BOUNDEDLY RATIONAL QUASI-BAYESIAN LEARNING IN COORDINATION GAMES WITH IMPERFECT MONITORING

HSIAO-CHI CHEN,* *National Taipei University*

YUNSHYONG CHOW,** *** AND

JUNE HSIEH,** **** *Academia Sinica, Taipei*

Abstract

In this paper we study players' long-run behaviors in evolutionary coordination games with imperfect monitoring. In each time period, signals corresponding to players' underlying actions, instead of the actions themselves, are observed. A boundedly rational quasi-Bayesian learning process is proposed to extract information from the realized signals. We find that players' long-run behaviors depend not only on the correlations between actions and signals, but on the initial probabilities of risk-dominant and non-risk-dominant equilibria being chosen. The conditions under which risk-dominant equilibrium, non-risk-dominant equilibrium, and the coexistence of both equilibria emerges in the long run are shown. In some situations, the number of limiting distributions grows unboundedly as the population size grows to infinity.

Keywords: Coordination game; quasi-Bayesian learning; imperfect monitoring

2000 Mathematics Subject Classification: Primary 91A22; 60G99; 60K99; 60F99

1. Introduction

The multiplicity of Nash equilibria greatly lowers the predictive power and potential for application of game theory. For instance, two pure Nash equilibria, risk-dominant and non-risk-dominant, exist in a coordination game. The naturally occurring question is: Which equilibrium will emerge? Thus, the literature on refining Nash equilibria from both static and dynamic viewpoints has developed. The evolutionary learning process is one of the dynamic approaches to refining Nash equilibria. In the pioneering studies of Foster and Young (1990), Kandori *et al.* (1993), Young (1993), and Ellison (1993), it was shown that the risk-dominant equilibrium emerges in the long run in 2×2 coordination games if mutation is added into players' decisions in each time period. In these models, mutation rates are fixed over time and with respect to player. However, players may not coordinate at risk-dominant equilibrium in the long run when mutation rates vary with state (see, e.g. Bergin and Lipman (1996)) or with time (see, e.g. Robles (1998) and Chen and Chow (2001)). Moreover, Robson and Vega-Redondo (1996) showed that players will choose Pareto-efficient rather than risk-dominant

equilibrium under multiple-round matching mechanisms. All the models above assume that players' actions and payoffs are perfectly monitored. However, examples of observable signals relating stochastically to players' unobservable underlying actions are numerous in the real world. For instance, although regulators may not know a firm's exact pollution-abatement activities, pollution emission can be measured and depends on the firm's abatement level and some random factors, such as weather conditions. In principal–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 unknown product qualities even though they are also affected by factors such as stochastic market demands. Thus, in this paper we plan to analyze equilibrium selection under evolutionary learning processes with imperfect monitoring of players' actions and payoffs.

An evolutionary coordination game is considered in Section 2. At the end of each time period, players know only their own actions and payoffs, while signals related to players' underlying actions are revealed publicly. For each action of each player, there are two possible signals: the 'prime' signal is the one more likely to occur. The correlations between actions and signals are assumed to be fixed over time and with respect to player. We propose the following boundedly rational quasi-Bayesian learning process. In each time period, players form their beliefs about their opponents' current plays based on the signals observed during the last period and Bayes' updating rule. Then players choose their actions to maximize their expected current payoffs. Our model differs from the repeated prisoner's dilemma game with imperfect monitoring (see, e.g. Fudenberg *et al.* (1994), Sekiguchi (1997), Compte (2002), and Mailath and Morris (2002)) in that our players are boundedly (myopically) rational instead of longsighted rational. In the latter case, players take actions to maximize the expected discounted average payoffs. In the Bayesian learning process studied by Kalai and Lehrer (1993) and Jordan (1991), longsighted rational players update their beliefs about opponents' strategies based on observed actions and Bayes' rule. Therefore, these kinds of model are also different from ours. With respect to methods of extracting information from observed signals, a relevant study by Chen and Chow (2003) employed both a simple learning algorithm and a maximum likelihood estimation learning process. In the former, players use the ratio of signals observed during the last time period to forecast their opponents' current plays, while, in the latter, the likelihood functions of observed signals are adapted to predict the opponents' current plays.

In Section 3 we explore the limiting distributions of our learning process for large population size. In evolutionary models with perfect monitoring, mutation is the driving force making transitions between equilibria possible, and the relative sizes of the two equilibria's basins of attraction determines the players' long-run choices. However, Chen and Chow (2003) showed that random signals replace the role of mutation in refining Nash equilibria when players' actions and payoffs are imperfectly monitored, and that the correlation between signals and their underlying actions completely determines the players' long-run behaviors. Their learning processes induce two-state (risk-dominant and non-risk-dominant equilibria) Markov chains. Hence, the limiting distributions are fully determined by transition probabilities between the two equilibria, by the ergodic theorem. On the other hand, transition probabilities are related to the frequencies at which signals occur. When the population size is large, signals' frequencies of occurrence approximate the correlation degrees between underlying actions and signals.

In contrast, our learning process is not a Markov chain, so the ergodic theorem does not apply. Instead, we obtain a recursion equation for its limiting distributions. A careful analysis of this equation allows us to find the limiting distributions in almost all cases. The limiting distributions depend not only on the correlations between signals and the underlying actions, but also on the

players' initial choices. In some situations, risk-dominant equilibrium will emerge in the long run unless non-risk-dominant equilibrium is chosen initially. This situation occurs when the correlation between the non-risk-dominant action and its prime signal is small but not less than the correlation between the risk-dominant action and its prime signal. In these circumstances, players are more likely to switch from non-risk-dominant to risk-dominant equilibrium than vice versa. In other situations, risk- or non-risk-dominant equilibrium could emerge in the long run if it is initially chosen with high enough probability. When starting from either equilibrium with any other probability, players will visit both equilibria with positive probability. This visiting probability is determined by the players' initial choices, the population size, and the correlations between underlying actions and signals. Moreover, there will be infinitely many limiting distributions in which both equilibria are visited with positive probability, as the population size grows to infinity.

In Section 4 we show that our results remain valid when mutation of the players' rational choices is considered. Finally, we draw our conclusions in Section 5.

2. The model

There are N players, where N is even and finite. In period t, $t = 0, 1, 2, \ldots$, players are randomly and independently matched in pairs to play the following coordination game once:

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

Here $\{A, B\}$ is the action set for all players and a, b, c, and d are payoffs, with $a > c$, $d > b$, and $a - c > d - b$. Hence,

$$q^* := \frac{d - b}{(a - c) + (d - b)} < \frac{1}{2} \tag{2.1}$$

and action A is risk-dominant. It is assumed that players know only their own actions and payoffs at the end of each time period. Players may infer the actions of their opponents from their own payoffs. However, the actions of the other players remain unknown. Nevertheless, signals corresponding to the players' underlying actions are publicly observable. Denote by $\{\bar{A}, \bar{B}\}$ the signal set for all players. The correlations between actions and their corresponding signals are described by the following conditional probabilities:

$$\mathrm{P}(\bar{A} \mid A) = u \quad \text{and} \quad \mathrm{P}(\bar{B} \mid B) = v, \quad \text{where } \tfrac{1}{2} < u, v < 1.$$

Hence, $\mathrm{P}(\bar{B} \mid A) = 1 - u < \frac{1}{2}$ and $\mathrm{P}(\bar{A} \mid B) = 1 - v < \frac{1}{2}$. This means that signal $\bar{A}$ or $\bar{B}$ is more likely to be seen when action A or, respectively, B is taken. Thus, $\bar{A}$ and $\bar{B}$ are respectively called the prime signals of actions A and B. The correlations between underlying actions and signals are assumed to be fixed over time and with respect to player.

