

EQUILIBRIUM SELECTION IN EVOLUTIONARY GAMES WITH IMPERFECT MONITORING

HSIAO-CHI CHEN,* *National Taipei University*

YUNSHYONG CHOW,** *Academia Sinica*

Abstract

In this paper we analyze players' long-run behavior in evolutionary coordination games with imperfect monitoring in a large population. Players can observe signals corresponding to other players' unseen actions and use the proposed simple or maximum likelihood estimation algorithm to extract information from the signals. In the simple learning process we find conditions for the risk-dominant and the non-risk-dominant equilibria to emerge alone in the long run. Furthermore, we find that the two equilibria can coexist in the long run. In contrast, the coexistence of the two equilibria is the only limit distribution under the maximum likelihood estimation learning algorithm. We also analyze the long-run equilibria of other 2×2 symmetric games under imperfect monitoring.

Keywords: Coordination game; risk-dominant equilibrium; imperfect monitoring

2000 Mathematics Subject Classification: Primary 91A22; 60J20

1. Introduction

The multiplicity of Nash equilibria has weakened the prediction power and application potentials of game theory on human behavior. The coordination game with risk- and payoff-dominant equilibria is a typical example. The evolutionary learning process is a dynamic method to refine Nash equilibria. Various hypotheses have been given in the literature to characterize players' boundedly rational behavior, resulting in the distinct emergence of Nash equilibrium in the long run. For instance, under the best-response dynamics, Young (1993), Ellison (1993), and Blume (1993), (1995) showed that players will eventually coordinate at a risk-dominant equilibrium in the long run. However, the equilibrium could be changed to a payoff-dominant equilibrium when state-dependent (see, e.g. Bergin and Lipman (1996)) or time-dependent mutations (see, e.g. Robles (1998) and Chen and Chow (2001)) are imposed. On the other hand, under Darwinian-type dynamics, Kandori *et al.* (1993) and Robson and Vega-Redondo (1996) demonstrated that different equilibria will emerge in the long run when players' matching methods are distinct. Kandori *et al.* (1993) showed that the risk-dominant equilibrium will be

Received 14 September 2007; revision received 23 January 2008.

* Postal address: Department of Economics, National Taipei University, 151 University Road, San-Shia, Taipei County 23741, Taiwan, R.O.C. Email address: hchen@mail.ntpu.edu.tw

Funding from the National Science Council (project number: NSC 92-2415-H-305-001) is gratefully acknowledged. Valuable suggestions provided by an anonymous referee are also highly appreciated. Of course, all remaining errors belong to us.

** Postal address: Institute of Mathematics, Academia Sinica, Taipei, Taiwan 115, R.O.C.

Email address: chow@math.sinica.edu.tw

Funding from the National Science Council in Taiwan is gratefully acknowledged.

chosen when individuals play with all other players, while Robson and Vega-Redondo (1996) proved that the payoff-dominant equilibrium will be selected when players are randomly paired to play the game multiple rounds per period. However, Miekisz (2005) showed that the two models would yield the same results if the number of players approaches infinity.

Although acquiring different conclusions, all the above models assume that players' actions and payoffs are observable. However, examples of unseen players' actions, with the visual signals relating stochastically to the underlying actions, are numerous in the real world. For instance, although the regulator does not know a firm's exact pollution-abatement activities, pollutant emissions can be measured and depend on a firm's abatement levels and some random factors, such as weather conditions. In principal and agent problems, agents' efforts are usually unseen to principals. However, outputs are correlated with agents' efforts and can be observed. In oligopolistic markets, output prices are indicators of the unknown product quality even though they are also affected by factors such as stochastic market demands.

In this paper we aim to analyze equilibrium selection for a large population under evolutionary learning processes with imperfect monitoring of players' actions and payoffs. Our models are built in Section 2. At the beginning of each time period, players know only their own actions and payoffs that occurred in the last period, while the signals corresponding to other players' underlying actions in the last period are revealed publicly. There are two possible signals for each action: the *prime* signal is the one more likely to occur. The correlations between signals and the underlying actions are assumed to be fixed over time and across players. Two learning schemes are proposed to represent players' distinct ways of using signals to form their beliefs about their opponents' current plays. Then, based on the beliefs, players choose actions to maximize their expected current payoffs. Different from previous evolutionary models examined in Vega-Redondo (1995), we suppose no mutation on players' rational choices as well as on players' expectation updating process.

In the simple learning process players are assumed to use signals' occurring frequencies in the last period to simultaneously forecast their opponents' current behavior. The associated dynamic system is a two-state Markov chain. In previous evolutionary models, mutation is the driving force making transitions between equilibria possible. In contrast, random signals in our model play the role of mutation. As shown in Theorem 1, below, players' long-run behavior in a large population is completely determined by the correlation between the signals and their underlying actions. This is because the law of large numbers (LLN) is in force as the population size grows to infinity. As a consequence, players would eventually select the risk-dominant equilibrium in case the non-risk-dominant strategy and its prime signal are not highly correlated. For the rest cases, the non-risk-dominant equilibrium could emerge alone or players could visit both equilibria with positive probability in the long run.

Inspired by the broad usage of the maximum likelihood estimation (MLE) in empirical works, we use the MLE as our second learning algorithm. In the MLE learning process, players are assumed to adopt the maximum likelihood estimate of the observed signals to form expectations about their opponents' current plays. The associated dynamic system is also a two-state Markov chain. However, unlike in the simple learning process, the coexistence of the risk-dominant and the non-risk-dominant equilibria is shown in Theorem 2, below, as the only long-run limit for a large population. This demonstrates that the equilibrium selection is quite sensitive to the signal handling scheme when players cannot observe other players' actions and payoffs.

Besides coordination games, we also examine the limit distribution of other 2×2 symmetric games under these two learning algorithms. Similar conclusions are reached.

Our models differ from the repeated prisoner’s dilemma games with imperfect monitoring (see, e.g. Fudenberg *et al.* (1994), Sekiguchi (1997), Compte (2002), and Mailath and Morris (2002)) in that our players are myopically rational while their players take actions to maximize the expected discounted payoffs (i.e. long sightedly rational). Chen *et al.* (2006) examined myopically rational players’ long-run behavior under imperfect monitoring, but constructed a model different from ours in two respects. Firstly, they adopted the Bayesian learning process, which is quite complicated and less compatible with the spirit of players’ bounded rationality. Secondly, they only analyzed the coordination games, while we consider all 2×2 symmetric games.

The rest of this paper is organized as follows. In Section 2 we present the coordination-game models and the associated outcomes. In Section 3 we analyze other 2×2 symmetric games. Finally, we draw our conclusions in Section 4 and give the proofs of Theorems 1, 2, and 4 in Section 5.

