Year: 2010

Best-response dynamics in a birth-death model of evolution in games

Alós-Ferrer, C; Neustadt, I


Posted at the Zurich Open Repository and Archive, University of Zurich. http://www.zora.uzh.ch

Best-response dynamics in a birth-death model of evolution in games

Abstract

We consider a model of evolution with mutations as in Kandori et al (1993) [Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29-56], where agents follow best-response decision rules as in Sandholm (1998) [Sandholm, W., 1998. Simple and clever decision rules for a model of evolution. Economics Letters 61, 165-170]. Contrary to those papers, our model gives rise to a birth-death process, which allows explicit computation of the long-run probabilities of equilibria for given values of the mutation rate and the population size. We use this fact to provide a direct proof of the stochastic stability of risk-dominant equilibria as the mutation rate tends to zero, and illustrate the outcomes of the dynamics for positive mutation rates.
Best-response dynamics in a birth-death model of evolution in games*

Carlos Alós–Ferrer† and Ilja Neustadt‡

This version: October 2005

Abstract

We consider a model of evolution with mutations as in Kandori et al (1993) [Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61, 29-56], where agents follow best-response decision rules as in Sandholm (1998) [Sandholm, W., 1998. Simple and clever decision rules for a model of evolution. Economics Letters 61, 165-170]. Contrary to those papers, our model gives rise to a birth-death process, which allows explicit computation of the long-run probabilities of equilibria for given values of the mutation rate and the population size. We use this fact to provide a direct proof of the stochastic stability of risk-dominant equilibria as the mutation rate tends to zero, and illustrate the outcomes of the dynamics for positive mutation rates.

Keywords: Coordination Games, Learning, Mutation, Birth-Death Processes.

JEL Classification Numbers: C72, D83.

Many models of learning in games follow the lead of Kandori, Mailath & Rob (1993) and Young (1993). In this stream of the literature, a basic behavioral rule for agents is postulated and perturbed with a mutation probability. The focus is then on the limiting outcome as the mutation probability tends to zero. There are well-established techniques (see e.g. Ellison (2000) for a review) which allow the identification of the long-run outcomes without explicitly analyzing the stochastic process.

In many such learning models, revision opportunities (learning draws) are independent across agents, so that all of them might revise simultaneously. An alternative, advocated e.g. by Blume (1993, 1995, 2003), Benaïm & Weibull (2003), and Friedman & Mezzetti (2001), is posed by models where each period, a single agent is randomly sampled to learn.¹

This way, the model becomes a birth-death process. In this note, we make use of the analytical tractability of such processes to obtain exact results (as opposed to only limit ones) in a model of learning in games.

---

*The authors gratefully acknowledge financial support from the Austrian Science Fund (FWF) under Projects P18141-G09 and P15281.

†Department of Economics, University of Konstanz, Box D-150, D-78457 Konstanz (Germany). E-mail: Carlos.Alos-Ferrer@uni-konstanz.de

‡Socioeconomic Institute, University of Zurich, Hottingerstrasse 10 CH-8032 Zurich (Switzerland). E-mail: ilja.neustadt@soi.unizh.ch

¹See Alós-Ferrer (2003) for a discussion. This formulation is basically equivalent to a more realistic model where agents receive revision opportunities in continuous time following independent Poisson processes. See Blume (1995, Section 3).
We consider a population of \( N \geq 2 \) players who are repeatedly randomly matched to play a symmetric, \( 2 \times 2 \) game with strategy set \( \{A, B\} \) and payoff matrix

\[
\begin{array}{c|cc}
 & A & B \\
\hline
A & a, a & b, c \\
B & c, b & d, d \\
\end{array}
\]

We assume \( a > c \), \( b < d \), and \( a + b < c + d \). This game has two strict, symmetric Nash equilibria given by \((A, A)\) and \((B, B)\), such that equilibrium \((B, B)\) is risk-dominant. However, \((A, A)\) could be payoff-dominant if \( a > d \). There is also a mixed-strategy Nash equilibrium \((\alpha^*, \alpha^*)\) where \( A \) is played with probability

\[
\alpha^* = \frac{d - b}{a - c + d - b}.
\]

The risk dominance of \((B, B)\) is equivalent to the fact that \( \alpha^* > \frac{1}{2} \).