Our boundedly rational quasi-Bayesian learning process is described below. In each time period $t + 1$, $t \geq 0$, players are assumed to use the posterior probabilities of action profiles $\boldsymbol{A} = (A, A, \ldots, A)$ and $\boldsymbol{B} = (B, B, \ldots, B)$, given the observed signal profile $\boldsymbol{y}_t = (y_{1t}, y_{2t}, \ldots, y_{Nt}) \in \{\bar{A}, \bar{B}\}^N$, as the respective predicted probabilities of their meeting opponents taking actions A and B at $t + 1$. Based on the posterior belief, players choose actions to maximize their expected current payoffs. This is equivalent to each player hiring a Bayesian expert to forecast the opponents' current plays; this expert takes the posterior probabilities of

the two equilibria, given the observed signals, as the predictions. This learning process is plausible for boundedly rational and myopic players, since they only care about themselves and current payoffs. However, each player has perfect knowledge of the actions of himself and his opponent at each time period, as well as of the signals of the remaining $N - 2$ players. The main reasons to assume that each player adopts all players' signals to forecast the future play are as follows. First, each player remains unaware of the actions of the other $N - 2$ players even though he may infer the action of his opponent from his own payoff. Second, the outcome of using N players' signals is the same as that of using or $N - 2$ players' signals in the limit of large population size, which is the focus of this paper. In this setup, players would have common beliefs, which induces a mathematically tractable dynamical process. Since revealed signals are public information and the same decision rule is employed by all players, at time $t \geq 1$, state $X_t(N)$ of the underlying dynamic system is either A or B. For convenience, the initial state, $X_0(N)$, is assumed to be distributed in $\{A, B\}$ as well. Moreover, for simplicity, we hereafter refer to the common action, A or B, of all players as the value of $X_t(N)$ for all $t \geq 0$ and N. Let

$$p_t(N) = P(X_t(N) = A) \quad \text{and} \quad 1 - p_t(N) = P(X_t(N) = B), \quad \text{for } t \geq 0.$$

In the following, $X_t(N)$ and $p_t(N)$ will sometimes be abbreviated as X_t and p_t, respectively, when no confusion occurs. Given signal y_t at time t, the posterior distribution, $P(X_t \mid y_t)$, is computed via Bayes' rule to be

$$P(X_t = A \mid y_t) = \frac{P(X_t = A, y_t)}{P(y_t)} = \frac{p_t P(y_t \mid X_t = A)}{p_t P(y_t \mid X_t = A) + (1 - p_t) P(y_t \mid X_t = B)}, \quad (2.2)$$

$$P(X_t = B \mid y_t) = \frac{P(X_t = B, y_t)}{P(y_t)} = \frac{(1 - p_t) P(y_t \mid X_t = B)}{p_t P(y_t \mid X_t = A) + (1 - p_t) P(y_t \mid X_t = B)}.$$

Note that $P(y_t \mid X_t = A) = u^{k'}(1 - u)^{N - k'}$ and $P(y_t \mid X_t = B) = (1 - v)^{k'} v^{N - k'}$, where $k' = n(\bar{A}, y_t)$ is the number of occurrences of $\bar{A}$ in y_t. The expected payoffs for the players at time $t + 1$ are

$$u(A, y_t) = P(X_t = A \mid y_t)a + P(X_t = B \mid y_t)b \quad \text{if action } A \text{ is taken,}$$

$$u(B, y_t) = P(X_t = A \mid y_t)c + P(X_t = B \mid y_t)d \quad \text{if action } B \text{ is taken.}$$

By maximizing the expected payoffs, players will choose action A at time $t + 1$ if and only if $P(X_t = A \mid y_t) \geq q^*$, where q^* is defined as in (2.1). A simple calculation using (2.2) shows that $P(X_t = A \mid y_t) \geq q^*$ if and only if

$$k' \log\left(\frac{uv}{(1 - u)(1 - v)}\right) \geq N \log\left(\frac{v}{1 - u}\right) + \log\left(\frac{q^*}{1 - q^*} \frac{1 - p_t}{p_t}\right). \quad (2.3)$$

Here $\log 0 = -\infty$ and $\log(1/0) = \infty$, by convention. Hence,

$$p_{t+1} = P(X_{t+1} = A) = \sum_{y_t \in \{\bar{A}, \bar{B}\}^N} P(y_t) P(X_{t+1} = A \mid y_t)$$

satisfies

$$p_{t+1} = \sum_{\{k' : (2.3) \text{ holds}\}} (p_t u^{k'} (1 - u)^{N - k'} + (1 - p_t)(1 - v)^{k'} v^{N - k'}) C_{k'}^N, \quad (2.4)$$

where $C_{k'}^N = N!/(k'!\,(N-k')!)$. Since, in (2.3), the range of k' depends on p_t, our boundedly rational quasi-Bayesian learning process $\{X_t : t \geq 0\}$ over the state space $\{A, B\}$ is not a Markov chain. Accordingly, the ergodic theorem for Markov chains is not applicable here. We have to analyze (2.4) carefully to find the limiting distributions of $\{X_t\}$.

3. The results

In this section, the limiting distributions of $\{X_t : t \geq 0\}$ are derived and characterized. First, the expression of (2.4) needs improvement. To neatly present the results, we introduce two functions: the odds ratio function, $p \mapsto r(p)$ with $r(p) = p/(1-p)$, and the logit function, $p \mapsto l(p)$ with $l(p) = \log(p/(1-p))$. We then let

$$\gamma = \frac{\log(v/(1-u))}{l(u) + l(v)} \quad \text{and} \quad \varepsilon(x) = \frac{l(q^*) - l(x)}{l(u) + l(v)} \quad \text{for } x \in [0, 1]. \tag{3.1}$$

Note that $\varepsilon(0) = \infty$ and $\varepsilon(1) = -\infty$, by the log convention. Define

$$g_N(x) = x \sum_{k' \geq N\gamma + \varepsilon(x)} u^{k'}(1-u)^{N-k'} C_{k'}^N + (1-x) \sum_{k' \geq N\gamma + \varepsilon(x)} (1-v)^{k'} v^{N-k'} C_{k'}^N \tag{3.2}$$

for $x \in [0, 1]$. It is then easily seen from (2.4) that

$$p_{t+1} = g_N(p_t) \quad \text{for } t \geq 0.$$

Before studying the fixed points of g_N, which are the natural candidates for the limiting distributions of $\{X_t\}$, we introduce independent, identically distributed 0–1 random variables $\{Y_i : i \geq 1\}$ and $\{\tilde{Y}_i : i \geq 1\}$ with $P(Y_i = 1) = u$ and $P(\tilde{Y}_i = 1) = 1 - v$. Let $S_N = \sum_{i=1}^N Y_i$ and $\tilde{S}_N = \sum_{i=1}^N \tilde{Y}_i$ be their respective partial sums. Here Y_i and $\tilde{Y}_i$ respectively represent the events that player i reveals signal $\bar{A}$ given that he takes actions A and B, while S_N and $\tilde{S}_N$ are the numbers of occurrences of signal $\bar{A}$ given that all players choose actions A and, respectively, B. Thus, (3.2) can be expressed as

$$g_N(x) = x \, P(S_N \geq N\gamma + \varepsilon(x)) + (1-x) \, P(\tilde{S}_N \geq N\gamma + \varepsilon(x)). \tag{3.3}$$

