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

Timothy Cason, Daniel Friedman, Ed Hopkins

Presented by Kevin He

April 21, 2014

Why Rock-Paper-Scissors?

Rock-paper-scissors (RPS) is an ancient game. Historians date RPS to the Chinese Han Dynasty more than 2000 years ago.

Why Rock-Paper-Scissors?

Rock-paper-scissors (RPS) is an ancient game. Historians date RPS to the Chinese Han Dynasty more than 2000 years ago.

Despite its antiquity, a comprehensive review of the theoretical properties of RPS was not available until Seinfeld (1994).

Why Rock-Paper-Scissors? (the serious version)

Rock-Paper-Scissors (RPS) is the simplest game with intransitive dominance: rock beats scissors, scissors beats paper, paper beats rock.

Intuitively, RPS seems the ideal setting for studying cycles in evolutionary dynamics.

Why Rock-Paper-Scissors? (the serious version)

Rock-Paper-Scissors (RPS) is the simplest game with intransitive dominance: rock beats scissors, scissors beats paper, paper beats rock.

Intuitively, RPS seems the ideal setting for studying cycles in evolutionary dynamics.

Questions we may ask about RPS:

- Under what conditions will play converge to unique mixed NE?
- Under what conditions will we observe cycles?

Why Rock-Paper-Scissors? (the serious version)

Rock-Paper-Scissors (RPS) is the simplest game with intransitive dominance: rock beats scissors, scissors beats paper, paper beats rock.

Intuitively, RPS seems the ideal setting for studying cycles in evolutionary dynamics.

Questions we may ask about RPS:

- Under what conditions will play converge to unique mixed NE?
- Under what conditions will we observe cycles?

These questions are well understood in theory. (Will review results later.)

They are less well studied in experiments involving human players.

Recent experiments on repeated RPS

Cason et al. (2010) studied a variation of RPS. They used a 4×4 game, adding a new action D for each player that is never the best response.

Recent experiments on repeated RPS

Cason et al. (2010) studied a variation of RPS. They used a 4×4 game, adding a new action D for each player that is never the best response.

Hoffman et al. (2012) studied a very similar problem to the present paper. They used different RPS payoff matrices that should have different dynamic implications for convergence of play.

Recent experiments on repeated RPS

Cason et al. (2010) studied a variation of RPS. They used a 4×4 game, adding a new action D for each player that is never the best response.

Hoffman et al. (2012) studied a very similar problem to the present paper. They used different RPS payoff matrices that should have different dynamic implications for convergence of play.

Both of these studies:

- used z-Tree, the standard software for social science lab experiments
- (as a result) studied repeated RPS in discrete time
- Found no evidence of cycle, even in games predicted to have cycling under best response dynamic

But continuous time may matter, at least in theory...

Dekel and Scotchmer (1992) give an example of a game that, under discrete-time replicator dynamics (growth rate of a strategy at time $t+1$ depends on how well it does compared to the population average of time t), has a limit point that places positive probability on a strictly dominated action.

- in fact, the game is same kind as the “augmented RPS” in Cason et al. (2010)
- by contrast, replicator dynamics under continuous time cannot place positive probability on strictly dominated action, Samuelson and Zhang (1992)

Experimental innovation: ConG

ConG is a program for running **Con**tinuous-time **G**ames in the lab. It was developed by **L**earning and **E**xperimental **E**conomics **P**roject of **S**anta Cruz (**LEEPS**), with Daniel Friedman as its director.

Innovations over z-Tree:

- "continuous" time games
- allows players to easily specify mixed strategies (receive expected payoff)
- better strategy and payoff visualization for the players

RPS under best response dynamic

Definition

A symmetric game with payoff matrix:

$$\begin{pmatrix} 0 & -\mu_2 & \lambda_3 \\ \lambda_1 & 0 & -\mu_3 \\ -\mu_1 & \lambda_2 & 0 \end{pmatrix},$$

with $\mu_i, \lambda_i > 0$, is called a RPS game.