The evolution of the population is modelled by a Markov process. The state of population \( n \in \{0, 1, \ldots, N - 1, N\} \) is the number of players currently selecting strategy \( A \). Then the (expected) payoff of a player choosing strategy \( A \) or \( B \) at state \( n \) is given by the expressions

\[
\pi(A, n) = \frac{n - 1}{N - 1} a + \frac{N - n}{N - 1} b,
\]

\[
\pi(B, n) = \frac{n}{N - 1} c + \frac{N - n - 1}{N - 1} d
\]

which take into account the fact that a player does not play against himself.\(^2\)

We follow Sandholm (1998) and Alós-Ferrer (2003) and assume that players are myopic best-responders.\(^3\) This decision rule can be summarized as follows:

When playing strategy \( A \), switch to strategy \( B \) if \( \pi(A, n) < \pi(B, n - 1) \).

When playing strategy \( B \), switch to strategy \( A \) if \( \pi(B, n) < \pi(A, n + 1) \).

It remains to specify players’ behavior in the (nongeneric) case of payoff ties. We postulate a fixed probability \( \eta \in [0, 1] \) of switching from the current strategy if a tie occurs, reflecting e. g. unmodelled switching costs.\(^4\)

Elementary computations show that an A-player switches to B if

\[
\pi(A, n) < \pi(B, n - 1) \iff (n - 1)(a - c) < (N - n)(d - b) \iff n < (N - 1)\alpha^* + 1 =: n_A
\]

A B-player switches to A if

\[
\pi(B, n) < \pi(A, n + 1) \iff n(a - c) > (N - n - 1)(d - b) \iff n > (N - 1)\alpha^* =: n_B
\]

\(^2\)If the number of players is odd, there is a probability \( \frac{1}{N} \) that a given player is not matched. The expected payoffs then are \( \pi'(s, n) = \frac{N-1}{N} \pi(s, n) + \frac{1}{N} k \), where \( k \) is the default payoff of not being matched.

\(^3\)This decision rule is not the one implicitly used by Kandori et al. (1993). See Sandholm (1998) for details.

\(^4\)As we will see, this is of no consequence for the analysis. Sandholm (1998) assumes that \( \eta = 0 \), which corresponds to a “switch only if strictly better” rule.
Clearly, \( n_A - n_B = 1 \). The thresholds \( n_A, n_B \) are closely linked to the mixed-strategy equilibrium strategy \( \alpha^* \). Notice that \( \frac{n_A}{N} \leq \alpha^* \leq \frac{n_B}{N} \).

We distinguish 2 cases. In the generic case, \( n_A, n_B \in \mathbb{N} \). In the non-generic case, \( n_A, n_B \notin \mathbb{N} \). We will only consider the generic case in detail and merely report the results for the non-generic case.

In the generic case, let \( n^* \) be the only integer in \([n_B, n_A]\). That is, \( n^* = \lfloor n_A \rfloor = \lceil n_B \rceil \). Notice that, since \( \alpha^* N \in [n_B, n_A] \), \( n^* \) can be taken to be an integer approximation to the mixed-strategy equilibrium. Note that \( n^* > n_B = (N - 1)\alpha^* > \frac{1}{2} N \) whenever \( \alpha^* > \frac{1}{2} \frac{N}{N - 1} \).

Since \( \alpha^* > \frac{1}{2} \) by risk dominance of \((B, B)\), it follows that, given a fixed game, this latter condition is fulfilled for \( N \) large enough. Given a fixed population size, it is fulfilled when \((B, B)\) is significantly risk-dominant, i.e. \( \alpha^* \) is not too close to \( \frac{1}{2} \).

In the generic case, payoff ties are not possible. \( A \)-players switch to \( B \) if and only if \( n \leq n^* \), while \( B \)-players switch to \( A \) if and only if \( n \geq n^* \).

Consider now the following dynamics. Each period, exactly one player is sampled at random (uniformly) from the population and receives the opportunity to revise his strategy. If the current state is \( n \), the probability for an \( A \)-player to be selected is \( n/N \), while the corresponding probability for a \( B \)-player is \((N - n)/N \).