Let $\lfloor x \rfloor$ denote the largest integer no greater than x and let $\{x\} = x - \lfloor x \rfloor$ denote the fractional part of x. Furthermore, denote by $x_{N,k}$ the unique solution to

$$\varepsilon(x) = -k - \{N\gamma\}, \quad \text{where } k \text{ is an integer.}$$

A simple calculation using (3.1) shows that

$$x_{N,k} = \frac{1}{1 + r(q^*)^{-1}[r(u)r(v)]^{-k - \{N\gamma\}}}. \tag{3.4}$$

Hence, the sequence $(x_{N,k})$ is increasing in k, with

$$\lim_{k \downarrow -\infty} x_{N,k} = 0 \quad \text{and} \quad \lim_{k \uparrow \infty} x_{N,k} = 1. \tag{3.5}$$

Let $I_{N,k} = [x_{N,k}, x_{N,k+1})$. Then $-k - 1 - \{N\gamma\} < \varepsilon(x) \leq -k - \{N\gamma\}$ for $x \in I_{N,k}$.
By (3.3), we have

$$g_N(x) = x \, P(S_N \geq \lfloor N\gamma \rfloor - k) + (1-x) \, P(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k) \quad \text{on } I_{N,k}. \tag{3.6}$$

It follows that g_N is linear on $I_{N,k}$, with

$$g_N(x) = 1 \quad \text{on } [x_{N,\lfloor N\gamma \rfloor}, 1] \qquad \text{and} \qquad g_N(x) = 0 \quad \text{on } [0, x_{N,-N+\lfloor N\gamma \rfloor}). \tag{3.7}$$

This means that, for any N, the risk-dominant equilibrium will emerge in the long run if it is chosen at time 0 with probability no less than $x_{N,\lfloor N\gamma \rfloor}$. Conversely, the non-risk-dominant equilibrium will emerge in the long run if it is chosen at time 0 with probability no less than $1 - x_{N,-N+\lfloor N\gamma \rfloor}$. These results imply that players' long-run behaviors depend on their initial choices because our learning process is not a Markov chain. However, both $[x_{N,\lfloor N\gamma \rfloor}, 1]$ and $[0, x_{N,-N+\lfloor N\gamma \rfloor})$ shrink to singletons as $N \to \infty$, according to (3.5). Thus, to find the sizes of the basins of attraction of the two equilibria, we must investigate the function g_N thoroughly. We do so in what follows. The results for large population size are contained in Theorems 3.1–3.2.

By (3.6), the jump of function g_N at $x_{N,k}$ is

$$g_N(x_{N,k}) - g_N(x_{N,k}^-) = x_{N,k}\, P(S_N = \lfloor N\gamma \rfloor - k) + (1 - x_{N,k})\, P(\tilde{S}_N = \lfloor N\gamma \rfloor - k) \geq 0.$$

It is well known that the binomial distribution is stochastically increasing in the success parameter. Thus, $\mathrm{E}(Y_i) = u \geq \frac{1}{2} \geq 1 - v = \mathrm{E}(\tilde{Y}_i)$ implies that

$$P(S_N \geq m) \geq P(\tilde{S}_N \geq m) \quad \text{for } 0 \leq m \leq N. \tag{3.8}$$

By (3.6) and (3.8), the derivative, g_N', of g_N on $I_{N,k}$ satisfies

$$1 > g_N'(x) = P(S_N \geq \lfloor N\gamma \rfloor - k) - P(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k) \geq 0. \tag{3.9}$$

This proves the first part of the following result.

Lemma 3.1. (a) g_N *is increasing, piecewise linear, and right continuous.*

(b) *Any sequence $\{c_n : n \geq 0\}$ recurrently defined by $c_{n+1} = g_N(c_n)$ converges monotonically to a limit, c_∞, with $c_\infty = g_N(c_\infty)$ or $c_\infty = g_N(c_\infty^-)$.*

Together with (3.7), Lemma 3.1(a) shows that g_N is a distribution function. However, we do not know the exact form of the distribution. Let x be a quasi-fixed point if $x = g_N(x^-)$. Note that $c_\infty > c_0$ if c_0 is a quasi-fixed point. It is shown in Appendix A that

$$1 - v < \gamma < u \quad \text{for } \tfrac{1}{2} < u, v < 1. \tag{3.10}$$

Using the strong law of large numbers, from (3.3) we obtain

$$\lim_{N \to \infty} g_N(x) = x \quad \text{for } x \in [0, 1].$$

Note that the results for $x = 0$ and $x = 1$ follow trivially from the facts that $\varepsilon(0) = \infty$ and $\varepsilon(1) = -\infty$. As a consequence, finding the fixed points of g_N could be a complicated matter, even before distinguishing from among them those that are quasi-fixed.

Part (b) of Lemma 3.1, as shown in Appendix A, is proved by considering the relation between the graph of g_N on $I_{N,k}$ and the line $y = x$. Since g_N is linear and increasing, with $0 \leq g_N' < 1$ on $I_{N,k}$ according to (3.9), there are four exclusive cases.

Case (i):

$$g_N \text{ has a fixed point } z \text{ in } I_{N,k}. \tag{3.11}$$

Case (ii):

$$g_N(x) > x \text{ on } I_{N,k} \text{ but } x_{N,k+1} = g_N(x_{N,k+1}^-) \text{ is a quasi-fixed point.} \qquad (3.12)$$

Case (iii):

$$g_N(x) > x \text{ on } I_{N,k} \text{ with } g_N(x_{N,k+1}^-) > x_{N,k+1}. \qquad (3.13)$$

Case (iv):

$$g_N(x) < x \text{ on } I_{N,k}. \qquad (3.14)$$

Note that in case (i), z is the limit point, c_∞, guaranteed by Lemma 3.1 if $c_0 \in I_{N,k}$. By (3.6), case (iv) pertains if and only if

$$r(x)^{-1} < \frac{\mathrm{P}(S_N < \lfloor N\gamma \rfloor - k)}{\mathrm{P}(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} \quad \text{for } x \in I_{N,k}. \qquad (3.15)$$

