Cycles and Instability in a Rock-Paper-Scissors Population Game: A Continuous Time Experiment[*]

Timothy N. Cason[†]   Daniel Friedman[‡]   Ed Hopkins[§]
Purdue University   UC Santa Cruz   University of Edinburgh

June 12, 2013

Abstract

We report laboratory experiments that use new, visually oriented software to explore the dynamics of 3×3 games with intransitive best responses. Each moment, each player is matched against the entire population, here 8 human subjects. A "heat map" offers instantaneous feedback on current profit opportunities. In the continuous slow adjustment treatment, we see distinct cycles in the population mix. The cycle amplitude, frequency and direction are consistent with standard learning models. Cycles are more erratic and higher frequency in the instantaneous adjustment treatment. Control treatments (using simultaneous matching in discrete time) replicate previous results that exhibit weak or no cycles. Average play is approximated fairly well by Nash equilibrium, and an alternative point prediction, "TASP" (Time Average of the Shapley Polygon), captures some regularities that NE misses.

JEL numbers: C72, C73, C92, D83

Keywords: experiments, learning, cycles, mixed equilibrium, continuous time.

1 Introduction

Rock-Paper-Scissors, also known as RoShamBo, ShouShiLing (China) or JanKenPon (Japan) is one of the world's best known games. It may date back to the Han Dynasty 2000 years ago, and in recent years has been featured in international tournaments for computerized agents and humans (Fisher, 2008).

The RPS game is iconic for game theorists, especially evolutionary game theorists, because it provides the simplest example of intransitive dominance: strategy R (Rock) beats strategy S (Scissors) which beats strategy P (Paper), which beats strategy R (Rock). Evolutionary dynamics therefore should be cyclic, possibly stable (and convergent to the mixed Nash equilibrium), or perhaps unstable (and nonconvergent to any mixture). Questions regarding cycles, stable or unstable, recur in more complex theoretical settings, and in applications ranging from mating strategies for male lizards (Sinervo and Lively, 1996) to models of equilibrium price dispersion (Hopkins and Seymour, 2002; Lahkar, 2011).

The present paper is an empirical investigation of behavior in RPS-like games, addressing questions such as: Under what conditions does play converge to the unique interior NE? Or to some other interior profile? Under what conditions do we observe cycles? If cycles persist, does the amplitude converge to a maximal, minimal, or intermediate level? These empirical questions spring from a larger question that motivates evolutionary game theory: To understand strategic interaction, when do we need to go beyond equilibrium theory?

Theoretical analysis of evolutionary dynamics in RPS-like games date back to Zeeman (1980), with extensions to best response dynamics due to Gaunersdorfer and Hofbauer (1995); see Hofbauer and Sigmund (1998) and Sandholm (2010) for reviews. The extensive theoretical literature yields the consensus prediction that RPS-like payoff matrices are of two major types. In one type the interior equilibrium is an Evolutionarily Stable Strategy (ESS) and thus an attractor, so that any cyclic behavior will die out over time and there will be convergence to equilibrium. In the other main type of RPS-like game, the interior equilibrium is unstable, so that cycles will persist and possibly increase in amplitude. Benaïm, Hofbauer and Hopkins (2009) introduce the idea of the TASP ("Time Average of the Shapley Polygon") and show in the unstable case that, although play will continue to cycle, average play can converge to the TASP, which in general is not equal to the Nash equilibrium.

Empirical work is surprisingly sparse; we were able to find only three other human subject experiments investigating RPS-like games. Cason, Friedman and Hopkins (2010) study variations on a 4x4 symmetric matrix game called RPSD, where the 4th strategy, D or Dumb, is never a best response. Using the standard laboratory software zTree, the authors conducted 12 sessions, each with 12 subjects matched in randomly chosen pairs for 80 or more periods. In all treatments the data were quite noisy, but in the most favorable condition (high payoffs and a theoretically unstable matrix), the time-averaged data were slightly better explained by TASP than by Nash equilibrium. The paper reports no evidence of cycles.

Hoffman, Suetens, Nowak and Gneezy (2012) is another zTree study begun about the same time as the present paper. The authors compare behavior with three different symmetric 3x3 matrices of the form $\begin{pmatrix} 0 & -1 & b \\ b & 0 & -1 \\ -1 & b & 0 \end{pmatrix}$, where the treatments are $b = 0.1, 1.0, 3.0$. The unique NE=(1,1,1)/3 is an ESS (hence in theory dynamically stable) when $b = 3$, but not in the other two treatments. The authors report 30 sessions each with 8 human subjects matched simultaneously with all others (mean-matching) for 100 periods. They find that time average play is well approximated by NE, and that the mean distance from NE is similar to that of binomial sampling error, except in the $b = 0.1$ (unstable) treatment, when the mean distance is larger. This paper also reports no evidence of cycles.

Earlier, using fixed pair matches (more relevant to repeated games than evolutionary games), Engle-Warnick and Hopkins (2006) compared play of a bimatrix version of RPS to play in a (rescaled) zero sum game. The mixed equilibrium in the former game is predicted to be unstable under learning, while the mixed equilibrium in the second is predicted to be stable. There is evidence for cyclical play in both games, with some evidence of stronger cycles in the unstable RPS game.

Section 2 reviews relevant theory and distills three testable hypotheses. Section 3 then lays out our experimental design. The main innovations are (a) new visually-oriented software called ConG, which enables players to choose mixed as well as pure strategies and to adjust them in essentially continuous time, and (b) 3x3 payoff bimatrices where the NE is distinct from the centroid (1,1,1)/3. As in previous studies, we compare matrices that are theoretically stable to those that are theoretically unstable, and in the latter case we can distinguish TASP from NE as well as from the centroid. We also compare (virtually) instantaneous adjustment to continuous but gradual adjustment ("Slow"), and to the more familiar synchronized simultaneous adjustment in discrete time ("Discrete").

Section 4 reports the results. After presenting graphs of average play over time in sample periods and some summary statistics, it tests the three hypotheses. All three enjoy considerable, but far from perfect, support. We introduce a new empirical measure called the Cycle Rotation Index. Among other things, we find that cycles persist in the continuous time conditions in both the stable and unstable games, but that cycle amplitudes are consistently larger in the unstable games. In terms of predicting time average play, Nash equilibrium is better than Centroid, and when it differs from the NE, the TASP is better yet.

A concluding discussion is followed by an Appendix that collects mathematical details and sketches a possible explanation for the persistence of observed cycles in stable games. A second Appendix, available online, reproduces instructions to subjects.

2   Some Theory

The games that were used in the experiments are, first, a game we call U_a

$$U_a = \begin{array}{|c|c|c|c|} \hline & \text{R} & \text{P} & \text{S} \\ \hline \text{Rock} & 60,\ 60 & 0,\ 72 & 66,\ 0 \\ \hline \text{Paper} & 72,\ 0 & 60,\ 60 & 30,\ 72 \\ \hline \text{Scissors} & 0,\ 66 & 72,\ 30 & 60,\ 60 \\ \hline \end{array} \tag{1}$$

where U is for unstable because, as we will show, several forms of learning will not converge in this game. The subscript a distinguishes it from U_b that follows. Second, we have the stable RPS game,

$$S = \begin{array}{|c|c|c|c|} \hline & \text{R} & \text{P} & \text{S} \\ \hline \text{Rock} & 36,\ 36 & 24,\ 96 & 66,\ 24 \\ \hline \text{Paper} & 96,\ 24 & 36,\ 36 & 30,\ 96 \\ \hline \text{Scissors} & 24,\ 66 & 96,\ 30 & 36,\ 36 \\ \hline \end{array} \tag{2}$$

Finally, we have a second unstable game U_b

$$U_b = \begin{array}{|c|c|c|c|} \hline & \text{R} & \text{P} & \text{S} \\ \hline \text{Rock} & 60,\ 60 & 72,0 & 30,\ 72 \\ \hline \text{Scissors} & 0,\ 72 & 60,\ 60 & 66,\ 0 \\ \hline \text{Paper} & 72,\ 30 & 0,\ 66 & 60,\ 60 \\ \hline \end{array} \tag{3}$$

Notice that in U_b the best response cycle is reversed so that it is a RSP game rather than RPS. (Of course, subjects in the experiment observed only the neutral strategy labels A, B and C in all games.)

All these games have the same unique Nash equilibrium which is mixed with probabilities (0.25, 0.25, 0.5). The equilibrium payoff is 48 in all cases.

While these games are identical in their equilibrium predictions, they differ quite substantially in terms of predicted learning behavior. Consider a population of players who play this game amongst themselves, either via repeated random matching or (as in our experiment, to reduce sampling variance) via matching against the average mixed strategy of all other players. Suppose that each player chooses a target mixed strategy that is (close to) a best response to the current strategies of their opponents, and that (as in our experiment) the actual mixture employed moves smoothly towards the chosen target. Then we would expect that the population average mixed strategy x would move according to continuous time best response (BR) dynamics, which assumes that the population average strategy moves smoothly in the direction of the best reply to itself. That is, formally,

$$\dot{x} \in b(x) - x \tag{4}$$

where $b(\cdot)$ is the best response correspondence. Equation (4) uses "$\in$" because the best response correspondence can be multivalued. In our games the correspondence is single

valued almost everywhere so the equation can be read as an ordinary differential equation with "∈" replaced by "=". Note also that equation (4) applies strictly only to an infinite population but can be used as an approximation in finite populations.[1]

Because of the cyclical nature of the best response structure of RPS games (Rock is beaten by Paper which is beaten by Scissors which is beaten by Rock), if the evolution of play can be approximated by the best response dynamics, then there will be cycles in play. The question is whether these cycles converge or diverge.