With probability \( 1 - \varepsilon \), the selected player takes the action prescribed by myopic best reply. With probability \( 0 < \varepsilon < \frac{1}{2} \), he mutates and takes the opposite action.

Let \( P_{ij} \) be the probability of transition from a state \( i \) to another state \( j \). Clearly, \( P_{ij} = 0 \) whenever \(|i - j| > 1\), i.e. we have a birth-death process. Further,

\[
P_{n,n-1} = \begin{cases} \frac{(1 - \varepsilon) n}{N} & \text{if } 1 \leq n \leq n^* \\ \varepsilon \frac{n}{N} & \text{if } n^* + 1 \leq n \leq N \end{cases}
\]

\[
P_{n,n+1} = \begin{cases} \varepsilon \frac{N-n}{N} & \text{if } 0 \leq n \leq n^*-1 \\ (1 - \varepsilon) \frac{N-n}{N} & \text{if } n^* \leq n \leq N - 1 \end{cases}
\]

and \( P_{nn} = 1 - P_{n,n-1} - P_{n,n+1} \) for all \( n = 0, \ldots, N \).

The invariant distribution \( \rho \) has full support and fulfills the detailed balance condition,\(^5\)

\[
\rho(n) P_{n,n+1} = \rho(n+1) P_{n+1,n}.
\]

Define, for convenience, \( \beta = \frac{\varepsilon}{1 - \varepsilon} \). Then,

\[
\frac{\rho(n+1)}{\rho(n)} = \begin{cases} \beta \frac{N-n}{N+1} & \text{if } 0 \leq n \leq n^*-1 \\ \frac{1}{\beta} \frac{N-n}{N+1} & \text{if } n^* \leq n \leq N - 1 \end{cases}
\]

Iterating, we obtain, for any \( 1 \leq n \leq n^* \),

\[
\rho(n) = \rho(0) \beta^{n-1} \prod_{j=0}^{n-1} \frac{N-j}{j+1}
\]

and for any \( n^* + 1 \leq n \leq N \)

\[
\rho(n) = \rho(n^*) \frac{1}{\beta^{n-n^*}} \prod_{j=n^*}^{n-1} \frac{N-j}{j+1} = \beta^{2n^*-n} \rho(0) \prod_{j=0}^{n-1} \frac{N-j}{j+1}.
\]

These computations allow us to obtain the following result.

---

\(^5\)Discrete-time birth-death processes are defined as Markov chains such that (i) the state space is either the nonnegative integers or a finite set \([0, \ldots, N]\), and (ii) for every state \( n \), the only positive-probability transitions are to states \( n, n - 1, \) and \( n + 1 \), with the last two being strictly positive. The detailed balance condition follows directly from the definition of invariant distribution applied to this particular case. See e.g. Feller (1968, p.396).
Theorem 1. Consider the generic case. Let \( 0 < \varepsilon < 1 \) and \( N \geq 2 \) such that \( \alpha^* > \frac{1}{2} \frac{N}{N-1} \).

Then,

(a) The invariant distribution fulfills

\[
\rho(0) = \left[ 1 + \sum_{n=1}^{N} \beta^n \binom{N}{n} + \sum_{n=n^*+1}^{N} \beta^{2n^*-n} \binom{N}{n} \right]^{-1}
\]

where \( \binom{N}{n} = \prod_{j=0}^{n-1} \left( \frac{N-j}{j+1} \right) \) are the binomial coefficients.

(b) \( \rho(N)/\rho(0) = \beta^{2n^*-N} \approx \beta^{(2\alpha^*-1)N} \).

(c) For fixed \( N \), \( \lim_{\varepsilon \to 0} \rho(0) = 1 \).

(d) For fixed \( \varepsilon \), \( \lim_{N \to \infty} \rho(0) = \lim_{N \to \infty} \rho(N) = 0 \).

Proof. (a) follows from (1) and (2) observing that \( \sum_{n=0}^{N} \rho(n) = 1 \). (b) is a direct consequence of (2). (c) follows from (a) noting that the condition \( \alpha^* > \frac{1}{2} \frac{N}{N-1} \) implies \( n^* > \frac{1}{2} N \).