Since $(1 - x)/x = (1/x) - 1$ is decreasing on $(0, 1)$, (3.15) holds if and only if it holds at $x = x_{N,k}$. Let $I(u, v) = [r(u)r(v)]^{-1}$. By using (3.4), we finally obtain

$$(3.14) \text{ holds} \quad \Longleftrightarrow \quad r(q^*)^{-1} < \frac{\mathrm{P}(S_N < \lfloor N\gamma \rfloor - k)}{\mathrm{P}(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} I(u, v)^{-k - \{N\gamma\}}. \qquad (3.16)$$

Similarly,

$$(3.13) \text{ holds} \quad \Longleftrightarrow \quad r(q^*)^{-1} I(u, v) > \frac{\mathrm{P}(S_N < \lfloor N\gamma \rfloor - k)}{\mathrm{P}(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} I(u, v)^{-k - \{N\gamma\}}, \qquad (3.17)$$

$$(3.12) \text{ holds} \quad \Longleftrightarrow \quad r(q^*)^{-1} I(u, v) = \frac{\mathrm{P}(S_N < \lfloor N\gamma \rfloor - k)}{\mathrm{P}(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} I(u, v)^{-k - \{N\gamma\}}, \qquad (3.18)$$

and, by exclusion,

$$(3.11) \text{ holds} \quad \Longleftrightarrow \quad r(q^*)^{-1} I(u, v) < \frac{\mathrm{P}(S_N < \lfloor N\gamma \rfloor - k)}{\mathrm{P}(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} I(u, v)^{-k - \{N\gamma\}} \leq r(q^*)^{-1}. \qquad (3.19)$$

Let

$$J(u, v, \gamma) = \frac{r(\gamma)r(u)^{-1}[1 - r(\gamma)^{-1}r(v)^{-1}]}{1 - r(\gamma)r(u)^{-1}}.$$

We defer the proofs of the following lemmas and theorems to Appendix A.

Lemma 3.2. (a) *For any fixed integer k, we have*

$$\lim_{N \to \infty} \frac{\mathrm{P}(S_N < \lfloor N\gamma \rfloor - k)}{\mathrm{P}(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} I(u, v)^{-k - \{N\gamma\}} = J(u, v, \gamma). \qquad (3.20)$$

(b) *Let $L > 0$ be fixed. For all k with $L \leq k \leq \lfloor N\gamma \rfloor$, and large N, we have*

$$\frac{\mathrm{P}(S_N < \lfloor N\gamma \rfloor - k)}{\mathrm{P}(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} I(u, v)^{-k - \{N\gamma\}} \leq \frac{J(u, v, \gamma)}{1 - (r(\gamma)r(v))^{-L-1}}. \qquad (3.21)$$

Equations (3.16)–(3.19) imply that the comparison between $r(q^*)^{-1} I(u, v)$ and $J(u, v, \gamma)$ is crucial. However, before proceeding, we need a lemma. Recall the definition of γ from (3.1).

Lemma 3.3. (a) *Let*

$$K(u, v) = \frac{v \log(v/[1 - u]) - (1 - v) \log(u/[1 - v])}{u \log(u/[1 - v]) - (1 - u) \log(v/[1 - u])} = \frac{v}{1 - u} J(u, v, \gamma).$$

Then, for $\frac{1}{2} < u, v < 1$,

$$J(u, v, \gamma) < 1 \quad \text{or, equivalently,} \quad \frac{1 - v}{u} < K(u, v) < \frac{v}{1 - u}. \tag{3.22}$$

(b) *For $\frac{1}{2} < v < u < 1$, we have $1 < K(u, v)$.*

We now present our main results.

Theorem 3.1. (a) *If $r(q^*)^{-1} I(u, v) > J(u, v, \gamma)$ then there exists, for any $\eta > 0$, an integer N_0 such that*

$$\lim_{t \to \infty} P(X_t(N) = A) = 1 \tag{3.23}$$

holds for the boundedly rational quasi-Bayesian learning process $\{X_t(N): t \geq 0\}$, given that $N \geq N_0$ and $P(X_0(N) = A) \in [\eta, 1]$. It holds, in particular, for $1 - v > q^*$ and $\frac{1}{2} < u \leq v$.

(b) *If $r(q^*)^{-1} I(u, v) < J(u, v, \gamma)$ then, for any fixed integer k, $g_N(x)$ has a fixed point in $I_{N,k}$ when the population size, N, is large. As a consequence, the number of possible limiting distributions of $\{X_t(N): t \geq 0\}$ grows to ∞ as N tends to ∞.*

Theorem 3.1 states how the players' long-run behaviors are determined by both the correlations between signals and underlying actions, and, for large population size, their initial choices. For $r(q^*)^{-1} I(u, v) > J(u, v, \gamma)$, Theorem 3.1(a) shows that players will eventually coordinate at the risk-dominant equilibrium unless the non-risk-dominant equilibrium is selected initially. In particular, the condition $r(q^*)^{-1} I(u, v) > J(u, v, \gamma)$ is fulfilled if $1 - v > q^*$ and $\frac{1}{2} < u \leq v$. Of the latter, $1 - v > q^*$ is reasonable since the smaller v is, the more likely it is that enough occurrences of signal A will be accumulated for players to switch from B to A. The other latter condition, $\frac{1}{2} < u \leq v$, is somewhat contrary to our intuition, since we might expect a larger value of u to favor action A and make (3.23) more likely to hold. In fact, this is not necessarily so. If $u > v > \frac{1}{2}$ then, by using Lemma 3.3(b) and letting $1 - v \approx q^*$ and $u \approx 1$, we obtain

$$J(u, v, \gamma) = \frac{1 - u}{v} K(u, v) > \frac{1 - u}{v} = \frac{u}{1 - v} I(u, v) > r(q^*)^{-1} I(u, v).$$

Thus, (3.23) no longer holds, by Theorem 3.1(b).

It is interesting to compare Theorem 3.1(a) with Theorem 3.1(a) of Chen and Chow (2003). Chen and Chow (2003) showed that under the simple learning process with $1 - v > q^*$, players will eventually coordinate at the risk-dominant equilibrium as the population size grows to ∞, whatever the players' initial choices are. In contrast, in our boundedly rational quasi-Bayesian learning process, the risk-dominant equilibrium will not emerge in the long run if players coordinate initially at the non-risk-dominant equilibrium. This conclusion remains true even when mutation of players' rational choices is included, as will be discussed in Section 4.

Theorem 3.1(b) implies the following results when $r(q^*)^{-1} I(u, v) < J(u, v, \gamma)$ and the population is large. First, the risk-dominant equilibrium will emerge in the long run if the players' initial probability of choosing risk-dominant equilibrium is close to 1. Second, the

non-risk-dominant equilibrium will emerge in the long run if the players' initial probability of selecting non-risk-dominant equilibrium is close to 1. However, when starting from either equilibrium with any other initial probability, players will visit both the risk-dominant and the non-risk-dominant equilibria with probabilities η and $1 - \eta$, $1 > \eta > 0$, respectively. The value of η is determined by the values of N, u, v, and q^*, and the initial state. Moreover, the number of possible values of η approaches infinity as the population size grows. Coexistence of the two equilibria was also found under the simple and maximum likelihood estimator learning processes by Chen and Chow (2003), but the unbounded number of limiting distributions for large population under the boundedly rational quasi-Bayesian learning process is a new and surprising discovery.

