampling that combines two successive adaptations to regret.

interesting to try similar elicitation procedures in repeated matching pennies games. It may also be worthwhile to seek new treatments that facilitate convergence to Nash equilibrium or other point predictions. In some pilot sessions, we tried displays of best and worst possible payoffs (the latter intended to help maximin predictions), and tried slowing adjustment speed in continuous time, but found little impact. Finally, it seems worthwhile to design experiments that can better distinguish among different models of adaptive learning.

We hope especially that our results encourage applied researchers to work in a more nuanced fashion with mixed strategy equilibrium and adjustment dynamics. Biologists since Lotka and Volterra (Lotka, 1925) have recognized that dynamics are crucial to understanding asymmetric 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 asymmetric matching pennies population game, and our work suggests how adaptive dynamics could supplement equilibrium analysis.

References

[1] Axel Anderson, J Rosen, J Rust, and Ping Wong. Disequilibrium play in tennis. 2020.

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

[3] George W Brown. Iterative solution of games by fictitious play. *Activity Analysis of Production and Allocation*, 13(1):374–376, 1951.

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

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

[6] 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.

[7] Hsiao-Chi Chen, James W Friedman, and Jacques-Francois Thisse. Boundedly rational nash equilibrium: a probabilistic choice approach. *Games and Economic Behavior*, 18(1):32–54, 1997.

[8] Yin-Wong Cheung and Daniel Friedman. Individual learning in normal form games: some laboratory results. *Games and Economic Behavior*, 19(1):46–76, 1997.

[9] 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.

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

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

[12] Sean Duffy, JJ Naddeo, David Owens, and John Smith. Cognitive load and mixed strategies: on brains and minimax. 2016.

[13] 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.

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

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

[16] 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.

[17] Daniel Friedman, Steffen Huck, Ryan Oprea, and Simon Weidenholzer. From imitation to collusion: long-run learning in a low-information environment. *Journal of Economic Theory*, 155:185–205, 2015.

[18] Romain Gauriot, Lionel Page, and John Wooders. Nash at wimbledon: evidence from half a million serves. *Available at SSRN 2850919*, 2016.

[19] Jacob K Goeree, Charles A Holt, Philippos Louis, Thomas R Palfrey, and Brian Rogers. Rank-dependent choice equilibrium: a non-parametric generalization of qre. *Handbook of Research Methods and Applications in Experimental Economics*, 2019.

[20] 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.

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

[22] Morris W Hirsch, Stephen Smale, and Robert L Devaney. *Differential equations, dynamical systems, and an introduction to chaos*. Academic Press, 2012.

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

[24] Ed Hopkins. Two competing models of how people learn in games. *Econometrica*, 70(6):2141–2166, 2002.

[25] 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.

[26] 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.

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

[28] James G MacKinnon and Matthew D Webb. The wild bootstrap for few (treated) clusters. *The Econometrics Journal*, 21(2):114–135, 2018.

[29] Richard D McKelvey and Thomas R Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38, 1995.

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

[31] Dilip Mookherjee and Barry Sopher. Learning and decision costs in experimental constant sum games. *Games and Economic Behavior*, 19(1):97–132, 1997.

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

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

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

[35] 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.

[36] Ryan Oprea, Keith Henwood, and Daniel Friedman. Separating the hawks from the doves: evidence from continuous time laboratory games. *Journal of Economic Theory*, 146(6):2206–2225, 2011.

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

[38] Edgar Allan Poe. *Selected tales*. Penguin UK, 2007 (originally 1844).

[39] Vitaly Pruzhansky. Maximin play in completely mixed strategic games. *Theory and Decision*, 75(4):543–561, 2013.

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

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

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

[43] David Roodman, Morten Ørregaard Nielsen, James G MacKinnon, and Matthew D Webb. Fast and wild: Bootstrap inference in stata using boottest. *The Stata Journal*, 19(1):4–60, 2019.

[44] Reinhard Selten. Anticipatory learning in two-person games. In *Game Equilibrium Models I*, pages 98–154. Springer, 1991.

[45] Lloyd Shapley. Some topics in two-person games. *Advances in Game Theory*, 52:1–29, 1964.

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

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

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

[49] Fang-Fang Tang. Anticipatory learning in two-person games: some experimental results. *Journal of Economic Behavior & Organization*, 44(2):221–232, 2001.

[50] Peter D Taylor and Leo B Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1-2):145–156, 1978.

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

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

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

7 Online Appendix: Supplementary Analysis

Computation of NE, MM and Set-valued Predictions for AMPa.

Here we use the AMPa game to illustrate the computation of NE, and maximin. We also show how to compute the QRE arc (shown in Figure 1) and the RDCE region.

To calculate Nash equilibrium, recall from Section 3 the notation (a, b) for mixed strategy profiles and the payoff difference functions D_R and D_C between pure strategies for row and column players, respectively. By definition, the unique mixed Nash equilibrium for AMPa is the solution to the system

$$\begin{aligned} 0 = D_R &= 1000b - 200 \\ 0 = D_C &= 200 - 400a. \end{aligned} \tag{6}$$

Hence $(a_{NE}, b_{NE}) = (0.5, 0.2)$.