It is easy to show that in the game S, under the best response dynamics, the average strategy would converge to the Nash equilibrium. This is because the mixed equilibrium in S is an evolutionarily stable strategy or ESS. In the games U_a and U_b, however, there will be divergence from equilibrium and play will approach a limit cycle.[2] For example, the case of U_a is illustrated in Figure 1, with the interior triangle being the attracting cycle. This cycle was named a Shapley triangle or polygon by Gaunersdorfer and Hofbauer (1995) after the work of Shapley (1964) who was the first to produce an example of non-convergence of learning in games.

More recently, Benaïm, Hofbauer and Hopkins (BHH) (2009) observe the following.[3] If play follows the BR dynamics then, in the unstable game, play will converge to the Shapley triangle; furthermore, the time average of play will converge to a point that they name the TASP (Time Average of the Shapley Polygon), denoted "T" on Figure 1. It is clearly distinct from the Nash equilibrium of the game, denoted "N" in Figure 1.

These results can be stated formally in the following proposition. The proof can be found in the Appendix.

Proposition 1 *(a) The Nash equilibrium $x^* = (0.25, 0.25, 0.5)$ of the game U_a is unstable under the best response dynamics (4). Further, there is an attracting limit cycle, the Shapley triangle, with vertices, $A_1 = (0.694, 0.028, 0.278)$, $A_2 = (0.156, 0.781, 0.063)$ and $A_3 = (0.018, 0.089, 0.893)$ and time average, the TASP, of $\tilde{x} \approx (0.24, 0.31, 0.45)$. Average payoffs on this cycle are approximately 51.1.*
(b) The Nash equilibrium $x^* = (0.25, 0.25, 0.5)$ of the game S is a global attractor for the best response dynamics (4).

---

[1] A formal model of how best response-like dynamics can arise in similar situation, though again with a large population, can be found in Hofbauer and Sandholm (2007).

[2] Intuitively, the instability arises in U_a and U_b because the normalized gain from winning (which ranges from 6 to 12) is much smaller than the absolute normalized loss from losing (which ranges from -30 to -60). By contrast, in the stable game S the normalized gain from winning (30 to 60) is much larger than the absolute normalized loss from losing (-6 to -12). In other words, in the unstable games draws are almost as good as wins, which pushes learning dynamics towards the corners of the simplex (see Figure 1 below) where draws are more frequent. In the stable game draws are much worse than wins and only a little better than losses, pushing the dynamics away from the corners and decreasing the amplitude of the cycles.

[3] Other recent theoretical results on cycles in RPS-like games can be found in Lahkar (2011), Hahn (2012) and Hommes and Ochea (2012).
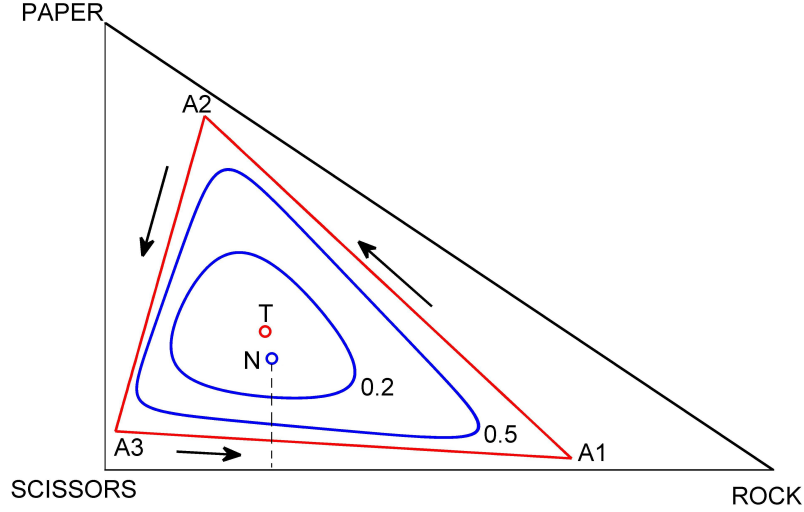
Figure 1: The Shapley triangle $A_1A_2A_3$ for game U_a with the TASP (T) and the Nash equilibrium (N). Also illustrated are orbits for the perturbed best response dynamics for precision parameter values 0.2 and 0.5.

(c) The Nash equilibrium $x^* = (0.25, 0.25, 0.5)$ of the game U_b is unstable under the best response dynamics (4). Further, there is an attracting limit cycle, the Shapley triangle, with vertices, $A_1 = (0.028, 0.694, 0.278)$, $A_2 = (0.781, 0.156, 0.063)$ and $A_3 = (0.089, 0.018, 0.893)$ and time average, the TASP, of $\tilde{x} \approx (0.31, 0.24, 0.45)$. Average payoffs on this cycle are approximately 51.1.

These theoretical arguments lead to the following testable predictions. Note that Hypothesis 2 competes with Hypothesis 1, while Hypothesis 3(a) elaborates on Hypothesis 1 and Hypotheses 3(b,c,d) elaborate on Hypothesis 2.

Testable Hypotheses

1. **Nash Equilibrium (NE):** average play will be at the NE $(0.25, 0.25, 0.5)$ and average payoff will be 48 in all treatments.

2. **TASP:**

   (a) The population average mixed strategy further averaged over time will be closer to the TASP than to the NE in U_a and U_b.

   (b) Average payoffs will be higher in U_a and U_b than in S.

3. **BR Dynamics:**

5

(a) In S, there will be counter-clockwise cycles that diminish in amplitude over time with ultimate convergence to NE.

(b) In U_a, there will be persistent counter-clockwise cycles that approach the Shapley triangle limit cycle.

(c) In U_b, there will be persistent clockwise cycles that approach the Shapley triangle limit cycle.

(d) Thus the average distance from NE will be consistently higher in U_a and U_b than in S.

After proofs of Propositions, the Appendix offers a theoretical analysis of robustness. Specifically, if subjects choose best responses imprecisely, then average frequencies will evolve according to the perturbed best response (PBR) dynamics rather than the strict BR dynamics. In games U_a and U_b, the PBR dynamics also give rise to cycles, two of which for U_a are illustrated in Figure 1 for specified levels of the PBR precision parameter.

A second phase of the robustness analysis focuses on discrete time settings, which constitute some of our experimental treatments. The Appendix shows that the above theoretical results still apply provided strategy adjustment step size is sufficiently small. However, if the step size exceeds some critical level, then cycles may not die out even for the stable payoff matrix. Factors associated with larger step sizes are noted in the next section.

3   Laboratory Procedures

Figure 2 displays an example of the subjects' decision screen during an experimental session. The upper left corner indicates the game payoff matrix B, and subjects choose actions by clicking on locations on the "heat map" triangle in the lower left. They can choose a pure action by clicking on a vertex, and can choose any desired mixture by clicking an interior point. The thermometer to the right of the triangle shows how heat map colors correspond to current payoff flow rates, given the current average mixture $x(t)$ in the population. This hugely reduces the subjects' cognitive load, since otherwise they would continually have to approximate, for each mixture y they might choose, the matrix product $y \cdot Bx(t)$ that gives the payoff flow.

The upper right side of the screen presents in real time the dynamic time path of strategies selected by the subject and the population average. The lower right panel shows the payoff flow received by the subject and the population average; the gray area represents the player's accumulated earnings so far in the current period.

Periods lasted 180 seconds each. Each session began with one unpaid practice period, providing subjects with an opportunity to familiarize themselves with the interface and display. The written instructions that were read aloud to subjects before this practice period
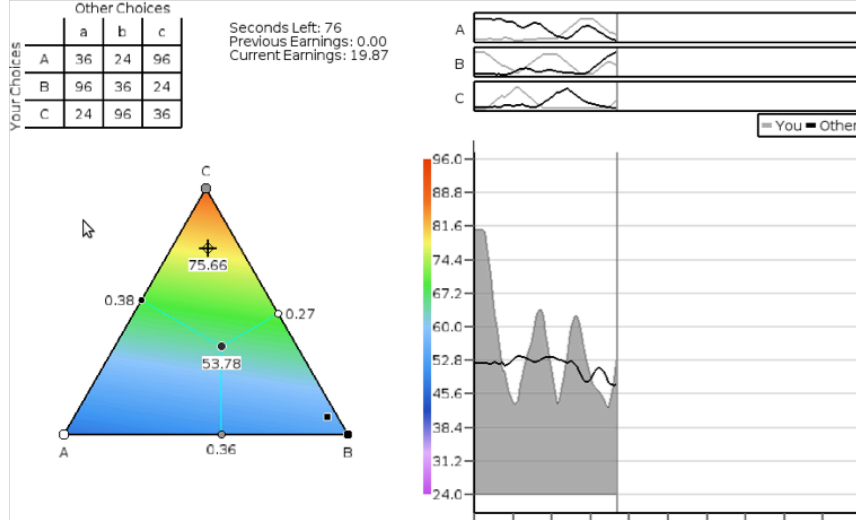
6

Figure 2: ConG Software: CS treatment (10 sec transit)

are included in the online supplementary material. Debriefings after the session was over, as well as experimenter impressions during the session, indicate that most subjects quickly became comfortable with the task, which they regarded as an enjoyable and not especially challenging video game.

Subjects participated in groups of 8 and played the three game matrices U_a, S and U_b in treatment "blocks" of 5 periods each. Treatments were held constant within blocks, while between blocks we switched the game matrix and changed the action set.