The boundary case $r(q^*)^{-1} I(u, v) = J(u, v, \gamma)$ seems too complicated to be analyzed here, as the approximation in (3.21) is not sharp. However, we have the following result for the most interesting case, namely $u = v > \frac{1}{2}$, which implies that

$$\gamma = \tfrac{1}{2}, \qquad J(u, u, \gamma) = r(u)^{-1}, \quad \text{and} \quad I(u, u) = (r(u))^{-2}. \tag{3.24}$$

A simple calculation shows that

$$r(q^*)^{-1} I(u, u) \begin{cases} > J(u, u, \gamma) & \text{if and only if } 1 - q^* > u, \\ = J(u, u, \gamma) & \text{if and only if } 1 - q^* = u, \\ < J(u, u, \gamma) & \text{if and only if } 1 - q^* < u. \end{cases} \tag{3.25}$$

By combining (3.25) with Theorem 3.1, we obtain the following result.

Theorem 3.2. *Suppose that $u = v$.*

(a) *If $1 - q^* > u$ then the conclusion of Theorem 3.1(a) holds.*

(b) *If $1 - q^* < u$ then the conclusion of Theorem 3.1(b) holds.*

(c) *If $1 - q^* = u$ then the following holds for any $N \geq 2$:*

$$\lim_{t \to \infty} \mathrm{P}(X_t(N) = A) = 1, \quad \text{given that } \mathrm{P}(X_0(N) = A) \in [x_{N,0}, 1]. \tag{3.26}$$

Moreover, for any fixed, negative integer k, $g_N(x)$ has a fixed point in $I_{N,k}$ when the population size, N, is large. As a consequence, the number of possible limiting distributions of $\{X_t(N): t \geq 0\}$ grows to ∞ as N tends to ∞.

Theorem 3.2(c) shows how the players' long-run behaviors in the case in which

$$r(q^*)^{-1} I(u, v) = J(u, v, \gamma)$$

rely on the initial state, $X_0(N)$. Note that when $u = v$ we have

$$x_{N,k} = \frac{1}{1 + r(q^*)^{-1}(r(u)r(v))^{-k}},$$

by (3.4), (3.24), and the assumption that N is even. In particular, $x_{N,0} = q^*$ is a constant independent of N. Thus, the limiting behavior of $\{X_t(N): t \geq 0\}$ in the boundary case, $1 - q^* = u$, is similar to that for $1 - q^* > u$ when the initial distribution satisfies $\mathrm{P}(X_0(N) = A) \in I_{N,k}$ with $k \geq 0$, since $x_{N,k} > x_{N,0}$, by (3.5), and is similar to that for $1 - q^* < u$ when

$P(X_0(N) = A) \in I_{N,k}$ with $k < 0$. In other words, players will eventually coordinate at the risk-dominant equilibrium when its initial probability is not less than q^*. Otherwise, either coexistence of the equilibria or the non-risk-dominant equilibrium alone will emerge in the long run.

It is worthwhile to discuss the long-run equilibrium when $u = v$ and $u \to 1$, under which conditions the players' actions are perfectly monitored. In these circumstances, $1 - q^* < u$. Consequently, Theorem 3.2(b) implies that, for large population size, players can coordinate at the risk-dominant equilibrium, select the non-risk-dominant equilibrium, or visit both equilibria with positive probability. This outcome differs from that of Chen and Chow (2003), namely that the players will coordinate at the risk-dominant equilibrium in the simple learning process with $u = v$ and $u \to 1$.

4. Extension

To make it possible for players to choose different actions in each time period, we include mutation of the players' rational choices in each time period. It is shown that the results in Sections 2 and 3 are still valid when mutation of the players' boundedly rational choices is considered. After choosing his boundedly rational action, a player is presumed to have a positive probability, ε, of disobeying the choice at the end of each time period. The mutation rate (ε) is fixed over time and with respect to player. Accordingly, the probability that signal $\bar{A}$ is revealed by a player at the beginning of a time period, given that action A was taken in the last period, becomes $\bar{u} = (1 - \varepsilon)u + \varepsilon(1 - v)$. Similarly, the probability that signal $\bar{B}$ is revealed by a player at the beginning of a time period, given that action B was taken in the last period, becomes $\bar{v} = (1 - \varepsilon)v + \varepsilon(1 - u)$. It is easy to check that $0 < \bar{u}, \bar{v} < 1$ given that $0 < u, v < 1$, and that $\bar{u}, \bar{v} > \frac{1}{2}$ when ε is sufficiently small. Then, given an $\varepsilon > 0$, the dynamics of players' rational choices under the original boundedly rational quasi-Bayesian learning process, $\{\hat{X}_t(\varepsilon) : t \geq 1\}$, still has state space $\{A, B\}$, but the frequencies of occurrence of the signals are $\bar{u}$ and $\bar{v}$. By the same argument used in Section 3, we can obtain the limiting distribution of $\{\hat{X}_t(\varepsilon) : t \geq 1\}$, denoted by

$$(\hat{p}(N, \varepsilon), \hat{q}(N, \varepsilon)) = \left(\lim_{t \to \infty} P(\hat{X}_t(\varepsilon) = A), \lim_{t \to \infty} P(\hat{X}_t(\varepsilon) = B) \right).$$

Accordingly, all the results of Section 3 hold here, for $\varepsilon > 0$. Given ε and $(\hat{p}(N, \varepsilon), \hat{q}(N, \varepsilon))$, we can derive the limiting distribution of the players' actual action profiles, $\{P(s, \varepsilon)\}_{s \in \{A, B\}^N}$, where

$$P(s, \varepsilon) = \hat{p}(N, \varepsilon)\varepsilon^{d(s, A)}(1 - \varepsilon)^{N - d(s, A)} + \hat{q}(N, \varepsilon)\varepsilon^{d(s, B)}(1 - \varepsilon)^{N - d(s, B)}$$

for all $s = (s_1, s_2, \ldots, s_N) \in \{A, B\}^N$. Here, $d(s, A) := |\{i \in \{1, 2, \ldots, N\} : s_i \neq A\}|$ and $d(s, B) := |\{i \in \{1, 2, \ldots, N\} : s_i \neq B\}|$ respectively count the numbers of players in s not taking actions A and B. As in Kandori *et al.* (1993), by letting $\varepsilon \downarrow 0$ we obtain the long-run equilibrium,

$$\lim_{\varepsilon \downarrow 0} P(s, \varepsilon) = \begin{cases} \hat{p}(N, \varepsilon) & \text{if } s = A, \\ \hat{q}(N, \varepsilon) & \text{if } s = B, \\ 0 & \text{otherwise.} \end{cases}$$

This shows that the long-run equilibria under the boundedly rational quasi-Bayesian learning process with diminishing-to-zero mutation are the same as those without mutation in Sections 2 and 3.