2. The models

There are N players, where N is finite and even. At each time period t , $t = 0, 1, 2, \dots$, players are randomly and independently matched in pairs to play the coordination game shown in Figure 1 once.

	A	B
A	a, a	b, c
B	c, b	d, d

FIGURE 1.

Here $\{A, B\}$ is the action set for all players and a, b, c , and d are payoffs with $a - c > d - b > 0$. Under the present coordination game, (A, A) is called the risk-dominant equilibrium and (B, B) is called the non-risk-dominant equilibrium. In conventional coordination games an extra condition, $d > a$, is imposed. Then (B, B) is referred to as the payoff-dominant equilibrium. Later, our results will show that it is the correlation between actions and signals which determines players’ long-run behavior and not whether (B, B) is payoff-dominant. Thus, the condition $d > a$ is not assumed to hold here. Define

$$q^* := \frac{d - b}{(a - c) + (d - b)} < \frac{1}{2}. \tag{1}$$

Then $(q^*, 1 - q^*)$ is the unique mixed Nash equilibrium. It is assumed that players know only their own actions and payoffs at each time period. However, signals corresponding to players’ underlying actions are publicly observable at each time period. Denote by $\{\bar{A}, \bar{B}\}$ the signal set for all players. The relations between signals and the underlying actions are described by the following conditional probabilities:

$$P(\bar{A} | A) = u \quad \text{and} \quad P(\bar{B} | B) = v, \quad \text{where } \frac{1}{2} < u, v < 1. \tag{2}$$

Hence, $P(\bar{B} | A) = 1 - u < \frac{1}{2}$ and $P(\bar{A} | B) = 1 - v < \frac{1}{2}$. This means that signal \bar{A} or \bar{B} is more likely to be seen when action A or, respectively, B is taken. Thus, \bar{A} is called the prime

signal of action A and \bar{B} is called the prime signal of action B . For simplicity, the correlation formula (2) is assumed to be fixed over time and across players.

Since players' actions are unobservable here, in the following two subsections we propose using the simple and the MLE processes as two possible learning mechanisms to extract information about players' current plays from signals.

2.1. The simple learning process

Let $Y_t = (Y_{1t}, Y_{2t}, \dots, Y_{Nt}) \in \{\bar{A}, \bar{B}\}^N$ be the revealed signal of all players at time t , and let $n(\bar{A}, Y_t) = |\{1 \leq j \leq N : Y_{j,t} = \bar{A}\}|$ be the number of signal \bar{A} in Y_t . Since \bar{A} and \bar{B} are respectively the prime signal for action A and B , our simple learning process assumes that players will naively treat the signals \bar{A} and \bar{B} as the actions A and B , respectively. Because we are mainly interested in a large population size N , $n(\bar{A}, Y_t)/N$ is used, instead of $|\{1 \leq j \leq N, j \neq i : Y_{j,t} = \bar{A}\}|/(N - 1)$, as the probability that player i will match with a player of action A when he is randomly paired with the other players in action updating. Therefore, the expected payoff of player i at time $(t + 1)$ is

$$u_i(A, Y_t) = \frac{n(\bar{A}, Y_t)}{N}a + \left(1 - \frac{n(\bar{A}, Y_t)}{N}\right)b \quad \text{if player } i \text{ takes action } A,$$

$$u_i(B, Y_t) = \frac{n(\bar{A}, Y_t)}{N}c + \left(1 - \frac{n(\bar{A}, Y_t)}{N}\right)d \quad \text{if player } i \text{ takes action } B.$$

A simple calculation shows that

$$u_i(A, Y_t) \geq u_i(B, Y_t) \quad \text{if and only if} \quad (a - c + d - b)n(\bar{A}, Y_t) \geq (d - b)N. \quad (3)$$

By maximizing one's expected payoff,

$$\text{player } i \text{ will choose action } A \text{ at time } t + 1 \text{ if and only if } u_i(A, Y_t) \geq u_i(B, Y_t). \quad (4)$$

Since $a - c > d - b > 0$ by assumption, (3) and (4) easily imply that

$$\text{player } i \text{ will choose action } A \text{ at time } t + 1 \text{ if and only if } n(\bar{A}, Y_t) \geq Nq^*. \quad (5)$$

In other words, player i will choose action A for the next period if and only if the number of signal \bar{A} observed at present is no less than Nq^* . It is worth mentioning that here no mutation or the possibility of experimentating with new strategies is considered after players have chosen their rational actions.

Since the revealed signals are public information and all players adopt the same decision rule to simultaneously update their next actions, it is clear from (5) that players would take an identical action at each period. As a result, the evolutionary process above can be represented by a time-homogeneous Markov chain $\{X_t : t \geq 1\}$ on the state space $\{\mathbf{A}, \mathbf{B}\}$, where $\mathbf{A} = (A, A, \dots, A)$ and $\mathbf{B} = (B, B, \dots, B)$ are the all- A and all- B action profiles, respectively. Moreover, its transition matrix

$$T_N = \begin{bmatrix} p_{AA}(N) & p_{AB}(N) \\ p_{BA}(N) & p_{BB}(N) \end{bmatrix}, \quad (6)$$

where $p_{x,y}(N) = P(X_{t+1} = y \mid X_t = x)$ for $x, y \in \{\mathbf{A}, \mathbf{B}\}$, satisfies

$$p_{AA}(N) = P(n(\bar{A}, Y_t) \geq Nq^* \mid X_t = \mathbf{A}) = \sum_{k \geq Nq^*} u^k (1 - u)^{N-k} C_k^N \quad (7)$$

and

$$p_{BA}(N) = P(n(\bar{A}, Y_t) \geq Nq^* \mid X_t = B) = \sum_{k \geq Nq^*} (1 - v)^k v^{N-k} C_k^N. \tag{8}$$

Here $C_k^N = N! / (k!(N - k)!)$. Note that $P(n(\bar{A}, Y_t) = k \mid X_t = A) = u^k (1 - u)^{N-k} C_k^N$, as players reveal their signals independently according to (2). The ergodic theorem for Markov chains implies that, independent of the initial distribution, the following ergodic distribution exists:

$$(p_N(A), p_N(B)) = \lim_{t \rightarrow \infty} (P(X_t = A), P(X_t = B))$$

with $(p_N(A), p_N(B))T_N = (p_N(A), p_N(B))$. Using $p_N(A) + p_N(B) = 1$, we easily obtain

$$p_N(A) = \frac{p_{BA}(N)}{p_{AB}(N) + p_{BA}(N)} \quad \text{and} \quad p_N(B) = \frac{p_{AB}(N)}{p_{AB}(N) + p_{BA}(N)}. \tag{9}$$

Remember that we are interested in $(p_N(A), p_N(B))$ for a large population. For instance, although $\min(p_N(A), p_N(B)) > 0$ for any fixed population size N by the ergodic theorem, we wonder when $\liminf_{N \rightarrow \infty} \min(p_N(A), p_N(B)) > 0$, meaning that players will visit both equilibria with positive probability.

