Heterogenous Heuristics in 3x3 Bimatrix Population Games

Marius I. Ochea*

14th April 2017

Abstract

We investigate population-level evolutionary dynamics resulting from individual-level, adaptive play both under homogenous ("self-play") and heterogenous ("mixed play") scenarios. In a class of bimatrix 3x3 normal form games (Sparrow and van Strien, 2008), that includes Rock-Paper-Scissors as a special case, rich limit behavior unfolds as game and heuristics parameters vary. In particular, a sequence of period-doubling bifurcations of limit cycles emerges under the perturbed best-reply dynamics and chaotic dynamics on the Hannan set appear under the no-regret dynamics.

JEL classification: C72, C73

Keywords: adaptive learning rules, imitation, regret-minimisation, logit choice

---

*THEMA and Dept. of Economics, Université de Cergy-Pontoise, 33, bd. du Port, 95011, Cergy-Pontoise Cedex, France. marius.ochea@u-cergy.fr
1 Introduction

A vast body of research on game dynamics focuses on limit behaviors (Nash equilibria, limit cycles) under homogenous learning rules. Indeed, the standard game dynamics - replicator, best-response, fictitious play - emerge when one imposes to each player the same procedure of making the game decisions. Dynamical systems arising from evolutionary games played in heterogenous populations - where players can use, and possibly revise based on realized performance, competing heuristics - are less studied. In this paper we investigate pairwise interaction of heuristics drawn from an adaptive toolbox containing various rules such as imitation, smooth best-reply, fictitious play, reinforcement learning, Brown-von Neumann-Nash (BNN) and regret-matching.

$3 \times 3$ bimatrix games are known to generate complex dynamics even in the case of homogenous learning. Sparrow and van Strien (2008) prove that in a class of $3 \times 3$ games that homogeneous fictitious play generates transition from periodic to chaotic behavior and back to periodic orbits. In a follow-up paper, van Strien and Sparrow (2011), study the topological aspects of such chaotic dynamics under Fictitious Play in the same class $3 \times 3$ bimatrix games. The chaotic patterns are also investigated by Aguiar and Castro (2010): chaos under homogeneous replicator dynamics in the class of Shapley $3 \times 3$ bimatrix games. Kaniovski et. al. (2000) document the existence of limit cycles even in $2 \times 2$ Coordination games played within heterogenous populations consisting of best-responders, conformists (imitate the majority) and non-conformists (imitate the minority). It is important to note that all these dynamics evolve on the product-space of marginal distribution of frequencies. Turning to the joint-space dynamics, Hart and Mas-Colell (2003) analyze homogenous population endowed with conditional (external) regret-minimising heuristics and show that the dynamics converge to the so-called Hannan set of non-positive regrets\footnote{To obtain sharper, point-predictions, e.g. correlated equilibrium, one has to to restrict the regret-minimising strategies to conditional (internal) regret.}. There is not much known about the game dynamics within the Hannan set and our simulations show that complicated, chaotic
trajectories may appear on this set.

The paper is organized as follows: Sector 2 briefly revisits the general model of (evolutionary) game dynamics, Sections 3 and 4 investigate the Shapley family of games under different homogeneous learning scenarios (imitation, logit, BNN). Section 5 introduces the heterogenous heuristics, bi-population dynamics whereas Section 6 looks at the limit sets of the no-regret dynamics.

2 Game Dynamics

For a normal form game with a finite set of available strategies $E = \{E_1, E_2, \ldots, E_n\}$ and the payoff matrix $A[n \times n]$, $x(t)$ denotes the $n$-dim vector of frequencies for each strategy/type $E_i$ at time $t$. $x$ belongs to the $n - 1$ dim simplex:

$$\Delta^{n-1} = \{ x \in \mathbb{R}^n : \sum_{i=1}^{n} x_i = 1, x_i \geq 0 \}$$

An evolutionary dynamic is a map assigning to each population game a differential equation on $\Delta^{n-1}$:

$$\dot{x} = V(x)$$

with the $i^{th}$ component of the vector field given by:

$$\dot{x}_i = V_i(x) = \text{inflow into strategy } i \text{ - outflow from strategy } i$$

Under the regular assumption of pairwise interaction and random matching of players the expected payoff vector $f(x)$, given the state of the population $x$, can be determined as:

$$f_i(x) = (Ax)_i$$
The population dynamic results from individual choices via revision protocols \( \rho_{ij}(f(x), x) \) which, for each pair \((i, j)\), defines the rate of switching \( \rho_{ij} \) from the currently played strategy \( i \) to strategy \( j \):

\[
\dot{x}_i = V_i(x) = \sum_{j=1}^{n} x_j \rho_{ji}(f(x), x) - x_i \sum_{j=1}^{n} \rho_{ij}(f(x), x).
\]