5. Conclusion

In this paper we have constructed a boundedly rational quasi-Bayesian learning process in evolutionary coordination games with imperfect monitoring. In the setup of imperfect monitoring, random signals play the role of mutation to make the movement between equilibria possible. Moreover, unlike previous evolutionary learning processes, our learning process is not a Markov chain. Thus, the players' long-run behaviors depend not only on the correlations between signals and underlying actions, but on the players' initial choices. The relative magnitudes of $r(q^*)^{-1}I(u, u)$ and $J(u, u, \gamma)$ classify the long-run equilibria. When $r(q^*)^{-1}I(u, u) > J(u, u, \gamma)$, players will eventually coordinate at the risk-dominant equilibrium unless their initial choice was the non-risk-dominant equilibrium. If $r(q^*)^{-1}I(u, u) < J(u, u, \gamma)$, the risk-dominant equilibrium, the non-risk-dominant equilibrium, or coexistence of the two equilibria can emerge in the long run, and players' initial choices will determine the final result. Moreover, the number of long-run equilibria becomes infinite as the population size grows to infinity. In the boundary case, $r(q^*)^{-1}I(u, u) = J(u, u, \gamma)$, either of the above outcomes could happen. To the authors' knowledge, Chen and Chow (2003) was the first paper to address the issue of equilibrium selection in coordination games with imperfect monitoring and myopic players. In the model of that paper, random signals are employed to refine multiple equilibria. However, the findings of this study undermine the efficacy of random signals in refining the evolutionary process with imperfect monitoring.

Appendix A.

A.1. Proof of Lemma 3.1

Part (a). The proof follows from (3.7)–(3.9).

Part (b). Suppose that $c_0 \in I_{N,k}$. Note that g_N is linear, with $0 \le g'_N < 1$ on $I_{N,k}$ by (3.9). For the four exclusive cases in (3.11)–(3.14), the sequence (c_n) converges monotonically to z and $x_{N,k+1}$ in cases (i) and (ii), respectively. In case (iv), (c_n) decreases in $I_{N,k}$ until it first exits $I_{N,k}$ at some $c_{m_1} \in I_{N,k_1}$, with $k_1 < k$. Since g_N is increasing in $[0, 1]$, $g_N(x) < x$ on $I_{N,k}$, and $c_{m_1-1} \in I_{N,k}$ by assumption,

$$c_{m_1} = g_N(c_{m_1-1}) < c_{m_1-1} \quad \text{and, thus,} \quad g_N(c_{m_1}) < g_N(c_{m_1-1}) = c_{m_1}. \tag{A.1}$$

Now repeat this procedure starting with $c_{m_1} \in I_{N,k_1}$. The results in cases (i) and (ii) follow as before (note that case (ii) is in fact excluded by (A.1)). Otherwise, we need only consider case (iv), in view of (A.1). By the same arguments as above, we are led to some $c_{m_2} \in I_{N,k_2}$ with $k_2 < k_1$ and $g_N(c_{m_2}) < c_{m_2}$. We thus restart with $c_{m_2} \in I_{N,k_2}$. Such reduction cannot be repeated indefinitely because (3.7) implies not only $g_N(0) = 0$ and $g_N(1) = 1$, but that only finite intervals $I_{N,j}$ with $-N + \lfloor N\gamma \rfloor \le j < \lfloor N\gamma \rfloor$ remain to be studied. In conclusion, the sequence (c_n) converges monotonically to either a fixed point or a quasi-fixed point of g_N in case (iv).

Case (iii) can be dealt with similarly. This completes the proof.

A.2. Proof of inequality (3.10)

Note that $1 - \gamma = \log(u/[1-v])/[l(u) + l(v)]$. By symmetry, it is enough to prove the first inequality in (3.10). A simple calculation shows that this is equivalent to

$$h(1 - v) - h(u) > 0,$$

where $h(x) = v \log(1-x) + (1-v) \log x$ for $x \in (0, 1)$. By differentiation,

$$h'(x) = -\frac{v}{1-x} + \frac{1-v}{x} = \frac{1-v-x}{x(1-x)},$$

which is positive on $(0, 1-v)$ and negative on $(1-v, 1)$. Hence, $h(x)$ attains its unique maximum at $x = 1-v$. This proves (3.10).

A.3. Proof of Lemma 3.2

Part (a). After an appropriate scaling, we have

$$P(S_N < \lfloor N\gamma \rfloor - k) = \sum_{\ell=1}^{\lfloor N\gamma \rfloor - k} P(S_N = \lfloor N\gamma \rfloor - k - \ell). \tag{A.2}$$

Recall that $C_m^N = N!/(m!\,(N-m)!)$. Using $\lfloor N\gamma \rfloor = N\gamma - \{N\gamma\}$ and the binomial theorem, we have

$$
\begin{aligned}
P(S_N = \lfloor N\gamma \rfloor - k - \ell) &= C_{\lfloor N\gamma \rfloor - k - \ell}^N u^{\lfloor N\gamma \rfloor - k - \ell}(1-u)^{N - \lfloor N\gamma \rfloor + k + \ell} \\
&= C_{\lfloor N\gamma \rfloor - k}^N (u^\gamma (1-u)^{1-\gamma})^N \left(\frac{1-u}{u}\right)^{\{N\gamma\} + k + \ell} A(N, k, \ell), \quad \text{(A.3)}
\end{aligned}
$$

where $A(N, k, \ell) = \prod_{j=1}^{\ell}(\lfloor N\gamma \rfloor - k - j + 1)/(N - \lfloor N\gamma \rfloor + k + j)$. Since $\lfloor N\gamma \rfloor \leq N\gamma$ and k is fixed, for each fixed $\ell \geq 1$ we have $\lim_{N\to\infty} A(N, k, \ell) = (\gamma/(1-\gamma))^\ell$ and, for any fixed $\varepsilon > 0$,

$$A(N, k, \ell) \leq \left(\frac{N\gamma - k}{N - N\gamma + k}\right)^\ell \leq \left(\varepsilon + \frac{\gamma}{1-\gamma}\right)^\ell$$

for large N. Furthermore, since $N\gamma - k \leq N\gamma$ for $k \geq 0$, we have

$$A(N, k, \ell) \leq \left(\frac{\gamma}{1-\gamma}\right)^\ell \quad \text{for } k \geq 0. \tag{A.4}$$

It follows from (A.2), (A.3), and Lebesgue's convergence theorem that

$$\lim_{N\to\infty} \frac{P(S_N < \lfloor N\gamma \rfloor - k)}{C_{\lfloor N\gamma \rfloor - k}^N (u^\gamma (1-u)^{1-\gamma})^N} \left(\frac{u}{1-u}\right)^{k+\{N\gamma\}} = \sum_{\ell=1}^{\infty} \left(\frac{\gamma(1-u)}{(1-\gamma)u}\right)^\ell. \tag{A.5}$$