The maximin problem for row players is:

$$\max_a \min\{f_R(a, 1), f_R(a, 0)\}, \tag{7}$$

where $f_R(a, b) = 1000ab - 200a - 200b + 200$ from equation (1). Since $f_R(a, 1)$ (resp. $f_R(a, 0)$) increases (resp. decreases) linearly in a, the max must occur where $f_R(a, 1) = f_R(a, 0)$, yielding $a_{MM} = 0.2$. The analogous problem for column players yields $b_{MM} = 0.5$.

To calculate the symmetric QRE(λ) curve for both players, recall that the logit payoff function in AMPa games is implicitly defined by the fixed point equations

$$a = \frac{\exp(\lambda f_R(1, b))}{\exp(\lambda f_R(1, b)) + \exp(\lambda f_R(0, b))} \tag{8}$$

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

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 is shown in Figure 1 as well as in Panel A of Figure 5.

For 2x2 bimatrix games, RDCE is defined by the following system of inequalities:

$$\text{sgn}D_R = \text{sgn}[a - \frac{1}{2}], \text{and} \;\; \text{sgn}D_C = \text{sgn}[b - \frac{1}{2}]. \tag{10}$$

From equation (6) we see that $\text{sgn}D_R = \text{sgn}[b - \frac{1}{5}]$ and $\text{sgn}D_C = -\text{sgn}[a - \frac{1}{2}]$ for the game AMPa. Plugging this into (10), we obtain the rectangular shaded region in Panel A of Figure 5. Similarly we obtain the non-convex shaded region in Panel B of Figure 5 as the RDCE for AMPb, where $\text{sgn}D_R = \text{sgn}[b - \frac{3}{4}]$ and $\text{sgn}D_C = -\text{sgn}[a - \frac{1}{3}]$.

Point Prediction Details

	row to NE	col to NE	row to MM	col to MM	row to Center	col to Center
Panel A: AMPa games						
mm	0.105*	0.089**	0.405***	-0.211***	0.105*	-0.211***
rp	-0.001	0.250**	0.299***	-0.050	-0.001	-0.050
Mixed	-0.006	0.262	0.294***	-0.038	-0.006	-0.038
Pure	0.033	0.191*	0.333***	-0.109	0.033	-0.109
Continuous	0.010	0.290**	0.310***	-0.010	0.010	-0.010
Discrete	0.012	0.168**	0.312***	-0.132**	0.012	-0.132**
Panel B: AMPb games						
mm	0.041	-0.118**	-0.379***	-0.038	-0.129***	0.132**
rp	0.190**	-0.240***	-0.230***	-0.160**	0.020	0.010
Mixed	0.157**	-0.257*	-0.263***	-0.177	-0.013	-0.007
Pure	0.195	-0.186**	-0.225*	-0.106*	0.025	0.064
Continuous	0.216*	-0.267**	-0.204**	-0.187*	0.046	-0.017
Discrete	0.126***	-0.182**	-0.294***	-0.102	-0.044**	0.068
Panel C: IDDS games						
mm	0.083**	-0.196**	0.083**	-0.196**	-0.417***	0.304***
rp	0.103**	-0.240***	0.103**	-0.240**	-0.397***	0.260**
Mixed	0.079*	-0.215*	0.079*	-0.215*	-0.421***	0.285*
Pure	0.129*	-0.261***	0.129*	-0.261***	-0.371***	0.239***
Continuous	0.130**	-0.197*	0.130**	-0.197*	-0.370***	0.303**
Discrete	0.068*	-0.276***	0.068*	-0.276***	-0.432***	0.224***

Table 6: Mean signed deviation of row and column players' time average profiles from predictions, by treatment. Asterisks indicate p-values of .10(*), .05(**) and .01(***) for subcluster wild bootstrap from MacKinnon and Webb (2018) (999 replications, wild weight drawn from the Webb six-point distribution, clustered at the session level).

Table 6 reports direct tests of the part (a) of hypotheses H1. It begins with the time

average profiles (a_τ, b_τ) of each instance τ of a given treatment, computes their mean $(\bar{a}, \bar{b})$ and subtracts the given point prediction (a_p, b_p). The mean matching (mm) population treatment of the AMPb game comes the closest to the NE prediction — as shown in the first line of Panel B, its mean row player action $\bar{a}$ is an insignificant 0.041 above the NE prediction $a_{NE} = 0.5$, while the column player mean action is 0.118 below $b_{NE} = 0.75$ and differs at the 5% significance level. Similarly, the mm treatment of the AMPa game is also the closest to the NE prediction. In every other game and treatment, we clearly reject H1-NE(a) and H1-MM(a) for row players or column players or both. On the other hand, we reject H1-Center(a) in only a few cases for either AMP game.

Less importantly, Panel C shows that for the dominance solvable game, all point predictions are rejected. Of course, for this game NE/MM is a corner of the state space, so deviations from that point all have the same sign and cannot cancel out. Despite not suffering that disadvantage, the Center prediction is even worse.

In order to test part (b) of hypotheses H1, Table 7 begins with the Euclidean distance $\sqrt{(a_\tau - a_p)^2 + (b_\tau - b_p)^2}$ between a point prediction and an instance of time average behavior, and reports the mean distances by treatment and point prediction. For example, the first line of the Table 7 shows that for mean matching instances in the AMPa game (24 periods, pooling over Continuous and Discrete time, and over Pure and Mixed action sets), the mean Euclidean distance between the time-average profile and the NE prediction is just 0.150. This is significantly ($p < 0.05$) less than 0.255, the mean distance between those same time average profiles and the Center point. The rest of that line shows that the mean distance to the maximin prediction, 0.463, is significantly larger. The next line of the Table shows that the time average distance to the Center is, among the 190 player pairs in rp treatments, on average 0.206, significantly less than the distance to either equilibrium point.