Taking advantage of the exact form (9), we now study the asymptotic behavior of $(p_N(A), p_N(B))$. By (9),

$$p_N(A) + p_N(B) = 1 \quad \text{and} \quad \frac{p_N(A)}{p_N(B)} = \frac{p_{BA}(N)}{p_{AB}(N)}. \tag{10}$$

Hence, it is enough to study the ratio of $p_{BA}(N)$ and $p_{AB}(N)$.

Introduce independent, identically distributed (i.i.d.) Bernoulli random variables $\{Y_i : i \geq 1\}$ and $\{Z_i : i \geq 1\}$ with

$$\begin{aligned} P(Y_1 = 1) &= u, & P(Y_1 = 0) &= 1 - u, \\ P(Z_1 = 1) &= 1 - v, & P(Z_1 = 0) &= v. \end{aligned} \tag{11}$$

Then (7) and (8) can be expressed as

$$p_{AA}(N) = P\left(\sum_{i=1}^N \frac{Y_i}{N} \geq q^*\right) \quad \text{and} \quad p_{BA}(N) = P\left(\sum_{i=1}^N \frac{Z_i}{N} \geq q^*\right). \tag{12}$$

By (1) and (2), $u > \frac{1}{2} > q^*$. Then (11), (12), and the LLN imply that

$$\lim_N p_{AB}(N) = 1 - \lim_N p_{AA}(N) = 0. \tag{13}$$

Similarly,

$$\lim_N p_{BA}(N) = 1 \text{ or } 0 \text{ depending on whether } 1 - v > q^* \text{ or } 1 - v < q^*, \text{ respectively.} \tag{14}$$

For the case in which $1 - v = q^*$, we may apply the central limit theorem (CLT) to obtain

$$p_{BA}(N) = P\left(\sum_{i=1}^N \frac{Z_i - (1 - v)}{N} \geq 0\right) = P\left(\sum_{i=1}^N \frac{Z_i - (1 - v)}{\sqrt{Nv(1 - v)}} \geq 0\right) \rightarrow \frac{1}{2} \quad \text{as } N \rightarrow \infty. \tag{15}$$

Combining (10) and (13)–(15), we can prove Theorem 1(a), below.

Theorem 1. Assume that $a - c > d - b > 0$ and that the simple learning algorithm (4) holds.

- (a) If $(1 - v) \geq q^*$ then $\lim_{N \rightarrow \infty} p_N(\mathbf{A}) = 1$.
- (b) Suppose that $(1 - v) < q^*$. Let $f(x) = q^* \log x + (1 - q^*) \log(1 - x)$. Then
 - (i) $\lim_{N \rightarrow \infty} p_N(\mathbf{A}) = 1$ if $f(1 - v) > f(u)$, which holds in particular for $u \geq v$;
 - (ii) $\lim_{N \rightarrow \infty} p_N(\mathbf{B}) = 1$ if $f(1 - v) < f(u)$; and
 - (iii) $\liminf_N \min(p_N(\mathbf{A}), p_N(\mathbf{B})) > 0$ if $f(1 - v) = f(u)$.

Since $(1 - v) < q^*$ in Theorem 1(b), $\lim_N p_{AB}(N) = \lim_N p_{BA}(N) = 0$ by (13) and (14). In view of (10), we have to compare the convergence rate to 0 of $p_{BA}(N)$ with that of $p_{AB}(N)$. Chernoff’s theorem of large deviations (see, e.g. Billingsley (1995, p. 151)) can be applied to show that, for large N ,

$$p_{AB}(N) = P\left(\sum_{i=1}^N \frac{Y_i}{N} < q^*\right) = P\left(\sum_{i=1}^N \frac{q^* - Y_i}{N} > 0\right) \approx \rho_1^N, \tag{16}$$

where $\rho_1 = \inf_t E \exp(t(q^* - Y_1)) = \inf_t [ue^{t(q^*-1)} + (1 - u)e^{tq^*}]$. Simple calculus yields the infimum at $t^* = \log(u(1 - q^*)/(1 - u)q^*)$ and, thus, $\log \rho_1 = f(u) - f(q^*)$. Similarly,

$$p_{BA}(N) = P\left(\sum_{i=1}^N \frac{Z_i}{N} \geq q^*\right) = P\left(\sum_{i=1}^N \frac{Z_i - q^*}{N} \geq 0\right) \approx \rho_2^N, \tag{17}$$

where $\rho_2 = \inf_t E \exp(t(Z_1 - q^*))$ and satisfies $\log \rho_2 = f(1 - v) - f(q^*)$. Hence,

$$\log \frac{\rho_2}{\rho_1} = \log \rho_2 - \log \rho_1 = f(1 - v) - f(u). \tag{18}$$

Theorem 1(b)(i) and (ii) follow easily from (10) and (16)–(18), except that we still need to check that $f(1 - v) > f(u)$ when $u \geq v > \frac{1}{2}$. Since $f'(x) = q^*/x - (1 - q^*)/(1 - x) = (q^* - x)/x(1 - x) < 0$ for $x > q^*$, $f(x)$ is decreasing on the interval $[q^*, 1]$. Then, for $u \geq v > \frac{1}{2} > q^*$,

$$\begin{aligned} f(1 - v) - f(u) &\geq f(1 - v) - f(v) \\ &= q^* \log\left(\frac{1 - v}{v}\right) + (1 - q^*) \log\left(\frac{v}{1 - v}\right) \\ &= (1 - 2q^*) \log\left(\frac{v}{1 - v}\right) \\ &> 0. \end{aligned}$$

The critical case, $f(1 - v) = f(u)$, is most subtle, as the previous argument is no longer decisive. Using brute-force computations, we will show in Section 5 that

$$\lim_N \frac{p_{AB}(N)}{p_{BA}(N)} \left(\frac{(1 - u)(1 - v)}{uv}\right)^{\{Nq^*\}} = \frac{(1 - u)(q^* + v - 1)}{v(u - q^*)} > 0, \tag{19}$$

where $\{x\} = [x] - x$ lies in $[0, 1)$ and $[x]$ is the least integer greater than or equal to x . Then Theorem 1(b)(iii) follows easily from (10).

In summary, Theorem 1 shows that players' long-run behavior is determined by q^* in (1) and the correlation between signals and their underlying actions, (2). Depending on the conditions specified in Theorem 1, the risk-dominant equilibrium, the non-risk-dominant equilibrium, or coexistence of both equilibria could emerge in the long run when the population size N is large enough.