By the same arguments, we have

$$P(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k) = \sum_{\ell=0}^{N - \lfloor N\gamma \rfloor + k} P(\tilde{S}_N = \lfloor N\gamma \rfloor - k + \ell) \tag{A.6}$$

and

$$P(\tilde{S}_N = \lfloor N\gamma \rfloor - k + \ell) = C_{\lfloor N\gamma \rfloor - k}^N ((1-v)^\gamma v^{1-\gamma})^N \left(\frac{1-v}{v}\right)^{-\{N\gamma\} - k + \ell} \tilde{A}(N, k, \ell), \tag{A.7}$$

where $\tilde{A}(N, k, \ell) = \prod_{j=1}^{\ell}(N - \lfloor N\gamma \rfloor + k + 1 - j)/(\lfloor N\gamma \rfloor - k + j)$ converges to $((1-\gamma)/\gamma)^\ell$ for each fixed $\ell \geq 1$. Thus, again by Lebesgue's convergence theorem, we have

$$\lim_{N\to\infty} \frac{P(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)}{C_{\lfloor N\gamma \rfloor - k}^N ((1-v)^\gamma v^{1-\gamma})^N} \left(\frac{1-v}{v}\right)^{k+\{N\gamma\}} = \sum_{\ell=0}^{\infty} \left(\frac{(1-\gamma)(1-v)}{\gamma v}\right)^\ell. \tag{A.8}$$

Note that both series above converge, by (3.10). Moreover,

$$\frac{N - \lfloor N\gamma \rfloor + k + 1 - j}{\lfloor N\gamma \rfloor - k + j} \geq \frac{N - N\gamma}{N\gamma} = \frac{1 - \gamma}{\gamma}$$

for all k with $k \geq L \geq \ell$. Thus,

$$\tilde{A}(N, k, \ell) \geq \left(\frac{1 - \gamma}{\gamma}\right)^\ell \quad \text{for all } k \text{ with } k \geq L \geq \ell. \tag{A.9}$$

Since $(1 - \gamma)/\gamma = \log(u/[1 - v])/\log(v/[1 - u])$, by (3.1), we have $u^\gamma (1 - u)^{1-\gamma} = (1 - v)^\gamma v^{1-\gamma}$. By taking the ratio of (A.5) to (A.8), (3.20) follows from a simple calculation based on the formula $\sum_{\ell=m}^\infty x^\ell = x^m/(1 - x)$, $0 \leq x < 1$.

Part (b). Note that we require N to be large to guarantee that there is indeed some k such that $L \leq k \leq \lfloor N\gamma \rfloor$. By (A.2)–(A.4), we have

$$\frac{\mathrm{P}(S_N < \lfloor N\gamma \rfloor - k)}{C^N_{\lfloor N\gamma \rfloor - k}(u^\gamma (1 - u)^{1-\gamma})^N} \left(\frac{u}{1 - u}\right)^{k+\{N\gamma\}} \leq \sum_{\ell=1}^\infty \left(\frac{\gamma(1 - u)}{(1 - \gamma)u}\right)^\ell. \tag{A.10}$$

By (A.6), (A.7), and (A.9), for all k with $L \leq k \leq \lfloor N\gamma \rfloor$ we have

$$\frac{\mathrm{P}(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)}{C^N_{\lfloor N\gamma \rfloor - k}((1 - v)^\gamma v^{1-\gamma})^N} \left(\frac{1 - v}{v}\right)^{k+\{N\gamma\}} \geq \sum_{\ell=0}^L \left(\frac{(1 - \gamma)(1 - v)}{\gamma v}\right)^\ell, \tag{A.11}$$

as the upper limit, $N - \lfloor N\gamma \rfloor + k$, in the sum in (A.6) is no less than L in this case. Equation (3.21) now follows by taking the ratio of (A.10) to (A.11) and using the formula $\sum_{\ell=0}^L x^\ell = (1 - x^{L+1})/(1 - x)$.

A.4. Proof of Lemma 3.3

Part (a). Note that (3.20) implies $J(u, v, \gamma) \geq 0$. A simple calculation using $(1 - \gamma)/\gamma = \log(u/[1 - v])/\log(v/[1 - u])$ shows that $J(u, v, \gamma) < 1$ if and only if

$$\frac{v \log(v/[1 - u]) - (1 - v) \log(u/[1 - v])}{u \log(u/[1 - v]) - (1 - u) \log(v/[1 - u])} = K(u, v) = \frac{v}{1 - u} J(u, v, \gamma) < \frac{v}{1 - u}. \tag{A.12}$$

This is the second inequality for $K(u, v)$ in (3.22). The first one can be obtained from it by noting that $K(u, v) = 1/K(v, u)$. It thus suffices to verify that (A.12) holds.

For $0 < a < 1$, it is easy to see, by differentiation, that the function

$$h_a(x) = a \log\frac{a}{x} - (1 - a) \log\frac{1 - x}{1 - a}$$

on $(0, 1)$ attains its unique minimum value, 0, at $x = a$. Hence, both the numerator and the denominator of the leftmost quotient in (A.12) are positive. A little rearrangement shows that (A.12) is equivalent to

$$2(1 - u)v \log\frac{v}{1 - u} < (uv + (1 - u)(1 - v)) \log\frac{u}{1 - v}.$$

By adding $(uv + (1 - u)(1 - v)) \log(v/[1 - u])$ to both sides of this, we see that (A.12) is equivalent to

$$(1 + v - u) \log\frac{v}{1 - u} < (uv + (1 - u)(1 - v)) \log\frac{uv}{(1 - u)(1 - v)}. \tag{A.13}$$

Now change variables by letting $x = v/[1-u]$ and $y = u/[1-v]$. Note that $x, y > 1$ because $\frac{1}{2} < u, v < 1$. It is easy to see that $u = (x-1)y/(xy-1)$ and $v = (y-1)x/(xy-1)$. Hence,

$$1 - u = \frac{y-1}{xy-1}, \qquad 1 - v = \frac{x-1}{xy-1},$$

$$1 + v - u = \frac{(x+1)(y-1)}{xy-1}, \quad \text{and} \quad 1 + u - v = \frac{(y+1)(x-1)}{xy-1}. \tag{A.14}$$

After some rearrangement and calculation, (A.13) can be rewritten as

$$\frac{x+1}{x-1} \log x < \frac{xy+1}{xy-1} \log(xy). \tag{A.15}$$

Introduce the functions $f(x) = [(x+1)/(x-1)] \log x$ and $e(x) = (x^2-1)/x - 2\log x$, for $x \in [1, \infty)$. Since $e(1) = 0$ and $e'(x) = 1 + x^{-2} - 2x^{-1} = (x-1)^2/x^2 > 0$ in $(1, \infty)$, we have $e(x) > 0$ in $(1, \infty)$. It is easy to check that $f'(x) = e(x)/(x-1)^2 \geq 0$. Hence, f is increasing in $[1, \infty)$. Because $y > 1$, we have $xy > x$. Therefore, $f(x) < f(xy)$. This proves (A.15) and, thus, (3.22), in view of (A.12) and (A.13).

Part (b). A simple rearrangement shows that the conclusion of part (b) is equivalent to

$$(1 + u - v) \log \frac{u}{1-v} < (1 + v - u) \log \frac{v}{1-u}.$$

In terms of x and y, the equation above can be rewritten, using (A.14), as

$$(y+1)(x-1) \log y < (x+1)(y-1) \log x$$

or, equivalently,

$$f(y) < f(x). \tag{A.16}$$

It is easy to check that $x > y$ under the assumption that $\frac{1}{2} < v < u$. Since f is known to be increasing in $[1, \infty)$, (A.16) holds and the conclusion follows.