We used four different action sets: Continuous Instant, Continuous Slow, Discrete Pure and Discrete Mixed. In the Continuous conditions, subjects click mixture targets and receive payoff and population strategy updates in essentially continuous time. In the Instant case the chosen mixture adjusted to their target click instantaneously; more precisely, lags are typically less than 50 milliseconds, hence imperceptible. In the Slow case the actual mixture moves at a constant rate towards the target click; the rate is such that it would take 10 seconds to move from one vertex to another.

In the Discrete conditions the 180-second period was subdivided into 20 subperiods of 9 seconds each, and subjects received payoff and population strategy updates at the end of each subperiod. In the Pure case subjects were restricted to specify a pure strategy in each subperiod, and in the Mixed case subjects could click any mixture or pure strategy in each subperiod. (If the subject clicked several points during the subperiod, only the last click counted.) Each subperiod after the first, the heat map displayed the potential payoffs given the population mixture chosen in the previous subperiod, and therefore the heat map remained static over the 9-second interval. In the Continuous conditions the heat map (and the displays on the right side of the screen) updated every 200 milliseconds to reflect the current population mixture.

7

The original motivation for these different action treatments is to facilitate comparison to earlier work that uses only discrete time and pure strategies. Our experiment holds constant one novelty (the visual interface featuring a simplex heat map) while varying the other two novelties, the discrete (pure) vs continuous (mixed) strategy set, and the discrete time vs continuous strategy adjustment. The effect of continuous time is further explored by running the different adjustment speeds, instantaneous and slow.

Fortuitously, the four action sets also shed light on a neglected theoretical issue raised in the Appendix – the impact of effective step size in discrete time adjustment models. Our discrete time treatments allow all subjects to change strategies simultaneously and thus adjustment is rapid relative to the time period (the 9 second interval). Allowing only pure strategies amplifies the impact of simultaneity. By contrast, in the continuous treatments, simultaneous adjustment by all subjects is highly unlikely in a time frame of 50 - 200 milliseconds, and such asynchronous adjustment is slowed further in the continuous slow adjustment treatment. We therefore expect effective step size in our data to be ordered by treatment, from smallest to largest, as follows: continuous slow, continuous instant, discrete mixed, discrete pure.

Table 1: Balanced Incomplete Block Design

	Block 1	Block 2	Block 3	Block 4	Block 5
Sess D1	U_a-DM	S-CI	U_a-DP	S-DM	U_b-CS
Sess D2	U_b-CS	U_a-CS	S-CS	U_a-CI	S-DP
Sess D3	S-CS	U_a-DM	S-CI	U_b-CS	S-DM
Sess D4	U_a-CI	S-DM	U_a-DM	S-CS	U_a-CI
Sess D5	S-DP	U_b-CS	U_a-DP	S-CI	U_a-CS
Sess D6	U_a-CS	S-DP	U_a-CI	S-DM	U_a-DP
Sess D7	S-CI	U_a-CS	U_b-CS	S-CS	U_a-DM
Sess D8	U_a-DP	S-DM	U_a-DM	S-DP	U_a-CI
Sess D9	S-CI	U_a-DP	S-DP	U_a-CS	S-CS
Sess D10	S-DM	U_a-CI	S-CS	U_a-DP	S-CI
Sess D11	U_a-CI	S-DP	U_a-CS	U_b-CS	U_a-DM

Note: Every treatment appears in Blocks 1 and 5, at least 8 out of 9 treatments appear in each of the other Blocks, and no treatment appears more than twice in any Block.

Matrices U_a and S were played in each of the four action sets. A bonus treatment U_b evaluates whether cycle rotation direction varied as predicted by BR dynamics. It was used only in conjunction with the treatment (due to its small effective step size) expected to be the most conducive to clean cycles: Continuous Slow. Thus we have 9 different combinations of matrix and action set, or treatments.

Each of the 11 sessions consisted of 5 blocks of 5 periods, with the treatment sequences

shown in Table 1. The within-subject design was chosen to change the matrix every block and to give 6 independent observations (i.e., from 6 different sessions) of each of the 9 treatments, while balancing treatments across block positions.[4] The within-subject variation in treatments also reduced subject boredom. To avoid confounding the treatment variation with learning, all treatments were included in early, middle and late blocks.

Each session lasted about 100 minutes, including instruction and payment time. No subject participated in more than one session, and all were recruited broadly from the student populations at Purdue University and UC-Santa Cruz. All 25 periods in the 5 blocks were paid periods, and subject earnings averaged approximately $25.

4   Results

We begin with graphs of the population mixtures during some sample periods in the Continuous Slow treatment. The figures show a projection of the mixture triangle into the (x, y) plane, so mixture probability for the third strategy ("Scissors") is suppressed. The vertical axis represents time remaining in the 180 second period, so moving downward in the figure corresponds to moving forward in time. The NE appears as a vertical red line at $(x, y) = (.25, .25)$. Figure 3 shows about a dozen irregular counterclockwise cycles of the population average mix around the NE in a sample period using the stable S matrix. Many of the cycles here have amplitude less than 0.1, but a few of them reach a distance of 0.2 or more from the NE.

Figure 4 shows ten counterclockwise cycles around the NE for a sample period using the unstable matrix U_a. The first few cycles (at the top) seem centered on the centroid $(x, y) = (0.33, 0.33)$ but last few cycles center closer to the NE. The amplitude is much larger than for the S matrix, and falls only slightly by the end of the period. Figure 5 shows 11 cycles for the reverse unstable matrix U_b. They are similar to those for U_a, with one major exception: as predicted, the cycles are clockwise.

These sample periods are typical in most respects. Time graphs for other periods suggest that cycles in Continuous Slow treatments persist even with the stable matrix S, as well as with (as predicted) the unstable matrices. The cycles typically seem to converge toward an approximate limit cycle, rather than having inward or outward spirals. As we document below, consistent with Hypothesis 3abc, the cycles under Continuous Slow are consistently counterclockwise for S and U_a and clockwise for U_b.

In the Instant treatment, the cycles are much more frequent and jagged, as Figures 6 and 7 illustrate, and also display greater amplitude for the U_a matrix. Note that only the

---

[4]One treatment, U_a-CI, was repeated in the final block of session D4. To preclude questions about the number of independent observations, this repeat block was excluded from the analysis dataset, but results are qualitatively unchanged when it is included.
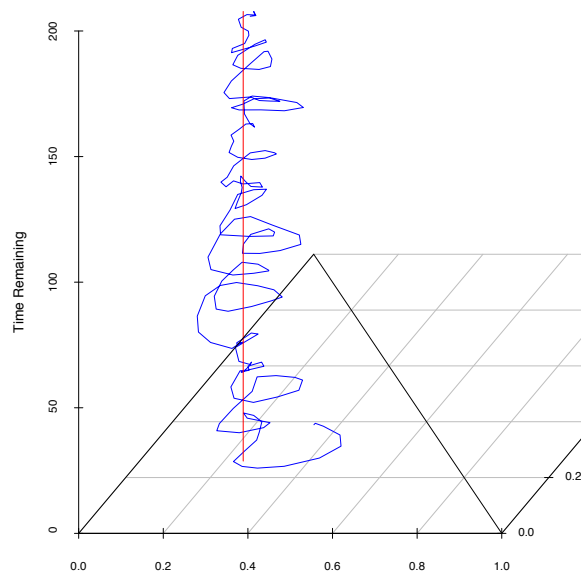
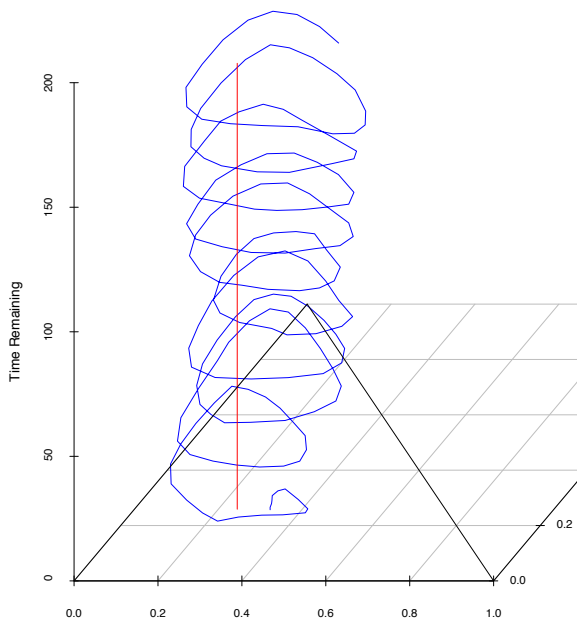Figure 3: Session 10, period 14: S matrix, Continuous-Slow.

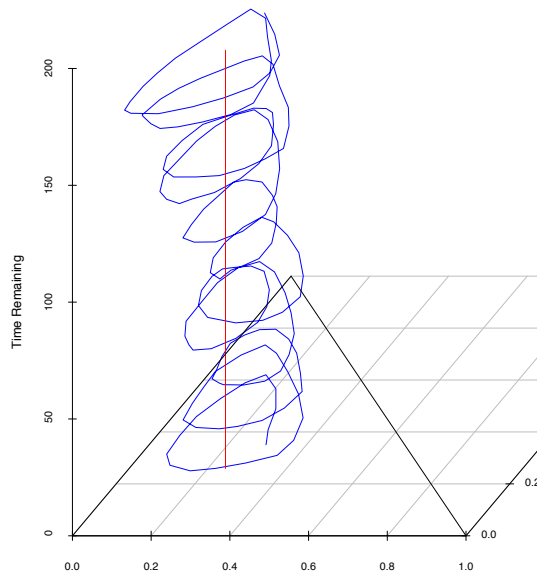Figure 4: Session 2, period 6: U_a matrix, Continuous-Slow.