RPS under best response dynamic

Definition

A symmetric game with payoff matrix:

$$\begin{pmatrix} 0 & -\mu_2 & \lambda_3 \\ \lambda_1 & 0 & -\mu_3 \\ -\mu_1 & \lambda_2 & 0 \end{pmatrix},$$

with $\mu_i, \lambda_i > 0$, is called a RPS game.

Definition

If the evolution of x satisfies $\dot{x} \in BR(x) - x$, then say x follows best response dynamic.

RPS under best response dynamic

Definition

A symmetric game with payoff matrix:

$$\begin{pmatrix} 0 & -\mu_2 & \lambda_3 \\ \lambda_1 & 0 & -\mu_3 \\ -\mu_1 & \lambda_2 & 0 \end{pmatrix},$$

with $\mu_i, \lambda_i > 0$, is called a RPS game.

Definition

If the evolution of x satisfies $\dot{x} \in BR(x) - x$, then say x follows best response dynamic.

Theorem

(Gaunersdorfer and Hofbauer 1995) If $\lambda_1\lambda_2\lambda_3 \geq \mu_1\mu_2\mu_3$, then best response dynamic converges to NE. If $\lambda_1\lambda_2\lambda_3 < \mu_1\mu_2\mu_3$, then the limit of best response dynamic is a triangle ("the Shapley polygon").

RPS under best response dynamic

Definition

The Time-Average of the Shapley Polygon (TASP) of a RPS game is defined as

$$x_{TASP} := \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t) dt$$

where x follows best response dynamic.

RPS under best response dynamic

Definition

The Time-Average of the Shapley Polygon (TASP) of a RPS game is defined as

$$x_{TASP} := \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t) dt$$

where x follows best response dynamic.

In light of previous theorem:

- If RPS game satisfies $\lambda_1 \lambda_2 \lambda_3 \geq \mu_1 \mu_2 \mu_3$ ("stable"), then $x_{TASP} = x_{NE}$.
- If RPS game has $\lambda_1 \lambda_2 \lambda_3 < \mu_1 \mu_2 \mu_3$ ("unstable"), then expect x_{TASP} to fit long run time average behavior better than x_{NE}.

Payoff matrices used in this study

Payoff matrices informed by theory. U_a and U_b are unstable while S is stable under best response dynamic.

	R	P	S
R	60,60	0,72	66,0
P	72,0	60,60	30,72
S	0,66	72,30	60,60

(a) The payoff matrix U_a

	R	S	P
R	60,60	72,0	30,72
S	0,72	60,60	66,0
P	72,30	0,66	60,60

(b) The payoff matrix U_b

	R	P	S
R	36,36	24,96	66,24
P	96,24	36,36	30,96
S	24,66	96,30	36,36

(c) The payoff matrix S

Payoff matrices used in this study

Some additional properties of the payoff matrices:

- In U_b, the strategies are ordered differently than in other two matrices
  - If we relabel the three strategies as ABC for each payoff matrix, then the cycles we observe in ABC space for U_a and for U_b should be in opposite directions
- unique NE is same in each: $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$ (in RPS space)
- NE payoff is 48 in all cases
- time average payoff along the Shapley polygon in U_a, U_b is about 51.1.

Experimental procedure

- Test subjects are students at Purdue University and UC Santa Cruz
- Each period lasts 180 seconds and has 8 subjects. Each player plays against the population average strategy of others.
- Each session starts with 1 unpaid, practice period, followed by 5 blocks of 5 periods each
- Each block fixes one game matrix (U_a, U_b, S) and one "action set", which determines how frequently players can modify their strategy:
  - continuous instant
  - continuous slow (takes 10 seconds to gradually change from one pure strategy to another)
  - discrete pure (20 periods of 9 seconds each)
  - discrete mixed

Experimental procedure

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

Figure : The block treatments in different sessions. Each treatment is played in at least 6 different sessions.

Qualitative evidence for cycles

Authors plot the evolution of population average strategy for several periods.

They do these plots in *xyz* space, where x is probability assigned to first action, y is probability assigned to second action, and z is time remaining until end of the period.

So, projected onto the *xy* plane, the plot lives in the triangle: $\{(x, y)|x, y \geq 0, x + y \leq 1\}$.

This is a little odd – the triangle is isosceles but not equilateral.