As a side result, this model can be compared with those under perfect monitoring. In our setup two nonprime signals reflecting different strategies can be regarded as two kinds of mutation. The noises of the signals, measured by $(1 - u) = \varepsilon^\alpha$ and $(1 - v) = \varepsilon^\eta$, are then interpreted as the associated mutation rates. Here, ε , α , and η are positive constants. Being the mutation rate, we may assume that ε is small. By (12),

$$p_{AB}(N) = P\left(\sum_{i=1}^N Y_i < \lceil Nq^* \rceil\right) \approx (1 - u)^{N - \lceil Nq^* \rceil + 1} = \varepsilon^{\alpha(N - \lceil Nq^* \rceil + 1)}.$$

Similarly,

$$p_{BA}(N) = P\left(\sum_{i=1}^N Z_i \geq \lceil Nq^* \rceil\right) \approx (1 - v)^{\lceil Nq^* \rceil} = \varepsilon^{\eta \lceil Nq^* \rceil}.$$

Using $\lceil Nq^* \rceil = Nq^* + \{Nq^*\}$ and (10),

$$\frac{p_N(\mathbf{A})}{p_N(\mathbf{B})} \approx \frac{\varepsilon^{\eta \lceil Nq^* \rceil}}{\varepsilon^{\alpha(N - \lceil Nq^* \rceil + 1)}} = \varepsilon^{-N(\alpha(1 - q^*) - \eta q^*) + (\{Nq^*\}(\eta + \alpha) - \alpha)}. \tag{20}$$

For the case in which $\alpha = \eta q^*/(1 - q^*)$, the above formula is simplified as

$$\frac{p_N(\mathbf{A})}{p_N(\mathbf{B})} \approx \varepsilon^{\{Nq^*\}(\eta + \alpha)}. \tag{21}$$

Since $\{Nq^*\} \in [0, 1)$, the following can be easily verified using (10), (20), and (21).

Proposition 1. *Let $1 - u = \varepsilon^\alpha$ and $1 - v = \varepsilon^\eta$ be the mutation rates of the actions A and B, respectively, where ε , α , and η are positive constants.*

- (a) *For large N , $\lim_{\varepsilon \rightarrow 0} p_N(\mathbf{A}) = 1$ if $\alpha > \eta q^*/(1 - q^*)$, and $\lim_{\varepsilon \rightarrow 0} p_N(\mathbf{A}) = 0$ if $\alpha < \eta q^*/(1 - q^*)$.*
- (b) *For the case in which $\alpha = \eta q^*/(1 - q^*)$, we have $p_N(\mathbf{A}) \rightarrow 0, 1$, or neither as $\varepsilon \rightarrow 0$, depending on whether $\{Nq^*\}(\eta + \alpha) - \alpha$ is positive, negative, or 0, respectively.*

Except for the critical case, $\alpha = \eta q^*/(1 - q^*)$, Proposition 1 is a special case in Bergin and Lipman (1996), where the Darwinian-type dynamics were adopted on the state space $\{A, B\}^N$. Kandori *et al.* (1993) studied the same dynamics, but with $1 - u = 1 - v$ (that is, with equal mutation rates).

Generally speaking, the long-run equilibrium would depend on the relative depths of the basins of attraction at the two equilibria. Here, the depths of the basins of attraction at the risk-dominant and the non-risk-dominant equilibria are proportional to the values of α and η , respectively. The larger α is, the more difficult it is to jump out from the risk-dominant equilibrium; hence, the deeper the basin of attraction is at the risk-dominant equilibrium. A similar remark holds for η . It is no surprise that $\lim_{\varepsilon \rightarrow 0} p_N(\mathbf{A}) = 1$ or 0 if α is large enough or small enough, respectively, as shown in Proposition 1(a). However, it may not be easy to see that $\alpha = \eta q^*/(1 - q^*)$ is the critical value and that in this case the ergodic distribution depends on $\{Nq^*\}(\eta + \alpha) - \alpha$. Note that (21) holds for any finite N .

2.2. The MLE learning process

In this subsection we explore players' long-run behavior under the MLE learning algorithm. Let $Y_t = (Y_{1t}, Y_{2t}, \dots, Y_{Nt}) \in \{A, B\}^N$ be the revealed signal of all players at time t and let $y_t = (y_{1t}, y_{2t}, \dots, y_{Nt})$ be its realization. Since the revealed signals are public information and all players adopt the same decision rule to simultaneously update their next actions, the process can be represented as in Section 2.1 by a time-homogeneous Markov process $\{\tilde{X}_t : t \geq 1\}$ with state space $\{A, B\}$. Players are presumed to compare $P(Y_t = y_t \mid \tilde{X}_t = A)$ with $P(Y_t = y_t \mid \tilde{X}_t = B)$. If

$$\frac{P(Y_t = y_t \mid \tilde{X}_t = A)}{P(Y_t = y_t \mid \tilde{X}_t = A) + P(Y_t = y_t \mid \tilde{X}_t = B)} \geq \beta,$$

where $\beta \in (0, 1)$ is a preset threshold, players are hypothesized to guess $\tilde{X}_t = A$ and, thus, adopt action A at time $t + 1$. Otherwise, players would adopt action B . It is easy to see from (2) that

$$P(Y_t = y_t \mid \tilde{X}_t = A) = u^k(1-u)^{N-k} \quad \text{and} \quad P(Y_t = y_t \mid \tilde{X}_t = B) = (1-v)^k v^{N-k}, \quad (22)$$

where $k = n(\bar{A}, y_t) := |\{1 \leq j \leq N : y_{j,t} = \bar{A}\}|$ is the number of signal \bar{A} in y_t . Thus,

$$\tilde{X}_{t+1} = A \quad \text{if and only if} \quad \frac{P(Y_t = y_t \mid \tilde{X}_t = A)}{P(Y_t = y_t \mid \tilde{X}_t = B)} \geq \frac{\beta}{(1-\beta)}. \quad (23)$$

By taking the logarithm, a simple calculation shows that

$$\begin{aligned} \tilde{X}_{t+1} = A \quad \text{if and only if} \quad n(\bar{A}, Y_t) &\geq N \frac{\log(v/(1-u))}{\log(uv/(1-u)(1-v))} \\ &+ \frac{\log(\beta/(1-\beta))}{\log(uv/(1-u)(1-v))} \\ &\equiv N\gamma^* + \delta. \end{aligned} \quad (24)$$

By Bayes rule,

$$P(\tilde{X}_{t+1} = A) = P(\tilde{X}_t = A) P(\tilde{X}_{t+1} = A \mid \tilde{X}_t = A) + P(\tilde{X}_t = B) P(\tilde{X}_{t+1} = A \mid \tilde{X}_t = B).$$