Figure 5: Session 11, period 20: U_b matrix, Continuous-Slow.
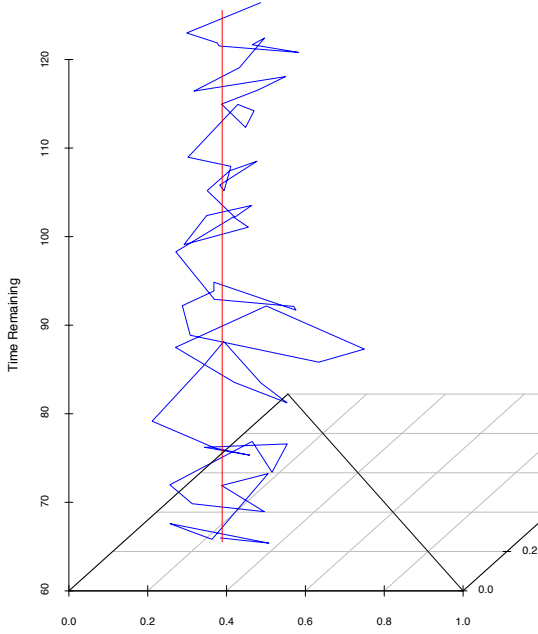
10

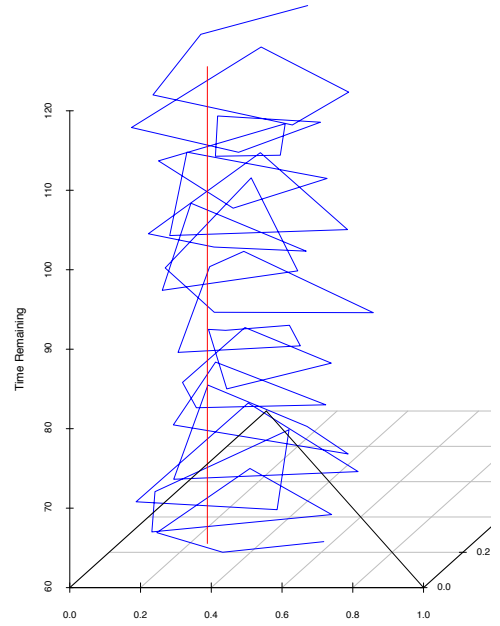Figure 6: Session 3, period 14, Middle 1/3: S matrix, Continuous-Instant.

Figure 7: Session 2, period 17, Middle 1/3: U_a matrix, Continuous-Instant.

middle 60 seconds are shown, but even so there are more than a dozen cycles. In the Discrete treatments, the path of population mixes by subperiod is not so obviously cyclic; see Figures 8 - 11 for typical examples.

4.1   Convergence across periods

Are there trends from one period to the next? To investigate, we plotted average population mixtures in each of the 5 periods within each block separately for each treatment in each session. The results are displayed in Figures 12 and 13 for four of the nine treatments.

Figure 12 shows that behavior remains quite unsettled in the action set featured in previous investigations — discrete time and pure strategy only — especially for the unstable matrix U_a. For example, the population mean frequency of Scissors in one session bounces from about 0.26 in period 4 to 0.61 in period 5. Although Scissors seems the most frequent strategy overall, and Rock the least, this is reversed in some period averages.

By contrast, Figure 13 shows much more consistent mean strategy choices in the Continuous-Instant treatment, particularly for the Stable matrix. The mean Rock frequency is always at or slightly below the NE value of 0.25, and the Paper frequency at or slightly above. The mean Scissors frequency is below the NE value of 0.50, but by period 5 in the Stable matrix it clusters tightly around that value. The other two action set conditions show behavior more settled than in Figure 12 but less than in Figure 13.
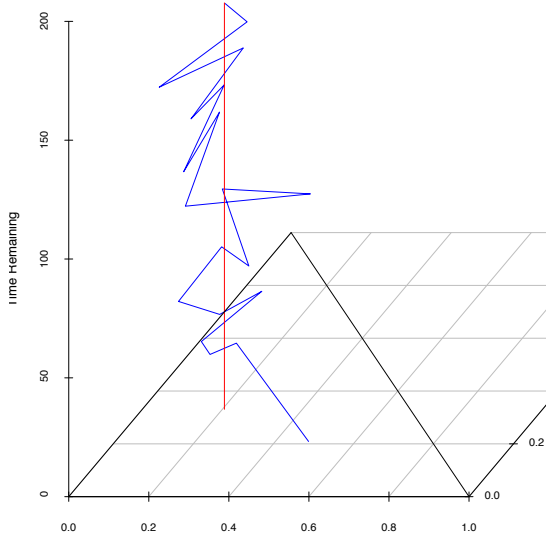
11

Figure 8: Session 3, period 23: S matrix, Discrete-Mixed.
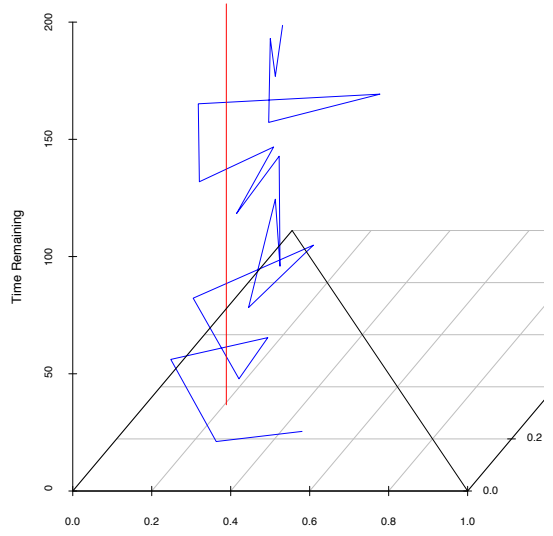

Figure 9: Session 11, period 21: U_a matrix, Discrete-Mixed.

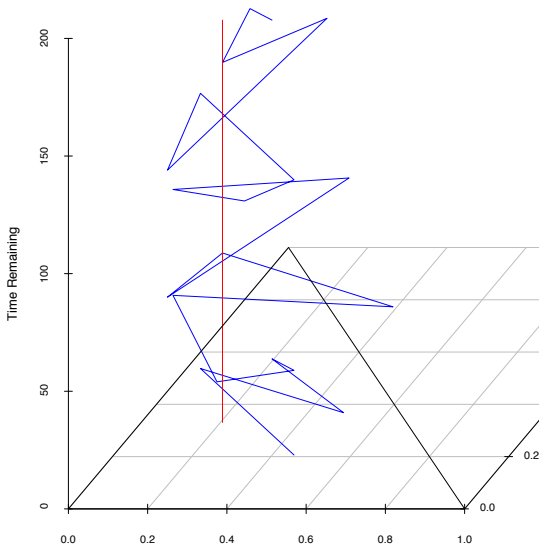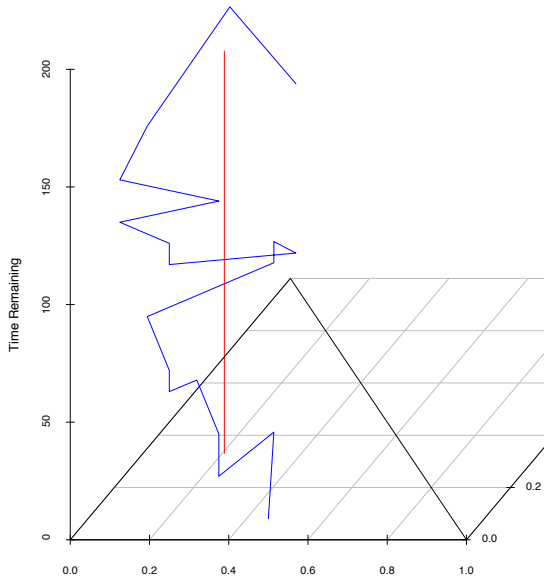Figure 10: Session 2, period 22: S matrix, Discrete-Pure