A.5. Proof of Theorem 3.1

Part (a). For any fixed k, the assumption of Theorem 3.1(a) and (3.20) together imply (3.17), and thus (3.13) holds for large N. Moreover, we can use (3.21) to conclude that (3.17) holds simultaneously for all k with $L \leq k \leq \lfloor N\gamma \rfloor$ if L is first chosen to be sufficiently large that

$$\frac{1 - q^*}{q^*} I(u, v) > \frac{J(u, v, \gamma)}{1 - (1-\gamma)(1-v)^{L+1}/(\gamma v)}.$$

In view of (3.5) and (3.7), $x = 1$ is the only possible limit point in $[\eta, 1]$ guaranteed to exist by Lemma 3.1(b). This proves (3.23).

It remains to verify that $[(1-q^*)/q^*]I(u, v) > J(u, v, \gamma)$ holds when $1 - v > q^*$ and $u \leq v$, which conditions imply that

$$\frac{u}{1-v} \leq \frac{v}{1-v} < \frac{1-q^*}{q^*}.$$

By (A.12), we have $J(u, v, \gamma) = [(1-u)/v]K(u, v)$. Since $K(u, v) = 1/K(v, u)$ and $K(u, u) = 1$, from Lemma 3.3(b) we have $K(u, v) \leq 1$ for $u \leq v$. Putting these results together, we obtain

$$J(u, v, \gamma) \leq \frac{1-u}{v} < \frac{1-q^*}{q^*} I(u, v),$$

as desired. Note that $I(u, v) = (1-u)(1-v)/(uv)$ by definition.

Part (b). Let k be fixed. Note that $(1 - q^*)/q^* > 1$ because $q^* < \frac{1}{2}$ (by assumption (2.1)). According to (3.20) and (3.22), the second inequality in (3.19) holds for large N. The first inequality also holds, by the assumption of Theorem 3.1(b) and (3.20). Hence, (3.11) holds for large N. The conclusion follows from the remark following (3.14).

A.6. Proof of Theorem 3.2

By (3.25) and Theorem 3.1, it remains to consider the boundary case, in which $1 - q^* = u$. When $u = v$, the random variables Y_i and $1 - \tilde{Y}_i$ have the same distribution, and so do S_N and $N - \tilde{S}_N$. In particular,

$$P(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k) = P(S_N \leq N - \lfloor N\gamma \rfloor + k). \tag{A.17}$$

In (3.21), we claim that for $-N + \lfloor N\gamma \rfloor \leq k \leq \lfloor N\gamma \rfloor$,

$$\frac{P(S_N < \lfloor N\gamma \rfloor - k)}{P(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} \begin{cases} < \left(\dfrac{1-u}{u}\right)^{2k+1+\{N\gamma\}} & \text{if } k \geq 0, \\[2mm] > \left(\dfrac{1-u}{u}\right)^{2k+1+\{N\gamma\}} & \text{if } k < 0. \end{cases} \tag{A.18}$$

Note that $\{N\gamma\} = 0$ as $\gamma = \frac{1}{2}$ by (3.24) and N is even by assumption, although (A.18) can be shown to be valid for odd N as well.

Temporarily suppose that (A.18) holds. Since $1 - q^* = u$, we have, from (3.24) and (A.18),

$$\frac{P(S_N < \lfloor N\gamma \rfloor - k)}{P(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} I(u, u)^{-k-\{N\gamma\}} \begin{cases} < \dfrac{1-u}{u} = \dfrac{1-q^*}{q^*} I(u, u) \ \text{ for } 0 \leq k \leq \lfloor N\gamma \rfloor, \\[2mm] > \dfrac{1-q^*}{q^*} I(u, u) & \text{for } -N + \lfloor N\gamma \rfloor \leq k < 0. \end{cases} \tag{A.19}$$

That is, (3.17) and, thus, (3.13) hold for all k, $0 \leq k \leq \lfloor N\gamma \rfloor$. Hence, (3.26) follows from the arguments used in the proof of Theorem 3.1(a). For any negative integer k, the first inequality in (3.19) holds by (A.19). Since $q^* < \frac{1}{2}$ by assumption (2.1), we have $1 < (1 - q^*)/q^*$; thus, the second inequality in (3.19) holds for large N by (3.20) and (3.22). Case (i) (see (3.11)) therefore pertains and the conclusion follows as in Theorem 3.1(b).

It remains to verify that (A.18) holds. Let $N = 2m$. By using $\gamma = \frac{1}{2}$ and (A.17), we obtain

$$\frac{P(S_N < \lfloor N\gamma \rfloor - k)}{P(\tilde{S}_N \geq \lfloor N\gamma \rfloor - k)} = \frac{P(S_N < m - k)}{P(S_N \leq m + k)} = \frac{\sum_{\ell=1}^{m-k} P(S_N = m - k - \ell)}{\sum_{\ell=1}^{m+k+1} P(S_N = m + k + 1 - \ell)}. \tag{A.20}$$

First consider the case in which $0 \leq k \leq \lfloor N\gamma \rfloor = m$. There are fewer terms in the numerator than in the denominator on the right-hand side of (A.20). Because C_j^{2m} is increasing for $0 \leq j \leq m$ and $C_{m+j}^{2m} = C_{m-j}^{2m}$, it is easy to see that $C_{m-k-\ell}^{2m} < C_{m+k+1-\ell}^{2m}$ for each ℓ, $1 \leq \ell \leq m - k$. Hence,

$$\frac{P(S_N = m - k - \ell)}{P(S_N = m + k + 1 - \ell)} = \frac{C_{m-k-\ell}^{2m} u^{m-k-\ell} (1-u)^{m+k+\ell}}{C_{m+k+1-\ell}^{2m} u^{m+k+1-\ell} (1-u)^{m-k-1+\ell}} < \left(\frac{1-u}{u}\right)^{2k+1}.$$

This together with (A.20) verifies (A.18) for $0 \leq k \leq \lfloor N\gamma \rfloor = m$. The remaining case, in which $-m = -N + \lfloor N\gamma \rfloor \leq k < 0$, can be treated similarly and is thus omitted.

References

BERGIN, J. AND LIPMAN, B. L. (1996). Evolution with state-dependent mutations. *Econometrica* **64,** 943–956.

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. AND CHOW, Y. (2003). Equilibrium selection in coordination games with imperfect monitoring. Working paper, National Taipei University.

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.

FOSTER, D. AND YOUNG, P. (1990). Stochastic evolutionary game dynamics. *Theoret. Pop. Biol.* **38,** 219–232.

FUDENBERG, D., LEVINE, D. AND MASKIN, E. (1994). The folk theorem with imperfect public information. *Econometrica* **62,** 997–1039.

JORDAN, J. S. (1991). Bayesian learning in normal form games. *Games Econom. Behavior* **3,** 60–81.

KALAI, E. AND LEHRER, E. (1993). Rational learning leads to Nash equilibrium. *Econometrica* **61,** 1019–1045.

KANDORI, M. G., MAILATH, G. J. 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.

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 repeated prisoner's dilemma with private monitoring. *J. Econom. Theory* **76,** 345–361.

YOUNG, H. P. (1993). The evolution of conventions. *Econometrica* **61,** 57–84.