Qualitative evidence for cycles

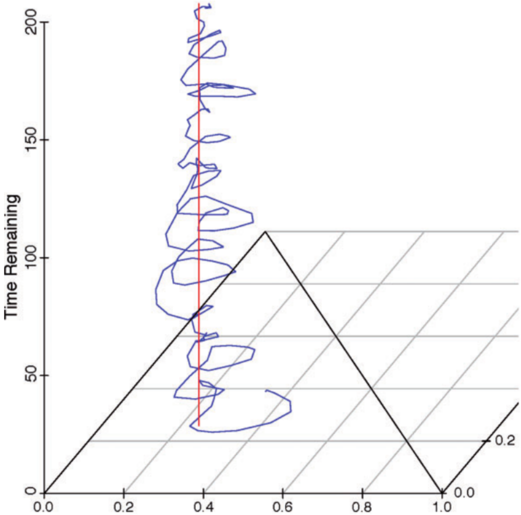

Figure : Population strategy evolution in a S, continuous-slow treatment period.

Qualitative evidence for cycles

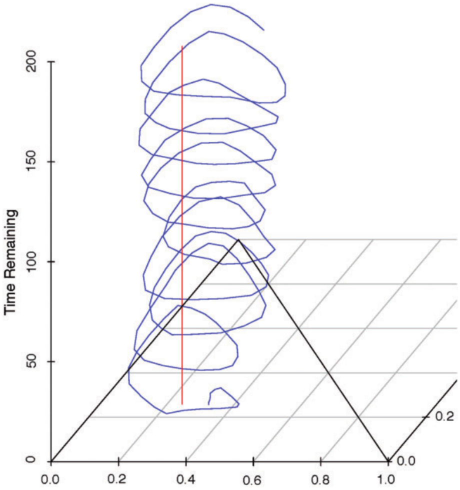

Figure : Population strategy evolution in a U_a, continuous-slow treatment period.

Qualitative evidence for cycles

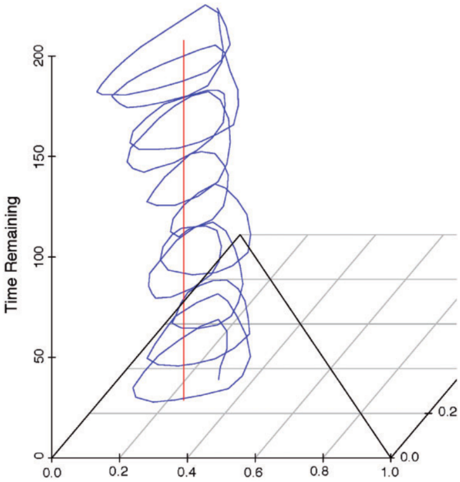

Figure : Population strategy evolution in a U_b, continuous-slow treatment period.

Qualitative evidence for cycles

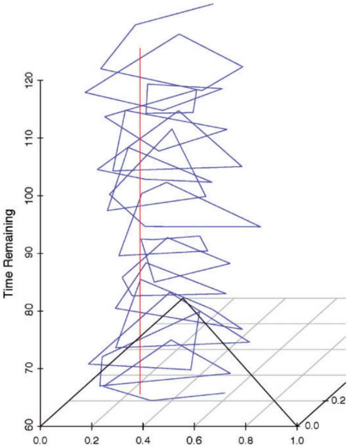

Figure : Population strategy evolution in a U_a, continuous-instant treatment period.

Qualitative evidence for cycles

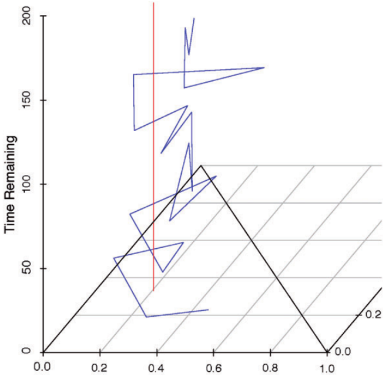

Figure : Population strategy evolution in a U_a, discrete-mixed treatment period.