To show (d), notice that the expression in brackets in (3) is bounded below by 1 + \( \beta N \), which diverges to \( +\infty \) (for fixed \( \varepsilon > 0 \)) as \( N \to +\infty \). Thus \( \rho(0) \) converges to 0. Since \( n^* > \frac{1}{2} N \), \( \beta^{2n^*-N} \) is bounded above by 1 (recall that \( 0 < \varepsilon < \frac{1}{2} \), thus \( 0 < \beta < 1 \)) and (b) implies that \( \rho(N) \) also converges to 0. \( \blacksquare \)

A state \( s \) is stochastically stable (Kandori et al. (1993)) if \( \lim_{\varepsilon \to 0} \rho(s) > 0 \). By the ergodic theorem, stochastically stable states are the only ones which are observed a significant fraction of time in the long run. Thus, part (c) states that the only stochastically stable state is full coordination on the risk-dominant equilibrium. Although this result can also be proven with the familiar “mutation-counting” techniques, the direct proof presented here has the advantage of providing the closed-form formula (a) for \( \rho(0) \) when \( \varepsilon \) is strictly positive.

Clearly, the best-reply process will spend most of its time at the states 0 and \( N \), where full coordination on the pure-strategy Nash equilibria is achieved. Thus, for \( \varepsilon \) small but positive, the odds ratio \( \rho(N)/\rho(0) \) gives an approximation for the proportion of time that the system spends on each equilibrium, for \( \varepsilon > 0 \) and finite \( N \). Since risk dominance amounts to \( \alpha^* > \frac{1}{2} \frac{N}{N-1} \), part (b) constitutes an extremely simple summary of the results.

Finally, part (d) observes that, for fixed mutation rates but increasing population, the long-run distribution concentrates its weight in the interior of the state space. This result is natural due to the birth-death nature of the process. Take for instance a state with low \( n \), where the best-reply dynamics points towards \( n-1 \). As the population grows but \( \varepsilon \) remains fixed, the probability \( \frac{n}{N} \) of sampling an \( A \)-player, who would then switch to \( B \), becomes smaller than the probability of sampling a \( B \) player who then mutates to \( A \), \( \varepsilon \frac{N-n}{N} \). Hence, the mutation-plus-best-reply dynamics points away from the corners. In a sense, this is the birth-death version of the well-known critique that, in models of learning with mutation, results become less plausible as the population size becomes large.\(^6\)

We proceed now to briefly report on the (analogous) results for the non-generic case. In this case, \( n_A, n_B \in \mathbb{N} \) and payoff ties occur at states \( n_A \) and \( n_B \). Recall that \( \eta \in [0, 1] \) is the probability of switching from the current strategy if a tie occurs.

Thus, \( A \)-players switch to \( B \) if \( n < n_A \), do not change if \( n > n_A \), and randomize if \( n = n_A \). \( B \)-players switch to \( A \) if \( n > n_B = n_A - 1 \), do not change if \( n < n_B \), and randomize if \( n = n_B \).

\(^6\)In Kandori et al. (1993), this leads to the observation that the expected time of convergence to the stochastically stable state is extremely long for large population sizes.
This induces the transition probabilities

\[
P_{n,n-1} = \begin{cases} 
\left(1 - \varepsilon\right) \frac{n}{N} & \text{if } 1 \leq n \leq n_B \\
\left(1 - \varepsilon\right)\eta + (1 - \eta)\varepsilon \frac{n}{N} & \text{if } n = n_A \\
\varepsilon \frac{n}{N} & \text{if } n_A + 1 \leq n \leq N 
\end{cases}
\]

\[
P_{n,n+1} = \begin{cases} 
\varepsilon \frac{N-n}{N} & \text{if } 0 \leq n \leq n_B - 1 \\
\left(1 - \varepsilon\right)\eta + (1 - \eta)\varepsilon \frac{N-n}{N} & \text{if } n = n_B \\
\left(1 - \varepsilon\right) \frac{N-n}{N} & \text{if } n_A \leq n \leq N - 1 
\end{cases}
\]

Letting again \( \beta = (1 - \varepsilon)/\varepsilon \), iteration of the detailed balance condition yields

\[
\rho(n) = \beta^n \rho(0) \binom{N}{n} \text{ for } 1 \leq n \leq n_B \\
\rho(n_A) = \frac{N-n_B}{n_A} \rho(n_B) = \beta^{n_A} \rho(0) \binom{N}{n_A} \\
\rho(n) = \beta^{2n_A-n-1} \rho(0) \binom{N}{n} \text{ for } n_A + 1 \leq n \leq N
\]