Figure 11: Session 1, period 11: U_a matrix, Discrete-Pure

12

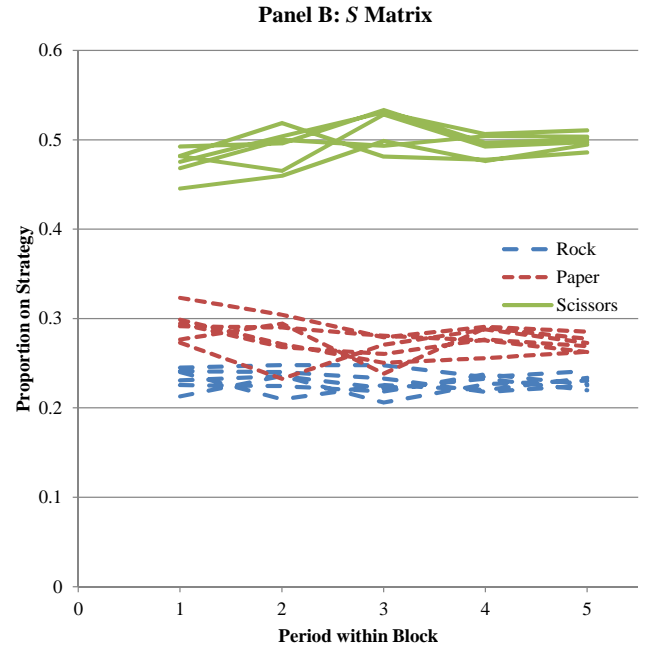**Panel A: U_a Matrix**

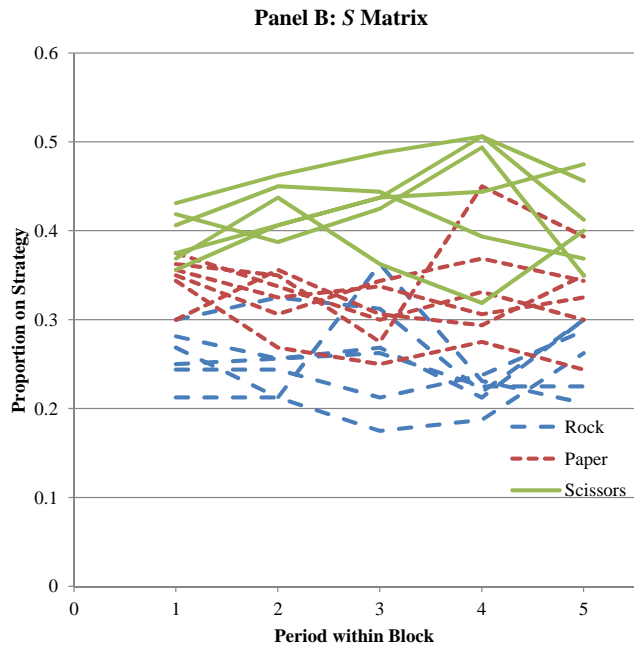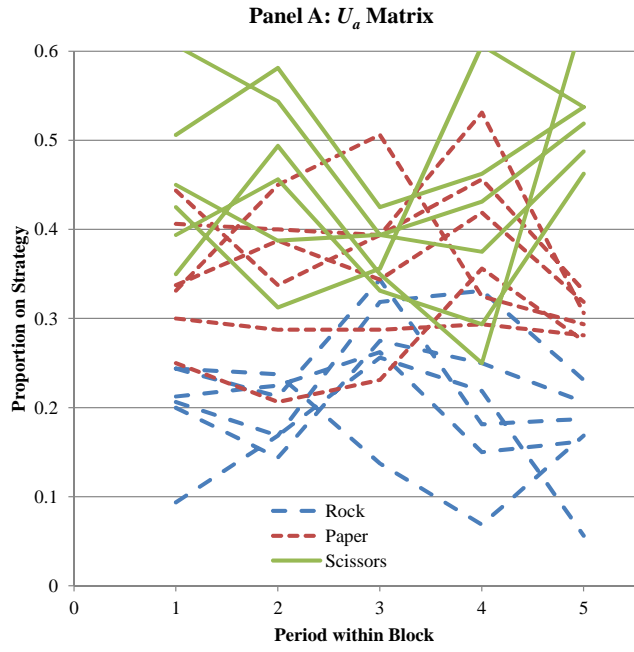**Panel B: S Matrix**
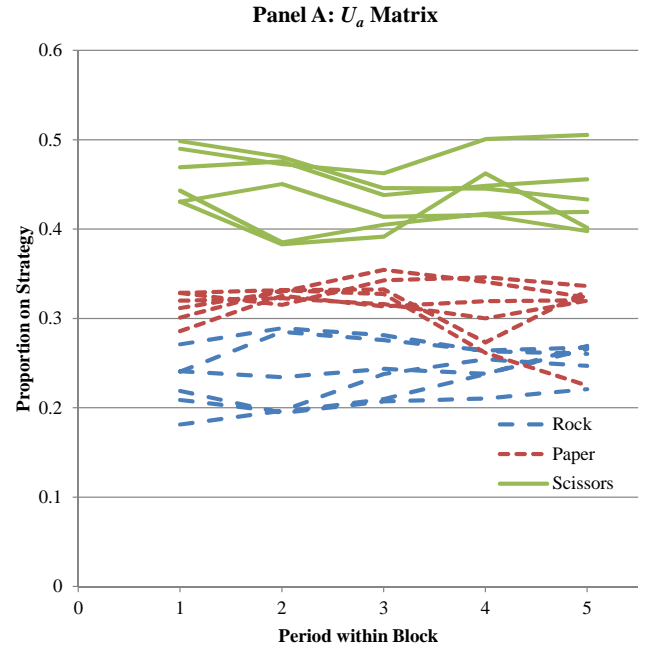
Figure 12: Mean choice by period within block for the Discrete-Pure treatment.

Panel A: U_a Matrix

Panel B: S Matrix

Figure 13: Mean choice by period within block for the Continuous-Instant treatment.

13

4.2   Hypothesis Tests

Table 2 displays predicted (top 3 rows) and actual (remaining 9 rows) overall mean frequency of each strategy and average payoffs. The Wald tests are based on regressions of the observed relative frequency of each strategy choice in each treatment and session block on dummy variables for each strategy. They test the joint hypothesis that estimates correspond to the theoretical values shown in the top of the table. Superscripts in the rightmost column refer to nonparametric Wilcoxon tests for average payoffs that conservatively treat each session as a single independent observation. We draw the following conclusions regarding the three testable Hypotheses given at the end of Section 2.

Result 1: In the Stable game, Nash Equilibrium is better than Centroid in predicting time average strategy frequencies, although the data reject both Nash and Centroid. Average payoff is significantly lower or not significantly different than in Nash Equilibrium. Thus Hypothesis 1 finds mixed support.

Evidence: Table 2 indicates that the data reject the Nash Equilibrium at the 5 percent significance level in all 4 action set conditions for the Stable game. The centroid is also rejected in all 4 cases for the Stable game, but 10 of the 12 average strategies are closer to NE than to the Centroid $(1, 1, 1)/3$. (The two exceptions are for Paper in the two Discrete conditions.) In each of the four treatments, the average payoff is at least 1.3 below the Centroid payoff of 49.33. but it is always within 0.50 of the Nash prediction of 48. The N superscripts indicate that the data reject the null hypothesis of Nash equilibrium payoffs (in favor of a mean payoff *lower* than in NE) in two of these four treatments.

Result 2: In the Unstable game, TASP is better than Nash Equilibrium (and *a fortiori* better than Centroid) in predicting time average strategy frequencies. The average payoff is always significantly higher than the Nash Equilibrium prediction and usually closer to the TASP prediction, albeit significantly lower than the TASP prediction in most cases. Moreover, consistent with TASP, payoffs are significantly higher in the Unstable game than the Stable game. Thus Hypothesis 2 also finds mixed support.

Evidence: The Wald tests in Table 2 indicate rejection of the NE null hypothesis for the strategy frequencies of the Unstable game in all 5 cases, and in all cases the deviations are in the direction of TASP. Rejection the TASP predictions occurs in only 2 of the 5 cases. The average payoff always lies between the NE and TASP predictions, but is closer to TASP in 4 of 5 cases. In all 5 cases the Nash prediction of 48 is rejected; in one case the TASP prediction is not rejected. In all 4 pairwise comparisons, average payoffs are significantly higher (at the 1% level in Mann-Whitney tests) in the Unstable than the Stable game. This is as predicted by TASP, while NE predicts no difference and Centroid predicts a difference in the wrong direction. Finally, TASP tracks the observed asymmetry between U_a and U_b in Rock and Paper time averages, while NE and Centroid predict no asymmetry. The observed asymmetry (for their shared Continuous-Slow action set) is significant according to a Mann-Whitney test.

Table 2: Time average behavior

Prediction//Treatment	Rock	Paper	Scissors			Payoff
Nash Equilibrium	0.25	0.25	0.50	Wald test	Wald test	48
TASP (U_a)	0.242	0.31	0.449	rejects	rejects	51.1
TASP (U_b)	0.31	0.242	0.449	Nash?	TASP?	51.1
S Continuous-Instant	0.226	0.269	0.504	Yes	–	47.59^N
S Continuous-Slow	0.236	0.265	0.500	Yes	–	48.03
S Discrete-Mixed	0.242	0.294	0.464	Yes	–	47.95
S Discrete-Pure	0.247	0.320	0.433	Yes	–	47.57^N
U_a Continuous-Instant	0.247	0.318	0.435	Yes	No	49.82^{NT}
U_a Continuous-Slow	0.228	0.281	0.491	Yes	Yes	49.08^{NT}
U_a Discrete-Mixed	0.225	0.342	0.433	Yes	Yes	49.70^{NT}
U_a Discrete-Pure	0.205	0.337	0.458	Yes	No	50.71^N
U_b Continuous-Slow	0.303	0.240	0.457	Yes	No	48.81^{NT}

Notes: Wald tests of the joint hypothesis that three coefficients on dummy variables for the three strategies correspond to indicated theoretical predictions (regression standard errors estimated with clustering on sessions), all 5% 2-tailed tests. Superscript N denotes average payoff significantly different from Nash (5% 2-tailed Wilcoxon test), and T denotes significantly different from TASP (5% 2-tailed Wilcoxon test, relevant only for U_a and U_b).

Result 3: Cycles are clockwise for the Unstable matrix U_b, and counter-clockwise for the other matrices, Stable S and Unstable U_a. Cycle rotations are also more regular for the Unstable matrix than the Stable matrix. These patterns are consistent with Hypothesis 3abc.

Evidence: Table 3 reports a new statistic called the cycle rotation index to document clockwise or counter-clockwise rotation of mean strategy choices. To calculate this index, we first constructed a line segment (in the projection of the simplex into the two-dimensional plane) between the Nash equilibrium and the simplex edge, illustrated in Figure 1 as the vertical dashed line extending below the point N. This line segment $\mathcal{S}$ is the set of all mixtures in the simplex in which the Rock frequency is at its Nash equilibrium level of 0.25 and the Paper frequency is below its Nash equilibrium level of 0.25. It serves as a Poincare section (see, for example, Chapter 17 of Hofbauer and Sigmund, 1988) or "tripwire" for counting cycles.