Thus the first lines of Panels A and B in Table 7 support hypothesis H1-NE(b), that NE is the best point prediction, for mean-matching treatments in our matching pennies games. However, consistent with hypothesis H1-Center(b), the non-equilibrium point prediction Center is best in all other treatments in these games, and significantly in random pairwise matching treatments and in continuous time treatments. In Panels A and B, there is no support for the maximin hypothesis H1-MM(b). Panel C confirms that for the dominance solvable game, the pure strategy NE/MM is a better point prediction than Center in all

	Distance to NE		Distance to Center		Distance to MM	#Instances
Panel A: AMPa games						
mm	0.150	$<_{**}$	0.255	$<_{***}$	0.463	24
rp	0.305	$>_{*}$	0.206	$<_{***}$	0.372	190
Mixed	0.307	$>$	0.216	$<_{***}$	0.375	122
Pure	0.261	$>$	0.205	$<_{***}$	0.391	92
Continuous	0.339	$>_{*}$	0.204	$<_{***}$	0.382	112
Discrete	0.231	$>$	0.220	$<_{***}$	0.382	102
Panel B: AMPb games						
mm	0.157	$<_{*}$	0.224	$<_{***}$	0.398	24
rp	0.356	$>_{**}$	0.220	$<_{***}$	0.365	190
Mixed	0.325	$>$	0.202	$<_{***}$	0.380	122
Pure	0.345	$>$	0.244	$<_{**}$	0.354	92
Continuous	0.394	$>_{**}$	0.253	$<_{*}$	0.393	112
Discrete	0.268	$>$	0.183	$<_{**}$	0.343	102
Panel C: IDDS games						
mm	0.219	$<_{**}$	0.521	$>_{**}$	0.219	12
rp	0.300	$<_{**}$	0.548	$>_{**}$	0.300	95
Mixed	0.257	$<_{**}$	0.561	$>_{**}$	0.257	61
Pure	0.335	$<_{*}$	0.524	$>_{**}$	0.335	46
Continuous	0.283	$<_{*}$	0.559	$>$	0.283	56
Discrete	0.299	$<_{***}$	0.529	$>_{***}$	0.299	51

Table 7: Mean distance of time average profiles to predictions by treatments. Subscripted asterisks indicate p-values of .10(*), .05(**) and .01(***) for subcluster wild bootstrap from MacKinnon and Webb (2018) (999 replications, wild weight drawn from the Webb six-point distribution, clustered at the session level).

treatments.

Table 8 tests hypothesis H2 and H3 with the interaction terms of the treatment indicators. In the first column of Table 8, the significantly negative entry for the mm dummy variable supports hypotheses H2-NE(a), that NE is more accurate under mean matching than under random pairwise matching. In the second column of the table, the large constant coefficient, together with mostly small other coefficients (often with offsetting interactions) confirm that maximin prediction errors are large in all treatments. Entries in the third column indicate no consistent treatment effect for Center predictions, which is similar to Result 2 of the main text. The last column also confirms Result 3 that the dispersion is larger in

discrete time than in continuous time, and in pure actions sets than in mixed action sets.

To summarize, the three tables statistically support Results 1-3 in Section 5.1.

	Distance to NE	Distance to MM	Distance to Center	Dispersion
continuous	0.11	0.04	0.06	-0.13**
	(0.157)	(0.163)	(0.179)	(0.031)
pure	-0.01	-0.04*	0.03	0.20***
	(0.788)	(0.083)	(0.361)	(0.010)
mm	-0.21**	-0.02	0.01	0.05
	(0.023)	(0.587)	(0.814)	(0.107)
AMPa	-0.01	0.01	0.06	0.01
	(0.778)	(0.209)	(0.135)	(0.817)
continuous_pure	0.06	0.01	0.02	0.04
	(0.231)	(0.928)	(0.691)	(0.185)
continuous_mm	-0.10	0.01	-0.01	0.06
	(0.132)	(0.874)	(0.818)	(0.115)
continuous_AMPa	-0.02	-0.05*	-0.09	0.04
	(0.746)	(0.073)	(0.156)	(0.380)
pure_mm	0.12	0.10	-0.01	-0.05
	(0.109)	(0.336)	(0.871)	(0.143)
pure_AMPa	-0.07	0.04	-0.06	0.01
	(0.415)	(0.119)	(0.284)	(0.901)
mm_AMPa	0.05	0.06	0.05	-0.03
	(0.340)	(0.177)	(0.219)	(0.250)
Constant	0.29**	0.36***	0.17***	0.28***
	(0.015)	(0.000)	(0.002)	(0.009)
#Instances	428	428	428	428
R-squared	0.216	0.067	0.047	0.607

Table 8: Coefficient (p values) for Table 3 with intersections terms of treatment indicators and subcluster wild bootstrap (MacKinnon and Webb, 2018). The bootstrap applies 999 replications with wild weights drawn from the Webb six-point distribution and clustering at the session level. Dependent variable in last column is geometric mean dispersion, and in other columns is Euclidean distance from mean observed profile to predicted profile. #Instances is the number of pairs in rp and periods in mm in all sessions. Significance levels 1, 5, and 10% respectively denoted ***,** ,*.