Then, from (22) and (24), we obtain

$$\begin{aligned} P(\tilde{X}_{t+1} = A) &= P(\tilde{X}_t = A) P(n(\bar{A}, Y_t) \geq N\gamma^* + \delta \mid \tilde{X}_t = A) \\ &+ P(\tilde{X}_t = B) P(n(\bar{A}, Y_t) \geq N\gamma^* + \delta \mid \tilde{X}_t = B) \\ &= P(\tilde{X}_t = A) P\left(\sum_{i=1}^N Y_i \geq N\gamma^* + \delta\right) + P(\tilde{X}_t = B) P\left(\sum_{i=1}^N Z_i \geq N\gamma^* + \delta\right), \end{aligned}$$

where $\{Y_i, Z_i\}$ are the i.i.d. Bernoulli random variables given in (11). A similar expression can be obtained for $P(\tilde{X}_{t+1} = B)$. As a consequence, the transition matrix of $\{\tilde{X}_t : t \geq 1\}$ is similar to (6) except that now

$$p_{AA}(N) = P\left(\sum_{i=1}^N Y_i \geq N\gamma^* + \delta\right) \quad \text{and} \quad p_{BA}(N) = P\left(\sum_{i=1}^N Z_i \geq N\gamma^* + \delta\right). \quad (25)$$

The method used in Section 2.1 can be repeated to find its ergodic distribution:

$$(\tilde{p}_N(\mathbf{A}), \tilde{p}_N(\mathbf{B})) := \lim_{t \rightarrow \infty} (\mathbb{P}(\tilde{X}_t = \mathbf{A}), \mathbb{P}(\tilde{X}_t = \mathbf{B})).$$

Since we still have the following analog to (10):

$$\tilde{p}_N(\mathbf{A}) + \tilde{p}_N(\mathbf{B}) = 1 \quad \text{and} \quad \frac{\tilde{p}_N(\mathbf{A})}{\tilde{p}_N(\mathbf{B})} = \frac{p_{\mathbf{B}\mathbf{A}}(N)}{p_{\mathbf{A}\mathbf{B}}(N)}, \tag{26}$$

it suffices to study the ratio of $p_{\mathbf{B}\mathbf{A}}(N)$ and $p_{\mathbf{A}\mathbf{B}}(N)$. It will be shown in Section 5 that

$$1 - v < \gamma^* < u \quad \text{for } \frac{1}{2} < u, v < 1, \tag{27}$$

where γ^* is given in (24). As $(N\gamma^* + \delta)/N \approx \gamma^*$ for large N , the LLN implies that $\lim_N p_{\mathbf{A}\mathbf{B}}(N) = \lim_N p_{\mathbf{B}\mathbf{A}}(N) = 0$. Then Chernoff’s theorem can be applied to show that, for large N ,

$$p_{\mathbf{A}\mathbf{B}}(N) \approx \tilde{\rho}_1^N \quad \text{and} \quad p_{\mathbf{B}\mathbf{A}}(N) \approx \tilde{\rho}_2^N, \tag{28}$$

where $\tilde{\rho}_1 = \inf_t \mathbb{E} \exp(t(\gamma^* - Y_1))$ and $\tilde{\rho}_2 = \inf_t \mathbb{E} \exp(t(Z_1 - \gamma^*))$. A similar calculation to that used in (16) and (17) shows that

$$\log \tilde{\rho}_1 = \gamma^* \log\left(\frac{u}{\gamma^*}\right) + (1 - \gamma^*) \log\left(\frac{1 - u}{1 - \gamma^*}\right)$$

and

$$\log \tilde{\rho}_2 = \gamma^* \log\left(\frac{1 - v}{\gamma^*}\right) + (1 - \gamma^*) \log\left(\frac{v}{1 - \gamma^*}\right).$$

Since

$$\gamma^* = \log \frac{v/(1 - u)}{\log(uv/(1 - u)(1 - v))},$$

$$1 - \gamma^* = \log \frac{u/(1 - v)}{\log(uv/(1 - u)(1 - v))}$$

and then

$$\log \tilde{\rho}_2 - \log \tilde{\rho}_1 = (1 - \gamma^*) \log\left(\frac{v}{1 - u}\right) - \gamma^* \log\left(\frac{u}{1 - v}\right) = 0.$$

Hence, $\tilde{\rho}_2 = \tilde{\rho}_1$ and the order estimates in (28) are not good enough. We have to estimate $p_{\mathbf{A}\mathbf{B}}(N)$ and $p_{\mathbf{B}\mathbf{A}}(N)$ directly as we did in Theorem 1(b)(iii). Similar to (19), we will show in Section 5 that

$$\lim_N \frac{p_{\mathbf{A}\mathbf{B}}(N)}{p_{\mathbf{B}\mathbf{A}}(N)} \left(\frac{(1 - u)(1 - v)}{uv}\right)^{\delta + \{N\gamma^* + \delta\}} = \frac{(1 - u)(\gamma^* + v - 1)}{v(u - \gamma^*)} > 0. \tag{29}$$

Note that the constants γ^* and δ are defined in (24). In view of (26), we obtain the following theorem.

Theorem 2. For the MLE learning algorithm (23), $\liminf_N \min(\tilde{p}_N(\mathbf{A}), \tilde{p}_N(\mathbf{B})) > 0$.

In view of (23) and (24), players’ optimal behavior is determined by whether the number of \bar{A} in the revealed signal exceeds $N\gamma^* + \delta$, which depends on u, v , and β but not on the payoff parameters a, b, c , and d . Thus, Theorem 2 remains true for other 2×2 symmetric games as long as players adopt the MLE learning algorithm.

Theorem 2 demonstrates that players will visit the risk-dominant and the non-risk-dominant equilibria with positive probability in the long run under the MLE learning process, rather than stick to one equilibrium. Because the exponent $\{N\gamma^* + \delta\}$ in (29) lies in $[0, 1)$ and varies with N by (27), $\lim_N(\tilde{p}_N(\mathbf{A}), \tilde{p}_N(\mathbf{B}))$ does not exist. Hence, the probability for players to visit the two equilibria in the long run is unlikely to equal the mixed Nash equilibrium, $(q^*, 1 - q^*)$, of the coordination game.

Compared with Theorem 1, we discover that players’ long-run behavior under imperfect monitoring is fairly sensitive to the information-extracting mechanism as N tends to ∞ .

Finally, these results can be compared with those under perfect monitoring. As in Proposition 1, we may regard the nonprime signals reflecting different strategies as two kinds of mutation with rates $(1 - u)$ and $(1 - v)$. Then it can be shown that players’ long-run behavior depends on how fast both mutation rates, $1 - u$ and $1 - v$, tend to 0. Since its proof and interpretation are the same as those of Proposition 1, they are omitted here.

3. Other symmetric 2×2 games

In addition to coordination games there are two other generic types of 2×2 symmetric games. They are games with a dominant strategy and games with a unique mixed Nash equilibrium. We analyze them below.

