Learning a Population Distribution*

Seung Han Yoo†

January 8, 2013

Abstract

This paper introduces a dynamic Bayesian game with an unknown population distribution. Players do not know the true population distribution and assess it based on their *private* observations using Bayes' rule. First, we show the existence and characterization of an equilibrium in which each player's strategy is a function not only of the player's type but also of experience. Second, we show that each player's initial belief about the population distribution converges almost surely to a "correct" belief.

Keywords and Phrases: Bayesian games, Dynamic games, Bayesian learning

JEL Classification Numbers: **C72, C73, D83**

---

*This paper is a revised version of the paper "A Dynamic Bayesian Game with an Unknown Population Distribution." I am grateful to Geir Asheim, Edward Green, Seungjin Han, James Jordan, Atsushi Kajii, Jeong-Yoo Kim, Andrew McLennan, Ehud Lehrer, Yaw Nyarko, Larry Samuelson and seminar participants at the Far Eastern Meeting of the Econometric Society, 5th Pan-Pacific Conference on Game Theory, 4h Asian Decentralization Conference, Midwest Economic Theory Meeting 2009 for helpful comments. I would especially like to thank Fernando Vega-Redondo for valuable comments and encouragement in the early stage of this project. Of course, all remaining errors are mine.

†Department of Economics, Korea University, Seoul, Republic of Korea 136-701 (e-mail: `shyoo@korea.ac.kr`).

1 Introduction

When a (true) population distribution is not known, as opposed to the Bayesian game by Harsanyi (1967), players who are repeatedly matched have incentives to learn from others' actions about the population distribution. For example, in reality, before people decide to take a particular action and anticipate others' strategies, they try to figure out what the relevant population distributions look like. In this paper, we are interested in those game-theoretic situations and their consequences.

We assume that there is a large population, and that players from the population are distributed according to a probability distribution. It is common knowledge that there is a true population distribution among two distributions in which one "locally" first-order stochastic dominates the other,[1] but players do not know which one is the true population distribution. In each period $t \in \mathbb{N}$, every two players are randomly paired to play a 2×2 game, and after the game, each player's belief about the population distribution is updated by Bayes' rule after observing the other's action. They are paired for just one period, and rematched again after each period. The *types* of each player are relevant for the player's own payoff, and observations in each match are *private*. Hence, the matching mechanism of the game is "public," but the information flow of the game is "private."[2]

Since each player's monitoring is private, each player's estimate about the population distribution is also private. Hence, each player's strategy is a function of both the player's type and observations. We call player i's history of observations up to $t-1$ player i's *experience* at t, and a vector of player i's type and experience player i's *characteristic,* since a type (resp. an experience) of player i can be defined as his or her innate (resp. acquired) characteristic. Given this private type and estimate,

[1] For a formal definition, see (A5) in section 3.

[2] As observed by Fudenberg and Levine (1998), this environment is most frequently used in experiments for game theory: "*Random-matching model* Each period all players are randomly matched. At the end of each round, each player observes only the play in his own match...... This is the treatment most frequently used in game theory experiments" (p. 6).

each player anticipates what samples and data the opponent has observed in the past as well as the opponent's type.

This paper provides three main results: *existence*, *monotonicity* and *convergence*. For each period, every two-paired players play a Bayesian game with their characteristics. Since a Bayesian game even with a *one*-dimensional type space may not have an equilibrium,[3] we adopt a simple model in which a Bayesian equilibrium strategy is parameterized by a threshold type. In addition, by assuming players have the same *initial* beliefs, we enable them to construct expectations without invoking the hierarchy of beliefs problem, which extends Harsanyi's ingenious idea (Theorem 1).

Second, the monotonicity result establishes that if given each period, a player believes that a stochastic dominant distribution is more likely, then the player's equilibrium strategy shows a certain monotonic pattern (Theorem 2).[4] In particular, if a player has an experience that generates beliefs such that a stochastic dominant distribution is more likely, then his or her optimal strategy is to make more types take an action that *were* chosen under the stochastic dominant one. Hence, given each period, the same type, which is determined in the beginning of a game, can make different decisions depending on different experiences, or sample paths. For example, if a person takes a "bad" action, it may be because the person is bad by nature or because the person has observed many bad actions on the part of others in the past.[5]

This also allows us new interpretations about experimental results, which often do not support theoretical predictions, especially, in two different ways: before an

[3] There are two main approaches to tackling the existence of a Bayesian game with general action and multidimensional type spaces. One is by McAdams (2003), extending Athey (2001), who suggests the single crossing condition (p. 866) for a *one*-dimensional type space, and the other is by Vives (1990) and Van Zandt and Vives (2007), who utilize supermodular *payoffs*.