Qualitative Dynamics Details

Table 9 reinforces Table 4 with regressions and show potential treatment effects. The dependent variables in both tables refer to the fraction of time players play each type of cycles at the instance level. The results are essentially similar to what we have found in the main text. Figures 7 and 8 display how the fraction of time moves over time within the periods and show that the fractions are quite stable over time.

	CW	CCW	Diagonal	Stay
continuous	-0.17*	-0.01	-0.20***	0.40**
	(0.078)	(0.354)	(0.001)	(0.026)
pure	-0.03	-0.06*	-0.04	0.19*
	(0.735)	(0.071)	(0.123)	(0.067)
mm	0.08	-0.07	0.10*	-0.18
	(0.657)	(0.343)	(0.081)	(0.168)
AMPa	-0.01	-0.01	0.00	0.01
	(0.612)	(0.272)	(0.893)	(0.718)
continuous_pure	-0.01	-0.06	0.02	0.03
	(0.973)	(0.414)	(0.191)	(0.891)
continuous_mm	0.03	0.11	0.02	-0.18
	(0.889)	(0.290)	(0.232)	(0.495)
continuous_AMPa	0.08**	0.02	-0.00	-0.10*
	(0.026)	(0.301)	(0.917)	(0.100)
pure_mm	0.04	0.07	-0.12**	0.06
	(0.883)	(0.414)	(0.028)	(0.801)
pure_AMPa	0.04	0.01	0.00	-0.04
	(0.161)	(0.479)	(0.960)	(0.233)
mm_AMPa	-0.09*	0.00	0.03	0.08
	(0.100)	(0.967)	(0.493)	(0.246)
Constant	0.47**	0.17***	0.25***	0.05
	(0.019)	(0.000)	(0.004)	(0.153)
#Instances	428	428	428	428
R-squared	0.127	0.198	0.383	0.459

Table 9: Regressions of directional changes on treatment dummies. Dependent variable observations are instance level averages. Baseline (omitted) indicators are for AMPb, mixed strategy, discrete time, and random pairwise matching. Significance level: *** 0.01, ** 0.05, * 0.1 for subcluster wild bootstrap from MacKinnon and Webb (2018) (999 replications, wild weight drawn from the Webb six-point distribution, clustered at the session level).