For each period in our data set, we counted how many times (CCT, counter-clockwise transits) the population mixture crosses $\mathcal{S}$ from left to right and how many times (CT, clockwise transits) from right to left. The average number of transits in each direction are shown in the center columns of Table 3 for each treatment. The cycle rotation index then is defined for each period as $CRI = \frac{CCT-CT}{CCT+CT} \in [-1,1]$. Thus CRI values near 1 (-1) indicate consistently counter-clockwise (clockwise) cycles, and values near 0 indicate no consistent

cycles. The last column of Table 3 reports CRI averaged over all periods in each treatment.

Table 3: Mean Transits and Cycle Rotation Indexes

Game-Condition	Number of Counter-Clockwise Transits	Number of Clockwise Transits	Cycle Rotation Index
S Continuous-Instant	24.1	5.8	0.64*
S Continuous-Slow	9.3	0.9	0.86*
S Discrete-Mixed	2.1	1.3	0.30
S Discrete-Pure	0.5	0.7	-0.04
U_a Continuous-Instant	30.4	1.2	0.92*
U_a Continuous-Slow	8.3	0.0	1.00*
U_a Discrete-Mixed	1.8	0.3	0.78*
U_a Discrete-Pure	0.9	0.2	0.68*
U_b Continuous-Slow	0.3	8.5	-0.94*

Note: * Denotes Index significantly (p-value $< 5\%$) different from 0 according to 2-tailed Wilcoxon test.

The large values of CCT and CT the Continuous-Instant treatments reflect a substantially higher cycle frequency than in the Continuous-Slow treatments. The Discrete treatments have fewer transits, in part because each period has only 19 potential strategy changes, versus 179 potential strategy changes each Continuous period, where the data are sampled at one-second intervals.

The Discrete treatments for the Stable game S do not exhibit clear cyclical behavior, as indicated by CRI's not significantly different from 0. All other conditions exhibit significant cycles, with only the Unstable game U_b displaying clockwise cycles, consistent with Hypothesis 3abc. Although not shown on the table, Mann-Whitney tests with p-values below 0.05 for all four cases confirm that CRI always is larger for the Unstable matrix U_a than for the Stable matrix S.

Result 4: Cycle amplitude, and thus average distance from the Nash Equilibrium, is significantly higher in the Unstable than in the Stable game (support for Hypothesis 3d). Although the cycle amplitudes usually decrease in size across periods for the Stable matrix, they do not converge to Nash equilibrium, contrary to Hypothesis 3a. Cycles also persist and have larger amplitudes for the Unstable matrices, but less than that of the Shapley triangle limit cycle. Thus Hypotheses 3bc find mixed support.

Evidence: Define *cycle amplitude* for a period as the time average over that period of the squared Euclidean distance $A(t) = ||x(t) - x^*||^2 = (x_0(t) - x_0^*)^2 + (x_1(t) - x_1^*)^2 + (x_2(t) - x_2^*)^2$ between Nash equilibrium mix x^* and the instantaneous actual mix $x(t)$. (The average squared deviation from the period average $\bar{x}$ yields similar results.) Figures 14 and 15 display cycle amplitude period by period for each block of the Continuous conditions; each
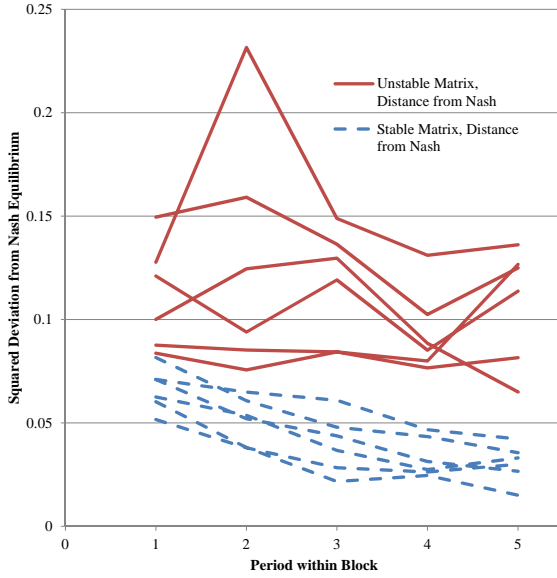
16

Figure 14: Mean squared deviation from NE by block in Continuous-Instant treatments
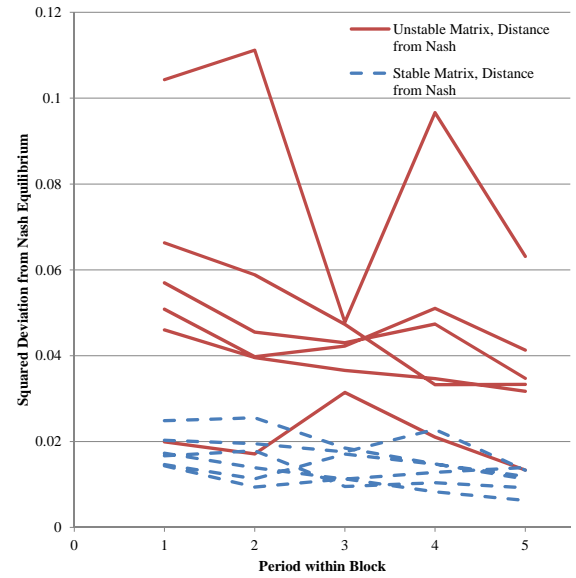
Figure 15: Mean squared deviation from NE by block in Continuous-Slow treatments.

line comes from an independent session. The amplitude declines between the first and last period in 22 out of the 24 blocks; the two exceptions are both in the Unstable - Continuous Instant treatment. But the amplitudes do not decline to zero; even in later periods of the 12 Stable matrix blocks, the squared deviations remain around .01 (i.e., trajectories remain about 0.10 away NE), contrary to Hypothesis 3a.

Table 4 reports average cycle amplitude in each treatment. Numerical calculations indicate that the Shapley triangle limit cycle has amplitude 0.187 for the U_a and U_b matrices. The entries in Table 4 are all smaller than that, and Wilcoxon tests indicate that they are significantly smaller (p-value$< 5\%$) for all treatments. The observed cycle amplitudes are consistent with the perturbed best response dynamics for the low precision parameters illustrated in Figure 1.

The Mann-Whitney p-values shown at the bottom of Table 4 indicate that the amplitude, conditioning on the action set condition, is always significantly greater in the Unstable than the Stable game. The amplitude is not significantly different between the two Unstable games for their shared Continuous-Slow action set, and no difference is consistent with our analysis of BR dynamics.

One loose end remains. What about cycle frequency, as opposed to cycle amplitude? The last piece of evidence under Result 3 noted that the Unstable matrices have significantly larger cycle rotation indexes and thus produced more consistent rotation direction than their Stable counterparts. However the CCT and CT entries in Table 3 suggest no clear ordering on the number of cycles per period, nor do eyeball impressions of individual period graphs.

17

Table 4: Mean Squared Deviation from Nash Equilibrium

	Continuous-Slow	Continuous-Instant	Discrete-Mixed	Discrete-Pure
Stable S	0.014	0.039	0.048	0.093
Unstable U_a	0.044	0.112	0.089	0.129
Unstable U_b	0.048			
p-value for M-W: S vs. U_a	0.004	0.004	0.004	0.037
p-value for M-W: U_a vs. U_b	0.200			

Notes: Excludes first period of each 5-period block to reduce impact of hysteresis. M-W is a 2-tailed Mann-Whitney test comparing MSDs by period across the given treatments.

We estimated cycle frequencies each period using standard frequency domain techniques, employing the cumulative spectral distribution function to identify the most significant cycle frequencies for the strategy and payoff time series.[5] Stronger cycles are evident in the continuous treatments, but overall the frequencies are estimated with substantial noise. Nevertheless, tests show significantly higher frequencies for the U_a-CI treatment than for U_a-CS, which comes as no surprise given the time series graphs and the CCT and CT counts noted earlier. More importantly for present purposes, we find no significant differences between S and U_a (or U_b) in any action treatment. This is consistent with the Conjecture stated in Appendix A based on perturbed best response dynamics.

5   Discussion

Evolutionary game theory predicts cyclic behavior in Rock-Paper-Scissors-like population games, but such behavior has not been reported in previous work.[6] In a continuous time laboratory environment we indeed found cycles in population mixed strategies, most spectacularly for the Unstable matrices with Slow adjustment. Moreover, we consistently observed counterclockwise cycles for one Unstable matrix and clockwise cycles for another, just as predicted.

Surprisingly, we also found very persistent cycles for Stable matrices, where the theory predicted damped cycles and convergence to Nash Equilibrium. The theory was partially vindicated in that these cycles had smaller amplitude than those for corresponding Unstable matrices, but the amplitude settled down at a positive value, not at zero.

---

[5]This procedure decomposes the time series into a weighted sum of sinusoidal functions to identify the principal cycle frequencies.

[6]In laboratory studies focusing on convergence to mixed strategy equilibrium, Friedman (1996) and Binmore et al (2001) both report some evidence of transient cyclic behavior in 2-population games. Neither study considered RPS-like games or persistent cycles.