[4] See Marimon (1997) for arguments that convergence alone is not sufficient to make learning theory interesting.

[5] Hence, the world one believes one knows can be just the reflection of what one perceives.

experiment, different "subjects" may have different preconceptions, given their past *private* experiences, about "objects" of the experiment, and during an experiment, subjects can learn even with a random matching; for example, van Huyck, Battalio and Beil (1990) find, using minimum effort games, that in random-pair experiments, the subjects' dynamic behavior shows learning features similar to those in *fixed*-pair experiments.[6]

Third, we show that each player's initial belief about the true distribution converges *almost surely* to a correct belief (Theorem 3). This also implies that the limit of equilibrium strategies from a sequence of observations is equal to the equilibrium strategy with common knowledge. In other words, Harsanyi's common distribution assumption is justifiable in the long run.[7] On the other hand, in the real world, people can observe a large but only *finite* number of samples, so their experiences will influence their beliefs about population distributions, such as the above second result, which provides a different perspective on "almost common knowledge" situation studied by Rubinstein (1989).

To the author's knowledge, this is the first paper to attempt to study Bayesian learning of an unknown type distribution through players' interactions. This paper also departs from papers in learning in games (for surveys, see Marimon (1997), Fudenberg and Levine (1998), Vega-Redondo (2003) and Nachbar (2004)) in two major ways; a signaling process is *endogenously* generated from players' actions, and each player receives *private* signals. Hence, monitoring is private in this paper,

---

[6]van Huyck, Battalio and Beil (1990) write that "Experiments six and seven randomly paired subjects with an unknown partner. Hence, experiments six and seven test whether the results obtained in experiments four and five were due to subjects repeating the period with the same opponent.... Moreover, the subjects' dynamic behavior was similar to that found in the fixed pair C treatment" (p. 244). However, they do not provide an explanation for how this occurs.

[7]This reports an optimistic prediction; if people can be matched infinitely often in a random manner, the *negative* biases or stereotypes they hold about others will disappear. For instance, if there are two groups (Black\White, etc), and members of each group can have certain beliefs about the other's distribution on characteristics, then the infinite random matching between them yields the convergence to a correct belief.

whereas public monitoring is allowed in papers on learning with repeated games (Jordan (1991, 1995), Kalai and Lehrer (1993a), and Nyarko (1994, 1998)). In addition, we focus on an equilibrium process like Jordan (1991, 1995) and Jackson and Kalai (1997), which is different from the literature on learning with "non-equilibrium processes" (Fudenberg and Levine (1993a, b), Kalai and Lehrer (1993b), Fudenberg and Kreps (1994), Nachbar (1997) and Dekel, Fudenberg and Levine (2004)) that dispense with the rational expectations about the other players' strategies and instead introduce other regularity conditions.[8]

We start by introducing an illustrative example with two periods in section 2. Section 3 provides a formal model, and section 4 presents the main results. Concluding remarks are in section 5, and all the proofs are collected in an appendix.

2 An illustrative example

Consider as an example a symmetric coordination game with two players and two actions.[9]

	I (invest)	N (not invest)
I (invest)	θ, θ	$\theta - 1, 0$
N (not invest)	$0, \theta - 1$	$0, 0$

(1)

If players have complete information about θ, I (resp. N) is a dominant strategy if $\theta > 1$ (resp. if $\theta < 0$), and (I, I) and (N, N) are two pure-strategy Nash equilibria if $\theta \in [0, 1]$. Following Harsanyi (1967), the game can be formulated with incomplete

[8]Regarding the literature on games with incomplete information, in which the common prior assumption is relaxed, following the seminal works by Schmeidler (1989) and Gilboa and Schmeidler (1989), papers study *ambiguity aversion* for games with incomplete information using either a maxmin expected utility or a Choquet expected utility with multiple priors (see Salo and Weber (1995), Lo (1998) and Chen, Katuscak and Ozdenoren (2007)), but this paper's focus is more on learning than on ambiguity aversion. See also Mertens and Zamir (1985) and Epstein and Wang (1996) for general conditions under which a state of types can be constructed.