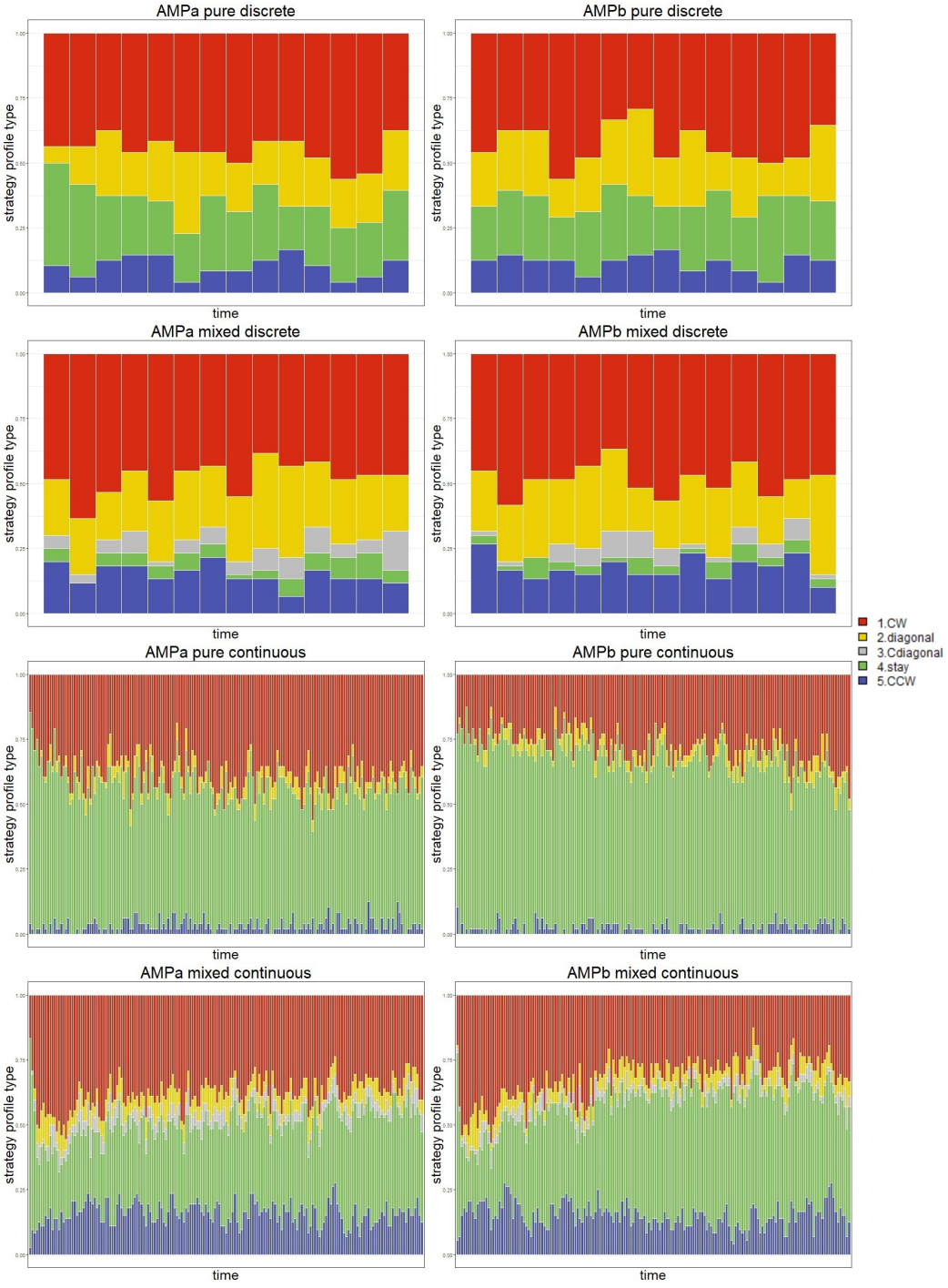

Figure 7: Classification over time in random pairwise matching sessions: clockwise (red), diagonal (yellow), counter-diagonal (grey), stay (green) and counter-clockwise (blue). AMPa on left, and AMPb on right, by subperiod or tick within period and averaged across periods.

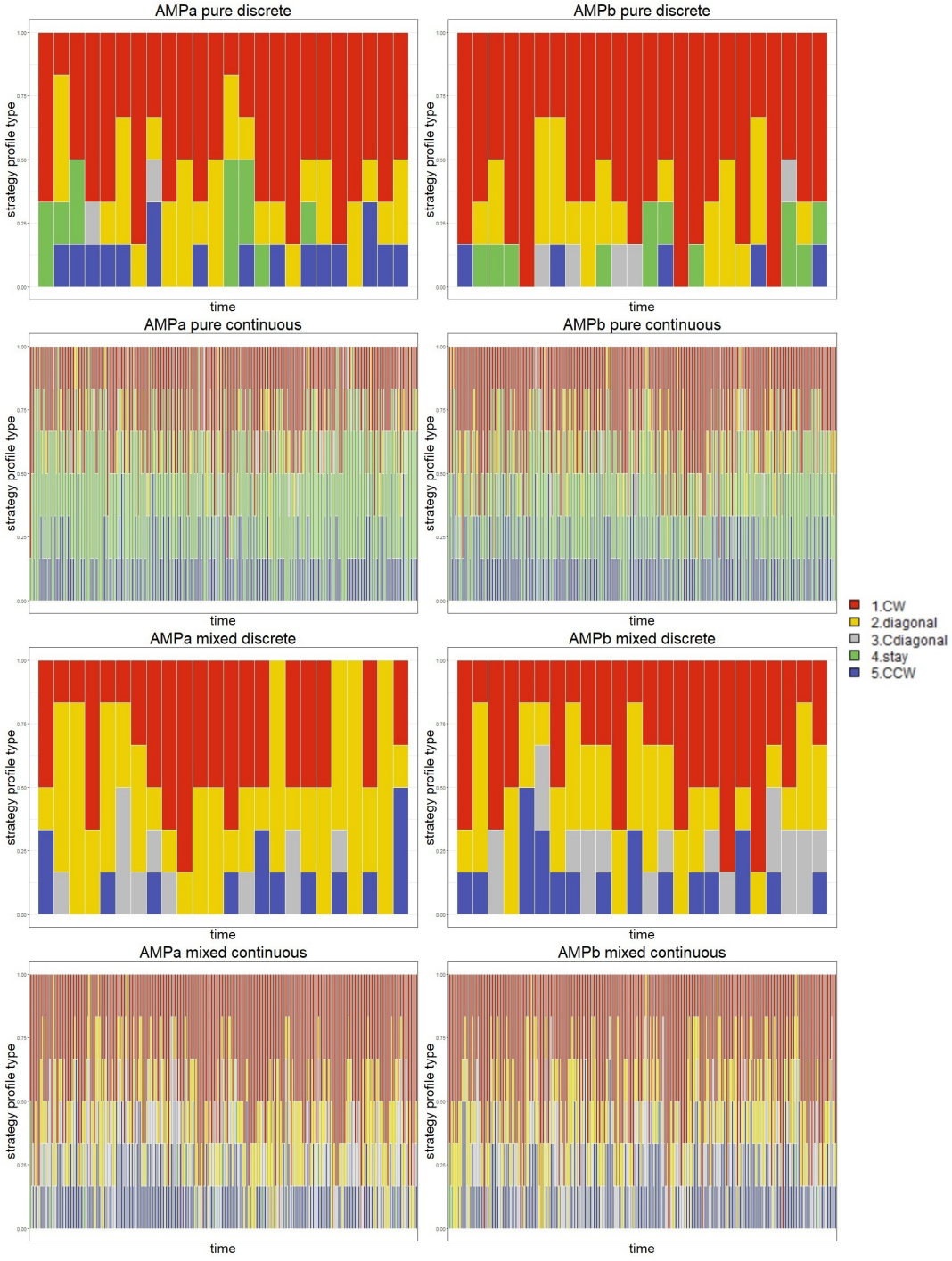

Figure 8: Classification over time in mean matching sessions: clockwise (red), diagonal (yellow), counter-diagonal (grey), stay (green) and counter-clockwise (blue). AMPa on left, and AMPb on right, by subperiod or tick within period and averaged across periods.