For instance, the pairwise proportional protocol (player \( i \) switches to strategy \( j \) at a rate proportional with the probability of meeting an \( j \)-strategist \( x_j \) and with the excess payoff of opponent \( j \), \( f_j(x) - f_i(x) \) iff positive:

\[
\rho_{ij}(f(x), x) = x_j[f_j(x) - f_i(x)]^+ 
\]

leads to the population-level homogenous Replicator Dynamics:

\[
\dot{x}_i = x_i[f_i(x) - \bar{f}(x)] = x_i[(A x)_i - x A x]
\]

The Logit revision protocol: \( \rho_{ij} \) represents the probability of player \( i \) switching to strategy \( j \):

\[
\rho_{ij}(f(x), x) = \frac{\exp[\eta^{-1} f_j(x)]}{\sum_k \exp[\eta^{-1} f_k(x)]} = \frac{\exp[\eta^{-1} (A x)_i]}{\sum_k \exp[\eta^{-1} (A x)_k]},
\]

and give rise to the population-level homogenous Logit dynamics \([\beta = \eta^{-1}]\):

\[
\dot{x}_i = \frac{\exp[\beta (A x)_i]}{\sum_k \exp[\beta (A x)_k]} - x_i
\]

If the intensity of choice parameter \( \beta \to \infty \), Logit approaches the Best-Reply Dynamics:

\[
\dot{x}_i = BR(x) - x_i
\]
3 Shapley 3x3 bi-matrix game

A bi-matrix - $A, B$ - game parameterized by the payoff-perturbation parameter $\mu$:

$$
\begin{bmatrix}
1, -\mu & 0, 0 & \mu, 1 \\
\mu, 1 & 1, -\mu & 0, 0 \\
0, 0 & \mu, 1 & 1, -\mu \\
\end{bmatrix}
$$

Payoff matrices for the row and column player, respectively:

$$
A = \begin{pmatrix}
1 & 0 & \mu \\
\mu & 1 & 0 \\
0 & \mu & 1
\end{pmatrix},
B^T = \begin{pmatrix}
-\mu & 1 & 0 \\
0 & -\mu & 1 \\
1 & 0 & -\mu
\end{pmatrix}, \mu \in (0, 1)
$$

Vectors of fractions in populations $A, B$, respectively:

$$
x = \begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix},
y = \begin{pmatrix}
y_1 \\
y_2 \\
y_3
\end{pmatrix}
$$

Expected payoff vectors:

$$
Ay = \begin{pmatrix}
y_1 + \mu y_3 \\
y_2 + \mu y_1 \\
y_3 + \mu y_2
\end{pmatrix},
B^T x = \begin{pmatrix}
x_3 - \mu x_1 \\
x_1 - \mu x_2 \\
x_2 - \mu x_3
\end{pmatrix}
$$

(1)

$$
B^T x = \begin{pmatrix}
-\mu & 0 & 1 \\
1 & -\mu & 0 \\
0 & 1 & -\mu
\end{pmatrix} \begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix} = \begin{pmatrix}
x_3 - \mu x_1 \\
x_1 - \mu x_2 \\
x_2 - \mu x_3
\end{pmatrix}
$$

(2)
Average payoffs in populations $A, B$:

$\mathbf{yB^T x} = y_1 (x_3 - \mu x_1) + y_2 (x_1 - \mu x_2) + y_3 (x_2 - \mu x_3) \quad (3)$

$\mathbf{xAy} = x_1 (y_1 + \mu y_3) + x_2 (y_2 + \mu y_1) + x_3 (y_3 + \mu y_2) \quad (4)$

With the standard simplex restrictions one obtains a 4-dimensional, continuous-time, dynamical system in $(x_1, x_2, y_1, y_2)$.