Evolutionary game theory considers several alternative dynamics. In our setting, replicator dynamics predicts that the cycles for the Unstable matrices have maximal amplitude (i.e., converge to the simplex boundary), while best response dynamics predict cycles that converge to the Shapley triangle and therefore have a particular amplitude less than maximal. The amplitude of cycles we observed with Unstable matrices varied by the action set available to each subject (instantaneous versus slow adjustment in continuous time, and pure only versus mixed strategies in discrete time), but it was always less than for the Shapley triangle. The data thus seem more consistent with perturbed best response dynamics with treatment-dependent step size. Indeed, the ordering of treatments by effective step size noted in section 3 corresponds very closely with the "predicted" ordering of cycle amplitudes observed for each payoff matrix.[7]

Classic game theory predicts that, on average, play will approximate Nash equilibrium. Indeed, time averages over our three minute continuous time periods (and over 20 subperiod discrete time periods) fairly closely approximated Nash equilibrium in all treatments. However, for the Unstable matrix, evolutionary game theory provides an alternative prediction of central tendency called the TASP, and it consistently outperformed Nash equilibrium.

Our results, therefore, are quite supportive of evolutionary game theory. EGT offers short run predictions, where classic game theory has little to say, and those predictions for the most part explained our data quite well. EGT's long run predictions either agreed with those of classic game theory or else were more accurate in explaining our data. Moreover, the predictions were most accurate in the treatment — continuous slow — that best corresponds to the assumptions of evolutionary game theory.[8]

While seeking answers to old questions, our experiment also raises some new questions. Granted that we observed very nice cyclic behavior, one now might want to know more about the necessary conditions. Does our "heat map" play a crucial role? Or is asynchronous choice in continuous time the key? Do cycles dissipate when subjects must choose simultaneously, and some choose to best respond to the previous population mix while others respond to their ("level k") anticipations of others' responses?

We hope that our work inspires studies investigating such questions. The current study already demonstrates that empirically grounded learning models and evolutionary game dynamics can help us grasp "instability," an increasingly important theme for social scientists.

---

[7]The "prediction" noted at the end of the Appendix was obtained from theory developed after the data were collected, hence the quote marks.

[8]As noted by an anonymous referee, our continuous time lab protocols enforce the asynchronous adjustment assumed in standard EGT, while the slow protocol "ensures that aggregate behavior adjusts with inertia, as it would in a large population [assumed in standard EGT], even though the [lab protocol] only used 8 subjects at a time."

Appendix

In this appendix, we state and prove some results on the behavior of the best response (BR) and perturbed best response (PBR) dynamics in the three games U_a, U_b and S.

When one considers stability of mixed equilibria under learning in a single, symmetric population, the most general criterion for stability is negative definiteness of the game matrix, which implies that a mixed equilibrium is an ESS (evolutionarily stable strategy). In contrast, mixed equilibria in positive definite games are unstable. As Gaundersdorfer and Hofbauer (1995) show, for RPS games there is a slightly weaker criterion for the stability/instability of mixed equilibria under the BR dynamics (see below). The RPS games we consider satisfy both criteria.

Proof of Proposition 1: Convergence to the Shapley triangle for the games U_a and U_b follows directly from the results of Gaunersdorfer and Hofbauer (1995). In particular, first one normalizes the payoff matrices by subtracting the diagonal payoff. Thus, for U_a, after subtracting 60, the win payoffs become (12, 12, 6) and the lose payoffs in absolute terms are (60, 60, 30). Thus clearly U_a satisfies the criterion (given in their Theorems 1 and 2) for instability that the product of the lose payoffs is greater than the product of the win payoffs.[9] One can then calculate the Shapley triangle directly from their formula (3.6, p. 286). The TASP is calculated by the procedure given in Benaïm, Hofbauer and Hopkins (2009). The average payoff can similarly be calculated. These results are easily extended to U_b.

Turning to S, subtracting 36, we find the win payoffs to be (60, 60, 30) and the lose payoffs are (12, 12, 6). Thus, clearly the Nash equilibrium is globally stable because it satisfies Gaundersdorfer and Hofbauer's condition that the product of the win payoffs are greater than the product of the lose payoffs. ∎

We note that these results are largely robust to mistakes, particularly the possibility that the choice of best response is not completely accurate. Suppose that subjects choose a mixed strategy that is only an approximate best response to the current average mixed strategy. Then we might expect that the population average might evolve according to the perturbed best response (PBR) dynamics

$$\dot{x} = a(\psi(x) - x) \tag{5}$$

where the function $\psi(\cdot)$ is a perturbed choice function such as the logit and $a > 0$ is a velocity term that we use below. These PBR dynamics are discussed in detail in Hopkins (1999, 2002) and Hofbauer and Sandholm (2002). The fixed points of the PBR dynamics are not in general identical to Nash equilibria but instead are Quantal Response (QRE) or Logit (LE) equilibria.

---

[9]Furthermore, U_a and U_B are also positive definite with respect to the set $\mathbb{R}_0^n = \{x \in \mathbb{R}^n : \sum x_i = 0\}$ which is a stronger criterion.

Perturbed best response choice functions such as the logit are typically parameterized with a precision parameter λ, which is the inverse of the amount of noise affecting the individual's choice. In such models, an increase of the precision parameter λ, for learning outcomes are the following. First, it is well known that the stability of mixed equilibria under the PBR dynamics (5) depend upon the level of λ. When λ is very low, agents randomize almost uniformly independently of the payoff structure and a perturbed equilibrium close to the center of the simplex will be a global attractor. This means that even in the unstable games U_a and U_b, the mixed equilibrium will only be unstable under PBR dynamics if λ is sufficiently large. For the games U_a and U_b, it can be calculated that the critical value of λ for the logit version of the dynamics is approximately 0.18. Second, in contrast, in the stable game S, the mixed equilibrium will be stable independent of the value of λ.

Proposition 2 *In U_a and U_b, the perturbed equilibrium (LE) $\hat{x}$ is unstable under the logit form of the perturbed best response dynamics (5) for all $\lambda > \lambda^* \approx 0.18$. Further, at λ^* there is a supercritical Hopf bifurcation, so that for $\lambda \in (\lambda^*, \lambda^* + \epsilon)$ for some $\epsilon > 0$, there is an attracting limit cycle near $\hat{x}$.*

Proof: Instability follows from results of Hopkins (1999). The linearization of the logit PBR dynamics (5) at $\hat{x}$ will be of the form $\lambda R(\hat{x})B - I$ where R is the replicator operator and B is the payoff matrix, either U_a or U_b. Its eigenvalues will therefore be of the form $\lambda k_i - 1$ where the k_i are the eigenvalues of $R(\hat{x})B$. $R(\hat{x})B$ has eigenvalues with real part positive as both U_a and U_b are positive definite. Thus, for λ sufficiently large, all eigenvalues of $\lambda R(\hat{x})B - I$ will have positive real parts. We find the critical value of 0.18 by numerical analysis. The existence of a supercritical Hopf bifurcation for RPS games has been established by Hommes and Ochea (2012). ∎

This result is less complete than Proposition 1, in that it does not give a complete picture of PBR cycles away from equilibrium. Numeric analysis for the logit form of the PBR dynamics suggests that as for the BR dynamics there is a unique attracting limit cycle (for $\lambda > 0.18$). The amplitude of this cycle is increasing in λ and approaches that of the Shapley triangle as λ becomes large. Two sample limit cycles are illustrated in Figure 1.

The game S is negative definite and hence its mixed equilibrium is a global attractor under both the BR and PBR dynamics. This implies it is also an attractor for (stochastic) fictitious play.

Proposition 3 *The perturbed equilibrium (QRE) of the game S is globally asymptotically stable under the perturbed best response dynamics (5) for all $\lambda \geq 0$.*

Proof: It is possible to verify that in the game S is negative definite with respect to the set $\mathbb{R}_0^n = \{x \in \mathbb{R}^n : \sum x_i = 0\}$. See e.g. Hofbauer and Sigmund (1998, p80) for the negative definiteness (equivalently ESS) condition. The result then follows from Hofbauer and Sandholm (2002). ∎

Conjecture 1 *For λ sufficiently large and for x sufficiently close to the perturbed equilibrium $\hat{x}$, cycles of the PBR dynamics, controlling for amplitude, should have the same frequency for all three games U_a, U_b and S.*

The reason behind the conjecture is the following. For λ large, the perturbed equilibrium $\hat{x}$ is close to the NE $x^* = (1/4, 1/4, 1/2)$. It is then possible to approximate the PBR dynamics by a linear system $\dot{x} = R(x^*)Bx$ where, for the logit version of the PBR dynamics, R is the replicator operator (see the proof of Proposition 2 above) and B is the payoff matrix, which could be any of U_a, U_b or S. One can calculate $R(x^*)B$, precisely for all three games and thus derive the eigenvalues for this linear system, which are exactly $6 \pm 9\sqrt{3}i$ for U_a and U_b and $-6 \pm 9\sqrt{3}i$ for S. It is the imaginary part of the eigenvalues that determines the frequency of the solutions of the linearized system $\dot{x} = R(x^*)Bx$, while the exponential of the real part determines the amplitude. The imaginary part is identical for all three games and thus one would expect similar frequencies in cycles. Admittedly, there are two approximations in making this argument. First, this is a linear approximation to the non-linear PBR dynamics and will only be valid close to equilibrium. Second, the linearization should properly be taken at $\hat{x}$ and not at the NE x^*. However, for λ large, the loss of accuracy should not be too great.