Fitted Dynamic Model Details

To close a minor gap in specifying equation (2), consider states where $argmax_x f_i(x, s_{-it})$ includes some values larger than s_{it} and other values smaller than s_{it}, e.g., at NE, the indifference condition tells us that both pure strategies are included in $argmax$. In such cases, the sign of directional change is ambiguous so we adopt the natural convention that $\text{sgn}\{\hat{s}_{it} - s_{it}\} = 0$.

	Row-Continuous	Row-Discrete	Col-Continuous	Col-Discrete
AMPb-Mrp	.301	.042**	.091*	.251
AMPb-Prp	.157	.035**	.035**	.188
AMPb-Mmm	.145	.050**	.073*	.140
AMPb-Pmm	.155	.028**	.062*	.144
AMPa-Mrp	.173	.039**	.071*	.216
AMPa-Prp	.158	.030**	.054*	.210
AMPa-Mmm	.422	.060*	.141	.287
AMPa-Pmm	.158	.022**	.096*	.219

Table 10: Wald tests on sums of coefficients reported in Table 5. The table reports the p values of the tests. e.g., AMPa-Prp is related to the sum of the coefficients including the baseline β_1 and the treatment regressors related to AMPa game, Pure action sets, and random pairwise matching. The sum of the coefficients can be calculated from Table 5.

	Row-Continuous	Row-Discrete	Col-Continuous	Col-Discrete
β_1	0.02	0.43*	0.05*	0.40**
	(0.743)	(0.060)	(0.073)	(0.034)
pure	0.41	0.18	0.40**	0.21*
	(0.136)	(0.128)	(0.016)	(0.076)
mm	0.04	-0.19	-0.02	-0.15
	(0.744)	(0.143)	(0.214)	(0.102)
AMPa	0.05	0.01	0.00	-0.02
	(0.678)	(0.802)	(0.893)	(0.523)
pure_mm	-0.23	0.28*	-0.29**	0.07
	(0.223)	(0.094)	(0.028)	(0.231)
pure_AMPa	-0.02	0.06	0.05	-0.00
	(0.833)	(0.446)	(0.169)	(0.940)
mm_AMPa	-0.10	-0.10	-0.01	0.05
	(0.587)	(0.294)	(0.748)	(0.337)
Observations	63,648	4,272	63,648	4,272
R-squared	0.267	0.375	0.281	0.289
#Instances	332	276	332	276

Table 11: BR Learning Model (eq (3)) coefficient estimates for Row and Col(umn) player actions in Continuous and Discrete time. Observations show number of time ticks in each regression. Least squares with pair fixed effects and subcluster wild bootstrap (MacKinnon and Webb, 2018). The bootstrap applies 999 replications, wild weight drawn from the Webb six-point distribution, and is clustered at the session level. Omitted baseline indicators are for AMPb, mixed strategy, discrete time, and random pairwise matching. Nominal significance levels 1, 5, and 10% denoted ***,**,*.

	Row-Continuous	Row-Discrete	Col-Continuous	Col-Discrete
β_1	0.01	0.19*	0.03*	0.15***
	(0.827)	(0.086)	(0.061)	(0.002)
pure	0.44	0.42*	0.42**	0.45***
	(0.115)	(0.054)	(0.019)	(0.001)
mm	0.02	-0.10	-0.01	-0.05***
	(0.728)	(0.171)	(0.176)	(0.001)
AMPa	0.03	0.04	0.00	0.00
	(0.647)	(0.427)	(0.956)	(0.110)
pure_mm	-0.23	0.17	-0.30**	-0.02***
	(0.204)	(0.125)	(0.021)	(0.007)
pure_AMPa	-0.02	0.03	0.05**	-0.01**
	(0.841)	(0.532)	(0.177)	(0.020)
mm_AMPa	-0.06	-0.09	-0.00	0.03***
	(0.602)	(0.227)	(0.854)	(0.008)
Observations	63,648	4,272	63,648	4,272
R-squared	0.266	0.341	0.281	0.252
#Instances	332	276	332	276

Table 12: Pure Directional Learning Model (eq (4)) coefficient estimates for Row and Col(umn) player actions in Continuous and Discrete time. Observations show number of time ticks in each regression. Least squares with pair fixed effects and subcluster wild bootstrap (MacKinnon and Webb, 2018). The bootstrap applies 999 replications, wild weight drawn from the Webb six-point distribution, and is clustered at the session level. Omitted baseline indicators are for AMPb, mixed strategy, discrete time, and random pairwise matching. Nominal significance levels 1, 5, and 10% denoted ***,** ,* .

Limit Cycle Details

Here we analyze limit cycles numerically with AMPa games by collecting distance between observation and Nash equilibrium over time when the trajectory crosses the Poincaré section from the simulation data. We also show how the limit cycle varies with the change of β_1 in eq (5).