As in Section 2.1, the simultaneous simple learning process for any of these two games is still a Markov chain on the state space $\{\mathbf{A}, \mathbf{B}\}$. Without causing any confusion, we still denote by $\{X_t : t \geq 1\}$ the associated dynamic process.

The payoff for games with a strictly dominant equilibrium is as shown in Figure 1, but with $b > d > a > c$. A typical example is the prisoners’ dilemma game. Under the present assumption, $(a - c + d - b)$ could be negative. For the case in which $a - c \leq b - d$, the inequality in (3) holds as $0 \geq a - c + d - b > d - b$ and $0 \leq n(\bar{A}, \mathbf{Y}_t) \leq N$. For the other case, $a - c > b - d$, the inequality in (3) holds trivially as $a - c + d - b > 0 > d - b$. All together, (4) implies that $X_t = \mathbf{A}$ for $t \geq 1$ no matter what the initial value of X_0 . This can be easily understood in view of the payoff matrix in Figure 1. Since $a > c$ and $b > d$, any player will choose action A for a better payoff, no matter what action his opponent will take.

Theorem 3. For 2×2 symmetric games, as depicted in Figure 1, with $b > d > a > c$, we have $X_t = \mathbf{A}$ for $t \geq 1$ under the simple learning algorithm (4). Hence, \mathbf{A} is the unique long-run equilibrium for a population size $N \geq 2$.

The payoff structure for games with a unique mixed Nash equilibrium is as shown in Figure 1, but with $c > a$ and $b > d$. Then (A, B) , (B, A) , and the mixed strategy $(q^*, 1 - q^*)$ are the Nash equilibria, where $q^* \in (0, 1)$ is defined in (1) but the condition that $q^* < \frac{1}{2}$ is not required here. Note that $a - c < 0$ and $d - b < 0$ now. For the simple learning algorithm, from (3) and (4), we find that, for $1 \leq i \leq N$,

$$\text{player } i \text{ will choose action } A \text{ at time } t + 1 \text{ if and only if } n(\bar{A}, \mathbf{Y}_t) \leq Nq^*.$$

Note that the inequality in (5) is reversed owing to $a - c + d - b < 0$ and $d - b < 0$. Then its transition probabilities satisfy

$$p_{AB}(N) = P(n(\bar{A}, \mathbf{Y}_t) > Nq^* \mid X_t = \mathbf{A}) = P\left(\sum_{i=1}^N \frac{Y_i}{N} > q^*\right) \tag{30}$$

and

$$p_{BA}(N) = P(n(\bar{A}, \mathbf{Y}_t) \leq Nq^* \mid X_t = \mathbf{B}) = P\left(\sum_{i=1}^N \frac{Z_i}{N} \leq q^*\right), \tag{31}$$

where $\{Y_i, Z_i\}$ are the i.i.d. Bernoulli random variables given in (11).

Let $(p_N(\mathbf{A}), p_N(\mathbf{B}))$ be the ergodic distribution of $\{X_t : t \geq 1\}$. Then (9) and (10) still hold since (30) and (31) look like the transition probabilities in (12). By repeating the arguments in Theorem 1, the following result is proved in Section 5.

Theorem 4. *Consider the 2×2 symmetric games of Figure 1 with $c > a$ and $b > d$ under the simple learning algorithm (4).*

- (a) *If $u > q^*$ and $1 - v > q^*$ then $\lim_{N \rightarrow \infty} p_N(\mathbf{A}) = 0$.*
- (b) *If $u = q^* = 1 - v$ or $u > q^* > 1 - v$ then $\lim_{N \rightarrow \infty} p_N(\mathbf{A}) = \lim_{N \rightarrow \infty} p_N(\mathbf{B}) = \frac{1}{2}$.*
- (c) *Depending on whether $u > q^* = 1 - v$ or $u = q^* > 1 - v$, $\lim_{N \rightarrow \infty} p_N(\mathbf{A}) = \frac{1}{3}$ or $\frac{2}{3}$, respectively.*
- (d) *If $1 - u \leq q^*$ and $v \leq q^*$ then $\lim_{N \rightarrow \infty} p_N(\mathbf{A}) = 1$.*
- (e) *Finally, for the case in which $u < q^* < 1 - v$, let $f(x) = q^* \log x + (1 - q^*) \log(1 - x)$. Then*
 - (i) *$\lim_{N \rightarrow \infty} p_N(\mathbf{A}) = 1$ or 0 depending on whether $f(1 - v) > f(u)$ or $f(1 - v) < f(u)$, respectively;*
 - (ii) *$\liminf_N \min(p_N(\mathbf{A}), p_N(\mathbf{B})) > 0$ if $f(1 - v) = f(u)$.*

Theorem 4 shows that players’ long-run behavior depends not only on the relative values of $1 - v$ and q^* , but also on the relative magnitudes of u and q^* . This is because the condition that $q^* < \frac{1}{2}$ is not required, as in Section 2.1. In the simple learning process, a players’ long-run behavior is mainly determined by signals’ occurring frequencies. Thus, it is not a surprise to have \mathbf{A} or \mathbf{B} alone in the long run even though they are not Nash equilibria. Theorem 4(b) and (c) indicate that coexistence of both \mathbf{A} or \mathbf{B} is probable. The unique symmetric mixed Nash equilibrium could possibly survive only when $u < q^* < 1 - v$ and $f(1 - u) = f(v)$.

For the MLE learning process, Theorem 2 still holds as remarked in the paragraph after Theorem 2.

4. Conclusions

In this paper we have analyzed players’ long-run behavior in evolutionary coordination games with imperfect monitoring in a large population. Players can only observe public signals relating to other players’ underlying actions, and can employ the simple or the MLE learning process to form expectations about their opponents’ current plays based on the signals. In our setup, random signals are like mutations in the perfect monitoring literature, which make transitions

between equilibria possible. However, unlike perfect-monitoring models with state- and time-invariant mutations, the random signals setup may not be able to refine Nash equilibria. For instance, Theorem 1 shows that in the simple learning process the risk-dominant equilibrium, the non-risk-dominant equilibrium, or coexistence of both equilibria could appear in the long run. Which equilibrium will survive in the long run would depend on the correlation degrees between signals and their underlying actions. In contrast, Theorem 2 shows that coexistence of the two equilibria is the unique long-run limit in the MLE learning process. All together, it suggests that players' means to extract information from the observed signals has a critical impact on their long-run behavior in evolutionary games with imperfect monitoring. Theorems 3 and 4 show that similar conclusions can be extended to other 2×2 symmetric games with imperfect monitoring.