Turning now to discrete time learning, such as in our discrete time treatments, suppose over a time period of length Δ a proportion γ of individuals in the population move their choice of strategy x a proportion θ towards the (perturbed) best response to the current population state x_t. The remaining individuals do not change their strategy.[10] Then, the population state will change to

$$x_{t+\Delta} = \gamma\theta\phi(x_t) + (1 - \gamma\theta)x_t \tag{6}$$

So γ and θ are complements in determining the step size of the discrete process. Then if $\gamma\theta$ is proportional to Δ, i.e. if $\gamma\theta = a\Delta$, we have

$$\frac{x_{t+\Delta} - x_t}{\Delta} = a(\phi(x_t) - x_t) \tag{7}$$

So if Δ is small, then the discrete time adjustment process (7) will approximate the PBR dynamics (5).[11] One can also verify that it has the same fixed points. So, if this discrete time process does converge, it will converge to the perturbed equilibrium $\hat{x}$ corresponding to the Nash equilibrium of the appropriate game.

---

[10]This is, in effect, a very simplified version of the model in Hofbauer and Sandholm (2007). They formally show how, in a large population with asynchronous adjustment and perturbed payoffs, the aggregate dynamics can be approximated by the PBR dynamics.

[11]There is one other well-known reason why discrete time learning can be approximated by continuous time dynamics - the long memory assumption associated with classical fictitious play. When beliefs about the current state are determined by the whole history of play, then the step size of adjustment of strategies becomes very small over time. We see this as being less relevant in the current experiment where subjects were clearly informed about the current population state than when information is less complete and subjects might use past experience to form beliefs.
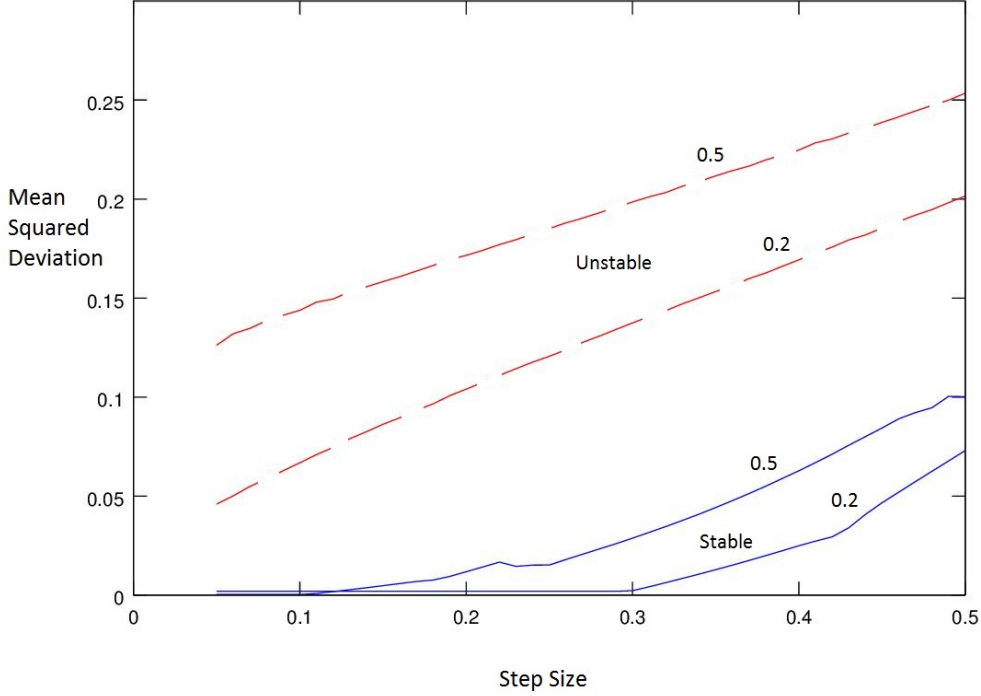
Figure 16: Simulated mean squared distance from Nash equilibrium for the discrete time PBR dynamics in the Stable game S and the unstable game U_a as a function of step size given precision parameter λ values 0.2 and 0.5.

Formally, for stability of a discrete system like (6), all eigenvalues have to be less than one in absolute value. The linearization of the logit version of (6) taken at a perturbed equilibrium $\hat{x}$ is

$$\frac{dx_{t+\Delta}}{dx_t} = \gamma\theta\lambda R(\hat{x})B + (1 - \gamma\theta)I \tag{8}$$

(see the proof of Proposition 2 above for derivation). We can verify directly that the eigenvalues of RB are $-6 \pm 9\sqrt{3}i$ for the stable game S with absolute values much greater than one (and similarly for U_a and U_b). We then have the following standard result.

Proposition 4 *For the game S there is a $\mu \in (0,1)$ such that if $\gamma\theta < \mu$, the perturbed equilibrium $\hat{x}$ is stable under the discrete time PBR dynamics (6). For $\lambda > \lambda^* \approx 0.18$ in the games U_a and U_b the perturbed equilibrium is unstable for any value of $\gamma\theta$ in $(0,1]$.*

Proof: The linearization is a convex combination of two matrices $\lambda R(\hat{x})B$ and I. Thus, its eigenvalues will be a convex combination of the eigenvalues of these two matrices. In the stable case, the eigenvalues of $\lambda R(\hat{x})B$ have negative real part, so their convex combination

23

with the unit eigenvalues of I will have absolute value less than one for $\gamma\theta$ sufficiently small. For $\lambda > \lambda^*$, the real part of the eigenvalues of $\lambda R(\hat{x})B$ for U_a and U_b is greater than 1, so the absolute value of the eigenvalues is definitely greater than one. Consequently, the convex combination of these eigenvalues with the unit eigenvalues of I always have absolute value greater than one. ∎

In effect, instability carries through into discrete time, but stability only carries through if the step size is sufficiently small. However, what if the step size is too big? Is there then no difference between the stable and unstable cases? We turn to numerical simulation and find that, in fact, the difference persists.

Figure 16 shows the simulated long run mean squared deviation from the Nash equilibrium for the discrete time system (6). It graphs for the games S (solid blue) and U_a (dashed red) for λ values 0.2 and 0.5 as marked against a varying step size ($\gamma\theta$) on the horizontal axis. For example, the lowest line is for the stable game S with $\lambda = 0.2$. One can calculate that for the equilibrium to be stable the step size must be below 0.3 and indeed the simulation replicates this with the distance from equilibrium being minuscule for a step size below 0.3. It then steadily rises. However, the distance is always below that of the unstable game. This leads to several important conclusions.

1. If the step size of the discrete time process is not very small, distance from equilibrium will be non-zero even for the stable game and even in the long run.

2. Nonetheless, distance from the equilibrium will be larger in the unstable games than in the stable.

3. Factors (such as the action treatments in the experiment) that increase the step size should be associated with increased observed distance from equilibrium.

References

Benaïm, M., Hofbauer, J., and Hopkins, E. (2009). "Learning in games with unstable equilibria", *Journal of Economic Theory*, **144**, 1694-1709.

Binmore, K., Swierzbinski, J. and Proulx, C. (2001). "Does Min-Max Work? An Experimental Study," *Economic Journal*, **111**, 445-464.

Cason, T., Friedman, D. and Hopkins, E. (2010) "Testing the TASP: an Experimental Investigation of Learning in Games with Unstable Equilibria", *Journal of Economic Theory*, **145**, 2309-2331.

Engle-Warnick, J., Hopkins, E. (2006). "A simple test of learning theory," working paper.

Fisher, Len (2008). *Rock, Paper, Scissors: Game Theory in Everyday Life*. Basic Books.

Friedman, D. (1996) "Equilibrium in Evolutionary Games: Some Experimental Results," *Economic Journal*, **106: 434**, 1-25.

Gaunersdorfer, A., and J. Hofbauer (1995). "Fictitious play, Shapley Polygons, and the Replicator Equation," *Games and Economic Behavior*, **11**, 279-303.

Hahn, Martin (2012). "An Evolutionary Analysis of Varians Model of Sales", *Dynamic Games and Applications*, **2**, 7196.

Hofbauer, J., Sigmund K., (1998) *Evolutionary Games and Population Dynamics.* Cambridge: Cambridge University Press.

Hofbauer, J., W.H. Sandholm, (2002). "On the global convergence of stochastic fictitious play", *Econometrica*, **70**, 2265-2294.

Hofbauer, J., W.H. Sandholm, (2007). "Evolution in games with randomly disturbed payoffs," *Journal of Economic Theory*, **132**, 47-69.

Hoffman, M., Suetens, S., Nowak, M., and Gneezy, U. (2011) "An experimental test of Nash equilibrium versus evolutionary stability", working paper.

Hommes, Cars H. and Marius I. Ochea (2012) "Multiple equilibria and limit cycles in evolutionary games with Logit Dynamics", *Games and Economic Behavior*, **74**, 434-441.

Hopkins, E. (1999). "A note on best response dynamics," *Games and Economic Behavior*, **29**, 138-150.

Hopkins, E. (2002). "Two Competing Models of How People Learn in Games," *Econometrica*, **70**, 2141-2166.

Hopkins, E., Seymour, R. (2002). "The Stability of Price Dispersion under Seller and Consumer Learning", *International Economic Review*, **43**, 1157-1190.

Lahkar, R. (2011). "The dynamic instability of dispersed price equilibria," *Journal of Economic Theory*, **146**, 1796-1827.

Sandholm, William H. (2010). *Population Games and Evolutionary Dynamics*, MIT Press.

Shapley, L. (1964). "Some topics in two person games," in M. Dresher et al. eds., *Advances in Game Theory*, Princeton: Princeton University Press.

Sinervo, B., and Lively, C. (1996). "The rock-paper-scissors game and the evolution of alternative male strategies," *Nature*, **380**, 240-243.

Zeeman, E. C. (1980) "Population dynamics from game theory," *Lecture Notes in Mathematics*, **819**, 471-497.