Given strategy profile (a_t, b_t), the Poincaré section we select is the isocline $(a = 0.5, b > 0.2)$. The trajectory (a_t, b_t) at time t is considered to cross the isocline when $b_t > 0.2$ and $(a_t - 0.5)(a_{t+1} - 0.5) \leq 0$ are satisfied. Whenever the trajectory crosses the isocline, the distance between the observation and Nash Equilibrium $(d = \sqrt{(a_t - 0.5)^2 + (b_t - 0.2)^2})$ is recorded and graphed. Figure 9 ($\beta_1 = 0.1$ for row players and $= 0.24$ for column players) and 10 ($\beta_1 = 1$ for both players) overlaps such distance over time with 20 simulations and 1500 iterations in each simulation. The blue (red) lines shows the distance over time when trajectory crosses the isocline and the initial position is outside (inside) the limit cycle. As shown in the figures, both blue and red lines converge to a limit cycle, which proves that instead of converging to the Nash equilibrium, the dynamics converge to a limit cycle that is affected by the payoff matrices and the speed of adjustment β_1. The result is robust for other 3 isoclines between Nash Equilibrium and the boundary of the action space.

Figure 11 shows how the time average (the center of the limit cycle) changes with β_1 and Figure 12 shows approximately how the radius of the limit cycle changes with β_1 using both NE and Center as reference points. Although there exists non-monotonicity when β_1 is large enough so that trajectories start to hit the edges of the square, we can find that the center of the limit cycle moves from NE to Center when β_1 increases, and the radius of the limit cycle increases with β_1. The figures confirm the results in Section 5.4.

When β_1 increases, the boundary of the limit cycle will eventually hit the boundary of the state space, which reshape the limit cycle. The reshaping process causes the non-monotonicity in Figures 11 and 12 (when β_1 is between 2 and 3), as the cycle is not monotonously expanding when its boundary hits a subset of the boundary of the state space. Figure 13 shows three typical limit cycles. Readers can get a sense of how the cycle looks like from the figure.

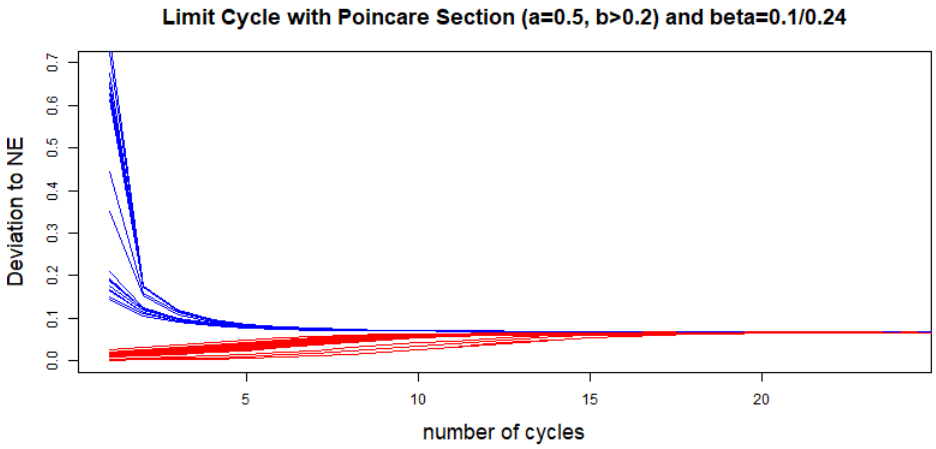

Figure 9: The figure shows the distance between observations and Nash equilibrium over time with 20 simulations and 1500 iterations in each simulation. The blue (red) lines shows the distance over time when trajectory crosses the isocline and the initial position is outside (inside) the limit cycle.

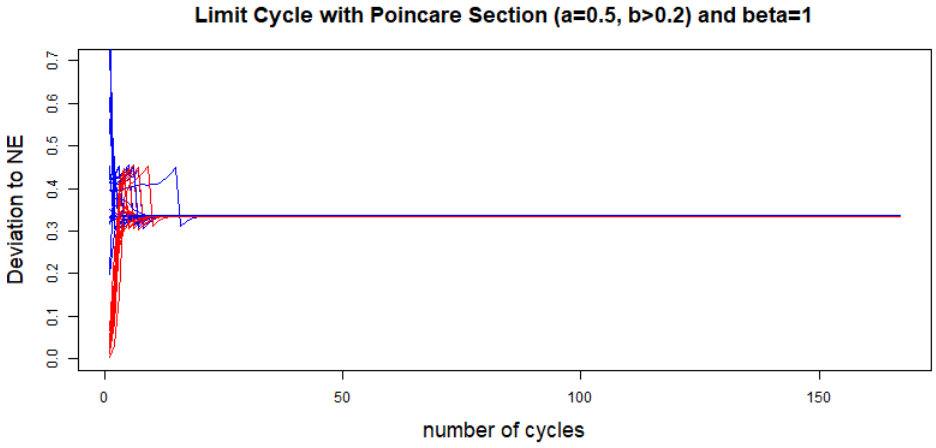

Figure 10: The figure shows the distance between observations and Nash equilibrium over time with 20 simulations and 1500 iterations in each simulation. The blue (red) lines shows the distance over time when trajectory crosses the isocline and the initial position is outside (inside) the limit cycle.