4 Homogenous Heuristics

4.1 Imitation

Bi-matrix Replicator Dynamics for $i, j = 1, 3$:

$\dot{x}_i = x_i((\mathbf{Ay})_i - \mathbf{x}^T \mathbf{Ay}) \quad (5)$

$\dot{y}_j = y_j((\mathbf{B^T x})_j - \mathbf{y}^T \mathbf{B^T x}) \quad (6)$
The dynamics under Replicator looks chaotic in certain parameter region, but this is hardly surprising as there are already examples of bimatrix games in the literature (Sato et al., 2002)
4.2 Logit choice

Logit Dynamics with precision level $\beta$ for $i, j = 1, 3$:

\[
\begin{align*}
\dot{x}_i &= \frac{e^{\beta(Ay)_i}}{\sum_{i=1}^{3} e^{\beta(Ay)_i}} - x_i \\
\dot{y}_j &= \frac{e^{\beta(B^T x)_j}}{\sum_{i=1}^{3} e^{\beta(B^T x)_j}} - y_i
\end{align*}
\]  

(7) (8)

Logit dynamics for $\mu = 0.8$

Phase plots for $\beta \in [0, 45]$

(a) $\beta = 19.1$-cycle 
(b) $\beta = 22.3$-cycle 
(c) $\beta = 41.1$-cycle 
(d) $\beta = 43.5$-cycle
Phase plots for $\beta > 50$

(a) $\beta = 50$

(b) $\beta = 100$

(c) $\beta = 300$

(d) $\beta = 700$

4.2.1 Continuation of a detected Hopf point

First, we fix $\mu = 0.8$ and we follow the Period-Doubling route to chaos as the intensity of choice $\beta$ increases (Matcont output). The barycentric equilibrium $\left[\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right]$ is "continued" with respect to $\beta$ and a stable limit cycle is created via a Hopf bifurcation: (label = H , $x = (0.333333, 0.333333, 0.333333, 0.333333, \beta = 18.669513$. First Lyapunov coefficient = -7.427464e+001). The sequence of three period-doubling (PD) bifurcations unfolds as follows:
1. first PD bifurcation of the stable limit cycle: 2-cycle (in red)

Period Doubling (period = 8.696211e-001, $\beta = 22.24087$)

Normal form coefficient = -8.485368e+002

2. second PD bifurcation of the 2-limit cycle: 4-cycle (in red)

Period Doubling (period = 1.202897e+000, $\beta = 41.07238$)

Normal form coefficient = -1.113931e+003

3. third PD bifurcation of the 4-limit cycle: 8-cycle (in red)

Period Doubling (period = 2.375447e+000, $\beta = 43.41395$)

Normal form coefficient = -1.200818e+003

The blue closed curves below are stable limit cycles of increasing amplitude as parameter $\beta$ increases. For each such cycles the package computes a certain test function\footnote{see the package documentation for the exact expression of the test function: http://sourceforge.net/projects/matcont/}. A 0 of this function means that a period-doubling bifurcation of the limit cycle is detected. The corresponding cycle is then colored in red and a PD label is attached to it. We re-initiate the continuation routine for each such Hopf-PD bifurcation detected, which means that, effectively, in the panels below 1-, 2-, 4-, 8-cycle are "continued", respectively with the resulting \textit{family} of such cycles depicted in blue.
Continuation of the Hopf and Hopf-Period Doubling points in the ($\beta, \mu$) space

(a) PD of 1-cycle, $\beta = 22.24087$

(b) PD of 2-cycle, $\beta = 41.07238$

(c) PD of 4-cycle, $\beta = 43.41395$

(d) cont. of 8-cycle, no PD of 8-cycle detected
No codim II singularities detected along these curves.

We could not detect a period-halving in "backward time" as we decrease $\beta$ from, say 100 to the "chaotic" region: $[50, 80]$. We need to understand better what happens after the sequence of 3 period-doubling bifurcations of limit cycles collapses into chaos, i.e. the system behavior for $\beta \in [50, 100]$ and also for $\beta > 100$ in the $\beta \to \infty$ "best-response" limit.

4.2.2 Chaotic Dynamics?

Largest Lyapunov exponent plots [numerical evidence for chaos]

a) Fix intensity of choice $\beta$ to increasingly higher values and let payoff perturbation $\mu$ vary in $[0, 1.2]$
b) Fix $\mu$ to Shapley (0.4), chaotic(0.7, 0.8), anti-Shapley (1) values, respectively and let
\( \beta \) vary in \([0, 500]\)

(b) \( \mu = 0.7 \)

(c) \( \mu = 0.8 \)

(d) \( \mu = 1 \)

4.3 BNN dynamics

Continuous-time, regret-minimisation dynamics for bi-matrix games \((A,B)\)

\[
\dot{x}_i = \frac{[(Ay)_i - x^T Ay]_+}{\sum_k [(Ay)_k - x^T Ay]_+} - x_i \tag{9}
\]

\[
\dot{y}_j = \frac{[(BTx)_j - y^T BTx]_+}{\sum_k [(BTx)_k - y^T BTx]_+} - y_j \tag{10}
\]
5 Heterogenous Heuristics