Qualitative evidence for cycles – comments and takeaway

The following seem clear from these pictures:

- The cycles in U_a, U_b are more regular and have larger amplitude than S
- Cycles run in the expected direction
- Cycles seem to disappear for discrete treatment

Qualitative evidence for cycles – comments and takeaway

The following seem clear from these pictures:

- The cycles in U_a, U_b are more regular and have larger amplitude than S
- Cycles run in the expected direction
- Cycles seem to disappear for discrete treatment

Why cycles in ConG but not in previous z-Tree studies?

- perhaps due to more frequent adjustment. Indeed, cycles less evident in discrete treatment even for ConG
- but perhaps also due to the salience of the heatmap in ConG!
  - encourages playing a myopic best response
  - so, ensures the best response dynamic $\dot{x} = BR(x) - x$ is reproduced in the data

Qualitative evidence for cycles – comments and takeaway

For comparison, here is the z-Tree interface used in Cason et al. (2010)

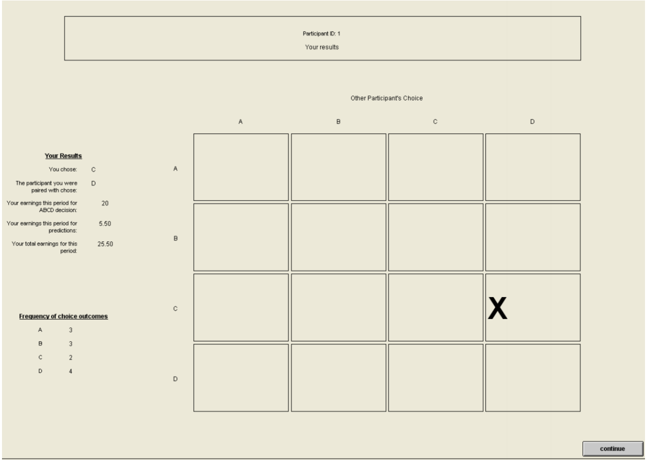

Figure : The feedback screen in Cason et al. (2010)

Qualitative evidence for cycles – comments and takeaway

For comparison, here is the z-Tree interface used in Cason et al. (2010)

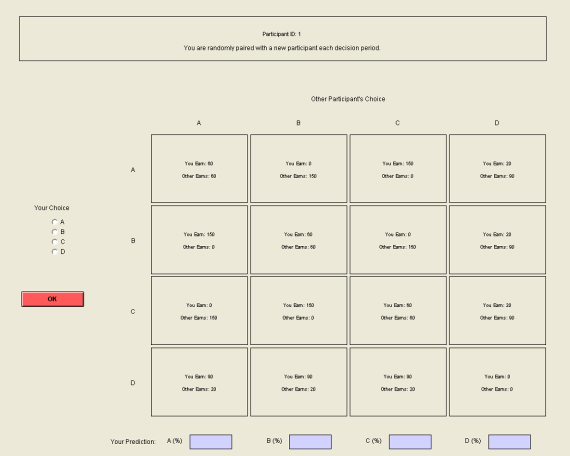

Figure : The strategy choice screen in Cason et al. (2010)

Cycle counting via the Cycle Rotation Index

It's clear that cycles exist under some treatments, just from the pictures.

But for some other treatments, the pictures are fuzzy. For example, hard to say if the plot for U_a under discrete-mixed treatment exhibits cycles or not.

Ideally: a way to describe how many cycles are in each plot.

No off-the-shelf statistic available. Therefore, the authors define their own: the Cycle Rotation Index (CRI).

Cycle counting via the Cycle Rotation Index

Defining CRI

1. Fix a center $(\alpha, \beta) \in \{(x, y) | x, y \geq 0, x + y \leq 1\}$. Authors choose NE as center.
2. Draw a perpendicular line segment from (α, β) to the $x-$axis. Call this line the tripwire.
3. Initialize CT $= 0$, CCT $= 0$
4. Every time the path of strategy evolution crosses the tripwire, update CT++ or CCT++ depending on direction of transit.
5. Finally compute CRI:$= \frac{\text{CCT}-\text{CT}}{\text{CCT}+\text{CT}} \in [-1, 1]$.