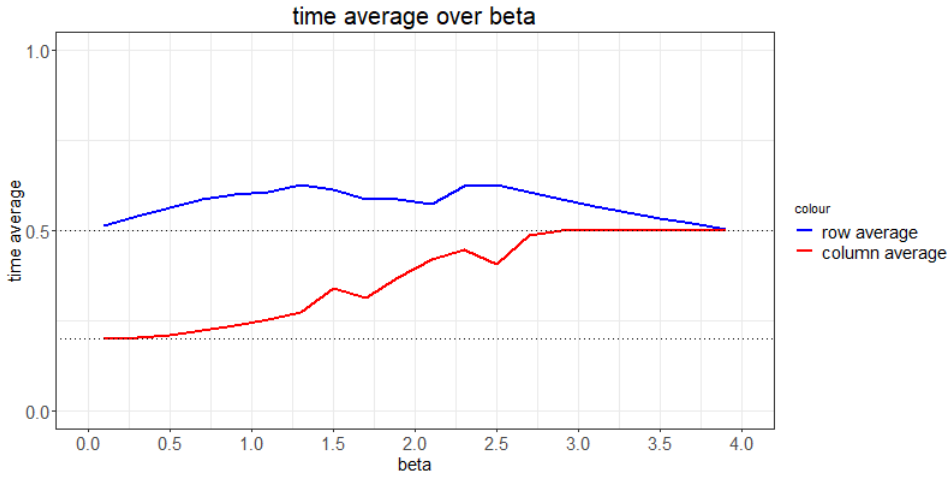

Figure 11: Time average over β_1 with 20 simulations (1500 iterations in each simulation, data is averaged by the last 500 iterations) for each β_1.

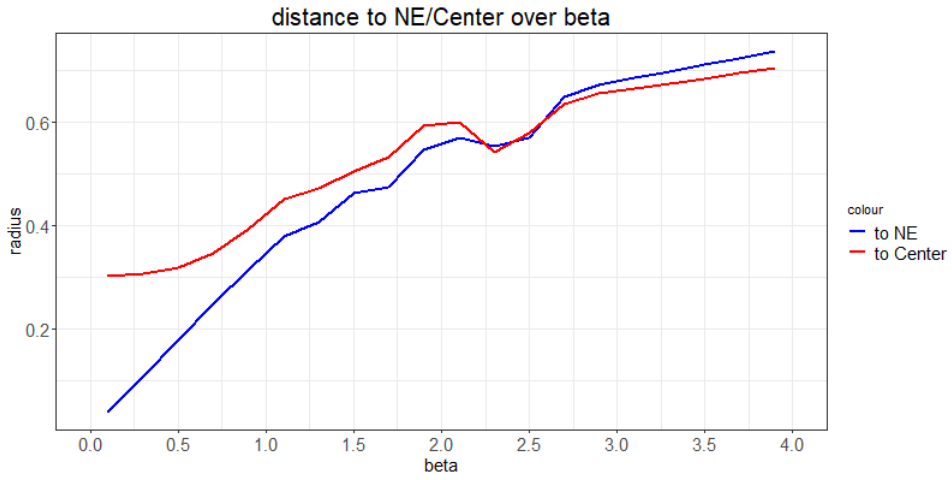

Figure 12: Radius of the limit cycle over β_1 with 20 simulations (1500 iterations in each simulation, data is averaged by the last 500 iterations) for each β_1.

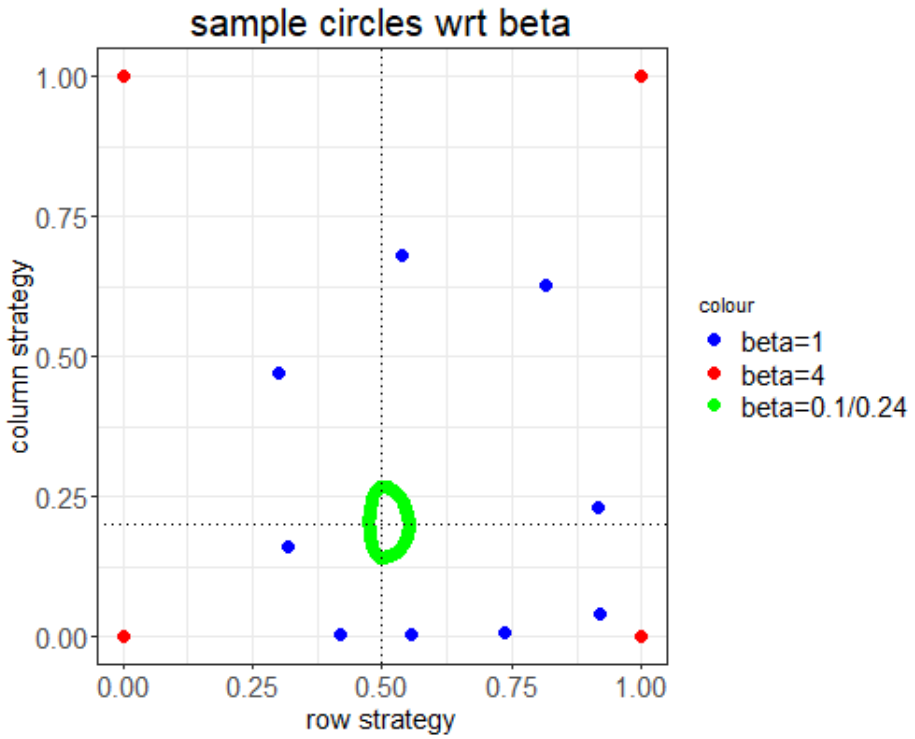

Figure 13: Sample dotted limit cycles with three values of β_1.