Hence, ties play no role whatsoever. This yields the following analogous equation to (3)

\[
\rho(0) = \left[1 + \sum_{n=1}^{n_A-1} \beta^n \binom{N}{n} + \beta^{n_A-1} \binom{N}{n_A} + \sum_{n=n_A+1}^{N} \beta^{2n_A-n-1} \binom{N}{n}\right]^{-1} \tag{4}
\]

To conclude, we make use of (3) and (4) to illustrate the outcomes of the dynamics for positive mutation rates and finite population size. Consider a coordination game as above with \( a = 1, b = c = 0, \) and \( d > 1 \). Figure 1(a) represents \( \rho(0) \) for \( d = 2 \) and \( N = 50 \) as \( \varepsilon \) varies. Stochastic stability implies that the full probability piles up on 0 as \( \varepsilon \to 0 \), but we can observe that \( \rho(0) \) has a significant probability for small but strictly positive values of \( \varepsilon \).

Figure 1(b) represents the invariant distribution for the same parameter values, fixing \( \varepsilon = 0.01 \). The shape is, of course, as expected, with probability piling up (60%) on state 0. This illustrates also that, for fixed values of \( \varepsilon \) and \( N \), full coordination is achieved only part of the time (60% in the particular case represented). The representation is truncated at the state \( n = 10 \); for \( n \geq 11 \), we have that \( \rho(n) < 3 \times 10^{-12} \).

We then perform a first sensitivity illustration in figure 1(c). This figure represents \( \rho(0) \) for \( \varepsilon = 0.01 \) and \( N = 50 \) as \( d \) varies between 1 and 4. The weight \( \rho(0) \) is almost constant (resulting in a practically flat representation), as long as \((B,B)\) is risk-dominant. As \( d \) approaches 1, though, \((B,B)\) becomes "less risk-dominant" and the condition \( \alpha^* > \frac{1}{2} \frac{N}{N-1} \) is violated,7 making both 0 and \( N \) stochastically stable and \( \rho(0) = \rho(1) = 0.3 \).

Figure 1(d) represents \( \rho(0) \) for \( d = 2 \) and \( \varepsilon = 0.01 \) as \( N \) varies. As predicted by Theorem 1(d), for a fixed mutation rate, coordination becomes less probable as population grows.

Figure 1(e) clarifies this point. This figure shows the invariant distribution for \( d = 2, \varepsilon = 0.01, \) and \( N = 1001 \), with the range truncated at \( N = 50 \) (\( \rho(n) < 2 \times 10^{-20} \) for all \( n \geq 51 \)). With this large population size,8 the dynamics points away from the corners. Risk dominance, though, leads to the distribution being skewed towards 0. Most of the probability piles up very close to but away from state 0 (recall that the figure shows only the range 0–50, whereas the population size is 1001). Figure 1(f) further illustrates this, showing the full invariant distribution for \( d = 2, \varepsilon = 0.1, \) and the original population size \( N = 50 \). For such a high mutation rate, a population size of 50 is already "large" and again the dynamics points away from the corners.

---

7 Equations (3) and (4) still hold, though. Examination of (3) reveals the reason for the flatness of the representation. As \( d \) changes, (3) changes only in terms \( \beta^n \) for \( n \geq n^* \). But, for \( d > 1 \), we have that \( n^* \geq 25 \).

8 We choose 1001 because 1000 lies in the non-generic case.
Figure 1: These figures show variations with respect to a default situation with population size $N = 50$, mutation rate $\varepsilon = 0.01$, and a risk-dominant equilibrium with payoff $d = 2$ ($a = 1$, $b = c = 0$). (a) represents the weight $\rho(0)$ of the risk-dominant equilibrium as $\varepsilon$ varies. (b) shows the (truncated) invariant distribution for the given parameters. (c) represents $\rho(0)$ as $d$ varies. (d) represents $\rho(0)$ as $N$ increases. (e) shows the (truncated) invariant distribution for $N = 1001$. (f) shows the full invariant distribution for $\varepsilon = 0.1$. 
References