5.1 Replicator vs. Logit

\[
\begin{align*}
\dot{x}_i &= x_i ((Ay)_i - x^T Ay) \\
\dot{y}_j &= \frac{e^{\beta(B'y)_j}}{\sum_{i=1}^{3} e^{\beta(B'y)_j}} - y_j
\end{align*}
\]
5.2 Replicator vs. BNN

\[
\begin{align*}
\dot{x}_i &= x_i((Ay)_i - x^T Ay) \\
\hat{y}_j &= \frac{[(B^T x)_j - y^T B^T x]_+}{\sum_k [(B^T x)_k - y^T B^T x]_+} - y_j
\end{align*}
\]
5.3 BNN vs. Logit

\[ \dot{x}_i = \sum_k [(Ay)_k - x^T Ay]_+ - x_i \]

\[ \dot{y}_j = \frac{e^{\beta(B'x)_j}}{\sum_{i=1}^3 e^{\beta(B^T x)_j}} - y_j \]

(a) \( \mu = 0.8, \beta = 200 \)  
(b) \( \mu = 0.8, \beta = 200 \)
6 Hart Mas-Colell Dynamics (no-regret dynamics)

Bi-matrix game $A, B \ [n \times n]$

Dynamics run on the joint probability space:

$$Z = \begin{pmatrix}
  z_{11} & \ldots & z_{1n} \\
  \vdots & \ddots & \vdots \\
  z_{n1} & \ldots & z_{nn}
\end{pmatrix}, \sum_{i,j} z_{ij} = 1$$

$$\dot{z}_{ij} = p_i q_j - z_{ij}$$

with,

$$p_i = \frac{[(Ay)_i - a(z)]_+}{\sum_k [(Ay)_k - a(z)]_+}$$

$$q_j = \frac{[(B^T x)_j - b(z)]_+}{\sum_k [(B^T x)_k - b(z)]_+}$$

where

$$a(z) = \sum_{i,j} a_{ij} z_{ij}, b(z) = \sum_{i,j} b_{ij} z_{ij}$$

Marginal probabilities:

$$x_i = \sum_j z_{ij}, y_j = \sum_i z_{ij}$$

6.1 Shapley game with no-regret dynamics

$$A = \begin{pmatrix} 1 & 0 & \mu \\ \mu & 1 & 0 \\ 0 & \mu & 1 \end{pmatrix}, \quad B^T = \begin{pmatrix} -\mu & 0 & 1 \\ 1 & -\mu & 0 \\ 0 & 1 & -\mu \end{pmatrix}$$
\[
a(z) = \sum_{i,j} a_{ij} z_{ij} = z_{11} + z_{22} + z_{33} + \mu z_{21} + \mu z_{13} + \mu z_{32}
\]

\[
b(z) = \sum_{i,j} b_{ij} z_{ij} = z_{21} + z_{13} + z_{32} - \mu z_{11} - \mu z_{22} - \mu z_{33}
\]

Attractors for different payoff perturbations $\mu$ in the $(z_{11}, z_{22})$ space (in numerics used $max(x, 0.001)$)

(a) $\mu = 0.2$

(b) $\mu = 0.5$

(c) $\mu = 0.615$

(d) $\mu = 1.15$
6.2 Symmetric game of cyclical dominance and no-regret

\[ A = \begin{pmatrix}
0 & -(1 + \varepsilon) & 1 \\
1 & 0 & -(1 + \varepsilon) \\
-(1 + \varepsilon) & 1 & 0
\end{pmatrix} \]

No-regret Dynamics:

\[ \dot{z}_{ij} = p_i p_j - z_{ij}, \quad z_{ji} = z_{ij} \]
\[ p_i = \frac{[(Ay)_i - a(z)]_+}{\sum_k [(Ay)_k - a(z)]_+} \]
\[ a(z) = \sum_{i,j} a_{ij} z_{ij} = -\varepsilon (z_{12} + z_{23} + z_{31}) \]

Attractors for different payoff perturbations \( \varepsilon \) in the \((z_{11}, z_{12}, z_{13})\) space (in numerics used \( \max(x, 0.0001) \))
7 Competing heuristics (in progress)

Row and column player actively chose from the adaptive toolbox (imitation, smooth best-response, regret-minimisers, etc) and update their heuristics according to their past performance. Key questions to address: which heuristic will strive in population $A, B$? can equilibrium/cyclical co-existence of heuristics arise?
References