In our model players are assumed to update their actions simultaneously at each time period. This presumption makes our dynamics a simple two-state Markov chain with state space $\{A, B\}$. The corresponding ergodic distribution can then be explicitly specified as in (10). This is the main reason why players' long-run behavior in a large population can be thoroughly determined, as in Theorems 1–4 of this paper.

There are two other updating schemes. In a sequential or random agent updating scheme (see, e.g. Blume (1995) and Miekisz (2005)), only one player is allowed to update his action per period. The other scheme lets all players have the same positive revision probability at each time period (see, e.g. Vega-Redondo (1997), Eshel *et al.* (1998), and Miekisz (2005)). Under these two updating schemes, the associated dynamic systems become a Markov chain with $N + 1$ and 2^N states, respectively. In either case it is much more complicated than ours. As shown by Eshel *et al.* (1998) and Miekisz (2005), players' long-run behavior could change accordingly. It would be interesting to see in the future how players' long-run behavior changes when these two updating schemes are adopted in our models.

5. Proofs of Theorems 1, 2, and 4

5.1. Proof of Theorem 1

It remains to verify (19) for part (b)(iii). By (16),

$$p_{AB}(N) = P\left(\sum_{i=1}^N Y_i < Nq^*\right) = P\left(\sum_{i=1}^N Y_i < \lceil Nq^* \rceil\right) = \sum_{l=1}^{\lceil Nq^* \rceil} P\left(\sum_{i=1}^N Y_i = \lceil Nq^* \rceil - l\right). \tag{32}$$

Using $\lceil Nq^* \rceil = Nq^* + \{Nq^*\}$,

$$\begin{aligned} & P\left(\sum_{i=1}^N Y_i = \lceil Nq^* \rceil - l\right) \\ &= C_{\lceil Nq^* \rceil - l}^N u^{\lceil Nq^* \rceil - l} (1 - u)^{N - \lceil Nq^* \rceil + l} \\ &= C_{\lceil Nq^* \rceil}^N (u^{q^*} (1 - u)^{1 - q^*})^N \left(\frac{u}{1 - u}\right)^{\{Nq^*\}} \left(\prod_{j=1}^l \frac{\lceil Nq^* \rceil - j + 1}{N - \lceil Nq^* \rceil + j}\right) \left(\frac{1 - u}{u}\right)^l. \end{aligned} \tag{33}$$

Since $\lceil Nq^* \rceil - j \leq Nq^*$ for $j \geq 1$, we find that, for each fixed $l \geq 1$,

$$\prod_{j=1}^l \frac{\lceil Nq^* \rceil - j + 1}{N - \lceil Nq^* \rceil + j}$$

converges to $(q^*/(1 - q^*))^l$ and is bounded by $(\lceil Nq^* \rceil / Nq^*)(q^*/(1 - q^*))^l$. It follows from Lebesgue’s bounded convergence theorem that

$$\lim_N \frac{p_{AB}(N)}{C_{\lceil Nq^* \rceil}^N (u^{q^*} (1 - u)^{1 - q^*})^N} \left(\frac{1 - u}{u}\right)^{\lceil Nq^* \rceil} = \sum_{l=1}^{\infty} \left(\frac{q^*(1 - u)}{(1 - q^*)u}\right)^l. \tag{34}$$

Similarly, from (12) we have

$$\begin{aligned} p_{BA}(N) &= P\left(\sum_{i=1}^N Z_i \geq Nq^*\right) \\ &= P\left(\sum_{i=1}^N Z_i \geq \lceil Nq^* \rceil\right) \\ &= \sum_{l=0}^{N - \lceil Nq^* \rceil} P\left(\sum_{i=1}^N Z_i = \lceil Nq^* \rceil + l\right) \\ &= \sum_{l=0}^{N - \lceil Nq^* \rceil} C_{\lceil Nq^* \rceil + l}^N (1 - v)^{\lceil Nq^* \rceil + l} v^{N - \lceil Nq^* \rceil - l} \\ &= \sum_{l=0}^{N - \lceil Nq^* \rceil} C_{\lceil Nq^* \rceil}^N ((1 - v)^{q^*} v^{1 - q^*})^N \left(\frac{1 - v}{v}\right)^{\lceil Nq^* \rceil} \\ &\quad \times \left(\prod_{j=1}^l \frac{N - \lceil Nq^* \rceil - j + 1}{\lceil Nq^* \rceil + j}\right) \left(\frac{1 - v}{v}\right)^l, \end{aligned} \tag{35}$$

and, thus,

$$\lim_N \frac{p_{BA}(N)}{C_{\lceil Nq^* \rceil}^N ((1 - v)^{q^*} v^{1 - q^*})^N} \left(\frac{v}{1 - v}\right)^{\lceil Nq^* \rceil} = \sum_{l=0}^{\infty} \left(\frac{(1 - q^*)(1 - v)}{q^*v}\right)^l. \tag{36}$$

Since $f(1 - v) = f(u)$ by assumption, $u^{q^*} (1 - u)^{1 - q^*} = (1 - v)^{q^*} v^{1 - q^*}$. Using $\sum_{l=m}^{\infty} x^l = x^m / (1 - x)$ for $0 < x < 1$, then (19) follows by taking the ratio of (34) and (35).

5.2. Proof of Theorem 2

It remains to verify (27) and (29). The first one is easy. Since

$$\gamma^* = \frac{\log(v/(1 - u))}{\log(uv/(1 - u)(1 - v))},$$

as given in (24), we have $1 - \gamma^* = \log(u/(1 - v)) / \log(uv/(1 - u)(1 - v))$. By symmetry, it is enough to prove the first inequality in (27). A simple calculation shows that this is equivalent to

$$g(1 - v) - g(u) > 0, \tag{37}$$

where $g(x) = v \log(1 - x) + (1 - v) \log x$ for $x \in (0, 1)$. By differentiation, we have $g'(x) = -v/(1 - x) + (1 - v)/x = (1 - v - x)/x(1 - x)$, which is positive on $(0, 1 - v)$ and

negative on $(1 - v, 1)$. Hence, $g(x)$ attains its unique maximum at $x = 1 - v$. This verifies (37) and, thus, (27).