Cycle counting via the Cycle Rotation Index

Wang et al. in “A Comment on RPS”

- correct a mistake in this paper’s tripwire transit detection
- point out the counting algorithm is sensitive to the choice of “center”
- in particular, why choose NE as center? This will become important later.

Cycle counting via the Cycle Rotation Index

We can use CRI to test whether there are cycles. Let's get a bit more technical...

- Each treatment (i.e. combination of payoff matrix and action set) was played in at least 6 distinct sessions. We can treat data from different sessions as independent.
- Consider the null hypothesis: CRI is drawn from some distribution with **median** of 0.
- If null is true, then the 6 data points on CRI should fall to the left and right of 0 with equal probability. If say 5 or 6 out of 6 data points are all positive / all negative, this is evidence against the null.
- Compute p-value of observation using binomial distribution. This is known as the Wilcoxon test.

Cycle counting via the Cycle Rotation Index

Game condition	CCT	CT	CRI
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^*

where * indicates a Wilcoxon test p-value less than 0.05.

Authors say they have replicated the result of "no cycle" under stable payoffs in discrete setting.

However, Wang et al. point out if the center for computing CRI is moved, data from S discrete-mixed and S discrete-pure can also become significantly different from 0.

Some other results: amplitude of cycles

This table displays the mean squared deviation from NE. For each action set, the difference between deviation statistics in stable and unstable treatments is significant at 5% level. Difference between U_a and U_b has p-value of 0.20.

	Stable S	Unstable U_a	Unstable U_b
continuous-slow	0.014	0.044	0.048
continuous-instant	0.039	0.112	-
discrete-mixed	0.048	0.089	-
discrete-pure	0.093	0.129	-

Some other results: NE vs. TASP in unstable payoff matrix

The following are the predictions of NE and TASP:

Prediction	Rock	Paper	Scissors	Payoff
NE	0.25	0.25	0.50	48
TASP (U_a)	0.242	0.31	0.449	51.1

Some other results: NE vs. TASP in unstable payoff matrix

The following are the predictions of NE and TASP:

Prediction	Rock	Paper	Scissors	Payoff
NE	0.25	0.25	0.50	48
TASP (U_a)	0.242	0.31	0.449	51.1

The following are behavior observed in data:

Treatment	Rock	Paper	Scissors	Payoff
U_a continuous-instant	0.247	0.318	0.435	49.82
U_a continuous-slow	0.228	0.281	0.491	49.08
U_a discrete-mixed	0.225	0.342	0.433	49.70
U_a discrete-pure	0.205	0.337	0.458	50.71

Some other results: NE vs. TASP in unstable payoff matrix

The following are the predictions of NE and TASP:

Prediction	Rock	Paper	Scissors	Payoff
NE	0.25	0.25	0.50	48
TASP (U_a)	0.242	0.31	0.449	51.1

The following are behavior observed in data:

Treatment	Rock	Paper	Scissors	Payoff
U_a continuous-instant	0.247	0.318	0.435	49.82
U_a continuous-slow	0.228	0.281	0.491	49.08
U_a discrete-mixed	0.225	0.342	0.433	49.70
U_a discrete-pure	0.205	0.337	0.458	50.71

Red means rejected against NE.

Some other results: NE vs. TASP in unstable payoff matrix

The following are the predictions of NE and TASP:

Prediction	Rock	Paper	Scissors	Payoff
NE	0.25	0.25	0.50	48
TASP (U_a)	0.242	0.31	0.449	51.1

The following are behavior observed in data:

Treatment	Rock	Paper	Scissors	Payoff
U_a continuous-instant	0.247	0.318	0.435	49.82
U_a continuous-slow	0.228	0.281	0.491	49.08
U_a discrete-mixed	0.225	0.342	0.433	49.70
U_a discrete-pure	0.205	0.337	0.458	50.71

Red means rejected against TASP.

Concluding remarks

- many experimental innovations over previous literature
- first to report persistent cycles in RPS
- at same time, some results may be due to parts of the new interface other than the headlining “continuous time design”
    - try removing the heatmap
- technically this is not a learning paper
    - investigates only best response dynamic – but where does this dynamic arise?
    - simple extension: use heatmap to show expected payoff against empirical average strategy instead of current strategy