Now we deal with (29), which is similar to (19) and will thus be proved by the same method. By (25),

$$p_{AB}(N) = P\left(\sum_{i=1}^N Y_i < N\gamma^* + \delta\right) = \sum_{l=1}^{\lceil N\gamma^* + \delta \rceil} P\left(\sum_{i=1}^N Y_i = \lceil N\gamma^* + \delta \rceil - l\right)$$

and

$$p_{BA}(N) = P\left(\sum_{i=1}^N Z_i \geq N\gamma^* + \delta\right) = \sum_{l=0}^{N - \lceil N\gamma^* + \delta \rceil} P\left(\sum_{i=1}^N Z_i = \lceil N\gamma^* + \delta \rceil + l\right),$$

which are similar to (32) and (35), respectively. Replacing Nq^* in (33) and (35) by $N\gamma^* + \delta$, we will obtain in a parallel way the following analogs to (34) and (36):

$$\lim_N \frac{p_{AB}(N)}{C_{\lceil N\gamma^* + \delta \rceil}^N (u\gamma^* (1 - u)^{1 - \gamma^*})^N} \left(\frac{1 - u}{u}\right)^{\delta + \{N\gamma^* + \delta\}} = \sum_{l=1}^{\infty} \left(\frac{\gamma^* (1 - u)}{(1 - \gamma^*)u}\right)^l$$

and

$$\lim_N \frac{p_{BA}(N)}{C_{\lceil N\gamma^* + \delta \rceil}^N ((1 - v)\gamma^* v^{1 - \gamma^*})^N} \left(\frac{v}{1 - v}\right)^{\delta + \{N\gamma^* + \delta\}} = \sum_{l=0}^{\infty} \left(\frac{(1 - \gamma^*)(1 - v)}{\gamma^* v}\right)^l.$$

Taking the ratio of these two formulae, we obtain (29) after a little calculation. This completes the proof.

5.3. Proof of Theorem 4

We follow the same procedure as in Theorem 1. Applying the LLN and CLT to (30),

$$\lim_N p_{AB}(N) = 0, 1, \text{ or } \frac{1}{2} \text{ depending on whether } u < q^*, u > q^*, \text{ or } u = q^*, \text{ respectively.} \tag{38}$$

Note that $E Y_1 = u$ and $E Z_1 = 1 - v$. Similarly, from (31) we obtain

$$\lim_N p_{BA}(N) = 0, 1, \text{ or } \frac{1}{2} \text{ depending on whether } 1 - v > q^*, 1 - v < q^*, \text{ or } 1 - v = q^*, \text{ respectively.} \tag{39}$$

Since (10) remains valid for the present ergodic distribution $(p_N(\mathbf{A}), p_N(\mathbf{B}))$, all the conclusions in parts (a)–(d) follow easily from (38) and (39).

For part (e), $\lim_N p_{AB}(N) = \lim_N p_{BA}(N) = 0$ as $u < q^* < 1 - v$ now. We are obliged to compare their convergence rates to 0 in order to find the limit via (10). As in (16) and (17), Chernoff’s theorem of large deviation implies that

$$p_{AB}(N) \approx \rho_1^N \quad \text{and} \quad p_{BA}(N) \approx \rho_2^N \quad \text{for } u < q^* < 1 - v,$$

where ρ_1 and ρ_2 are the same as in (16) and (17). Hence, (18) still holds. Part (e)(i) follows as a consequence. It remains to treat the critical case, $f(1 - v) = f(u)$.

Rewrite (30) and (31) respectively as

$$p_{AB}(N) = P\left(\sum_{i=1}^N (1 - Y_i) < N(1 - q^*)\right) = P\left(\sum_{i=1}^N (1 - Y_i) < \lceil N(1 - q^*) \rceil\right)$$

and

$$p_{BA}(N) = P\left(\sum_{i=1}^N (1 - Z_i) \geq N(1 - q^*)\right) = P\left(\sum_{i=1}^N (1 - Z_i) \geq \lceil N(1 - q^*) \rceil\right).$$

These equations are the same as (32) and (35), respectively, except that Y_i , Z_i , and q^* there are replaced by $(1 - Y_i)$, $(1 - Z_i)$, and $1 - q^*$. Hence, (34), (36), and then (19) still hold but with u , v , and q^* there replaced by $(1 - u)$, $(1 - v)$, and $1 - q^*$. In particular, (19) becomes

$$\lim_N \frac{p_{AB}(N)}{p_{BA}(N)} \left(\frac{uv}{(1-u)(1-v)} \right)^{\{N(1-q^*)\}} = \frac{u(1-v-q^*)}{(1-v)(q^*-u)} > 0.$$

Part (e)(ii) is verified in view of (10). This completes the proof.

References

- BERGIN, J. AND LIPMAN, B. L. (1996). Evolution with state-dependent mutations. *Econometrica* **64**, 943–956.
- BILLINGSLEY, P. (1995). *Probability and Measure*, 3rd edn. John Wiley, New York.
- BLUME, L. E. (1993). The statistical mechanics of strategic interaction. *Games Econom. Behavior* **5**, 387–424.
- BLUME, L. E. (1995). The statistical mechanics of best-response strategy revision. *Games Econom. Behavior* **11**, 111–145.
- CHEN, H.-C. AND CHOW, Y. (2001). On the convergence of evolution processes with time-varying mutations and local interaction. *J. Appl. Prob.* **38**, 301–323.
- CHEN, H.-C., CHOW, Y. AND HSIEH, J. (2006). Boundedly rational quasi-Bayesian learning in coordination games with imperfect monitoring. *J. Appl. Prob.* **43**, 335–350.
- COMPTE, O. (2002). On failing to cooperate when monitoring is private. *J. Econom. Theory* **102**, 151–188.
- ELLISON, G. (1993). Learning, local interaction, and coordination. *Econometrica* **61**, 1047–1071.
- ESHEL, I., SAMUELSON, L. AND SHAKED, A. (1998). Altruists, egoists, and hooligans in a local interaction model. *Amer. Econom. Rev.* **88**, 157–179.
- FUDENBERG, D., LEVINE, D. AND MASKIN, E. (1994). The folk theorem with imperfect public information. *Econometrica* **62**, 997–1040.
- KANDORI, M. G., MAILATH, G. AND ROB, R. (1993). Learning, mutation, and long-run equilibria in games. *Econometrica* **61**, 29–56.
- MAILATH, G. J. AND MORRIS, S. (2002). Repeated games with almost-public monitoring. *J. Econom. Theory* **102**, 189–228.
- MIEKISZ, J. (2005). Equilibrium selection in evolutionary games with random matching of players. *J. Theoret. Biol.* **232**, 47–53.
- ROBLES, J. (1998). Evolution with changing mutation rates. *J. Econom. Theory* **79**, 207–223.
- ROBSON, A. J. AND VEGA-REDONDO, F. (1996). Efficient equilibrium selection in evolutionary games with random matching. *J. Econom. Theory* **70**, 65–92.
- SEKIGUCHI, T. (1997). Efficiency in the prisoner's dilemma with private monitoring. *J. Econom. Theory* **76**, 345–361.
- VEGA-REDONDO, F. (1995). Expectations, drift, and volatility in evolutionary games. *Games Econom. Behavior* **11**, 391–412.
- VEGA-REDONDO, F. (1997). The evolution of Walrasian behavior. *Econometrica* **65**, 375–384.
- YOUNG, P. (1993). The evolution of conventions. *Econometrica* **61**, 57–84.