[9]This static game is from Morris and Shin (2003). In nature, however, this model is related *neither* to games with complementarities *nor* to global games.

information about θ such that each player's *type* θ_i is independently and identically drawn (i. i. d.) from a differentiable population distribution F where F is *common knowledge*.[10] A (pure) *strategy* s_i of each player i of this incomplete-information game is a mapping from $\left[\underline{\theta}, \overline{\theta}\right]$ to $\{I, N\}$, and its equilibrium strategy is always parameterized with a threshold type k_i such that

$$s_i\left(\theta_i\right) = \begin{cases} I & \text{if } \theta_i > k_i, \\ N & \text{if } \theta_i < k_i. \end{cases}$$

The existence of a *symmetric* equilibrium s^* with a threshold $k^* \in \left(\underline{\theta}, \overline{\theta}\right)$ follows from the condition $k^* - F\left(k^*\right) = 0$.

Now, we present a two-period game that is based on a model in the next section. The (true) population distribution F is no longer common knowledge; F can be either F_a or F_b. For this example, we make the following simplifying assumptions:

(i) F_a strongly first-order stochastic dominates F_b; that is,

$$F_a\left(\theta\right) < F_b\left(\theta\right) \text{ for all } \theta \in \left(\underline{\theta}, \overline{\theta}\right).$$

(ii) The support $\left[\underline{\theta}, \overline{\theta}\right]$ has values $\underline{\theta} < 0$ and $1 < \overline{\theta}$.

(iii) Either F_a or F_b is uniform.

We would like to emphasize that in the next model section, apart from a common 2×2 game, each above assumption is substantially generalized. In particular, (i) is replaced with a weaker condition, a local stochastic dominance (A5), (ii) is a general support (A6), and (iii) is a non-uniform assumption (A7). In addition, a differentiable distribution is generalized to a continuous distribution (A4).

There are two periods: period 1 and period 2. In each period, every two players are randomly paired, and they play the game (1). Denote by π_{i1} player i's subjective probability in period 1 that the population distribution is F_a, and by k_1 the threshold in period 1. We let all the players have the same initial belief $\pi_1 = \pi_{i1} = \pi_{j1} \in (0, 1)$ for $i \neq j$. A strategy s_{i1} of each player i in period 1 of this unknown population

[10]If the low (column) player is player i (j), then θ in the first (second) entry changes to θ_i (θ_j) with the incomplete information.

game is a mapping from $[\underline{\theta}, \overline{\theta}]$ to $\{I, N\}$, and its symmetric equilibrium s_1^* is derived with a threshold $k_1^* \in (\underline{\theta}, \overline{\theta})$ such that

$$k_1^* - [\pi_1 F_a(k_1^*) + (1 - \pi_1) F_b(k_1^*)] = 0. \tag{2}$$

After they play the game, each player observes I or N from his or her opponent in the end of period 1. Their beliefs about F_a are updated using Bayes' rule. Denote by $\pi_2(\omega_i)$ player i's posterior probability in period 2 that the population distribution is F_a if player i has observed $\omega_i \in \{I, N\}$ in period 1. Since there may be multiple equilibria in the previous period, we assume that in the beginning of period 2, all players correctly expect an equilibrium threshold in period 1, k_1^*. Then, $\pi_2(\omega_i)$ can be derived as

$$\begin{aligned} \pi_2(I) &= \frac{(1 - F_a(k_1^*))\pi_1}{(1 - F_a(k_1^*))\pi_1 + (1 - F_b(k_1^*))(1 - \pi_1)}, \\ \pi_2(N) &= \frac{F_a(k_1^*)\pi_1}{F_a(k_1^*)\pi_1 + F_b(k_1^*)(1 - \pi_1)}. \end{aligned} \tag{3}$$

A strategy s_{i2} of each player i in period 2 may depend not only on the player's *type* θ_i but also on *experience* ω_i from the previous period, so s_{i2} is a mapping from $[\underline{\theta}, \overline{\theta}] \times \{I, N\}$ to $\{I, N\}$, given by

$$s_{i2}(\theta_i, \omega_i) = \begin{cases} I & \text{if } \theta_i > k_{i2}(\omega_i), \\ N & \text{if } \theta_i < k_{i2}(\omega_i). \end{cases}$$

Denote by G the probability that player j plays N in period 2.

$$\begin{aligned} & G(\omega_i, k_1^*, k_{j2}(I), k_{j2}(N)) \\ & \equiv \pi_2(\omega_i) A(k_1^*, k_{j2}(I), k_{j2}(N)) + (1 - \pi_2(\omega_i)) B(k_1^*, k_{j2}(I), k_{j2}(N)), \end{aligned} \tag{4}$$

where

$$\begin{aligned} A(k_1^*, k_{j2}(I), k_{j2}(N)) &\equiv (1 - F_a(k_1^*)) F_a(k_{j2}(I)) + F_a(k_1^*) F_a(k_{j2}(N)), \\ B(k_1^*, k_{j2}(I), k_{j2}(N)) &\equiv (1 - F_b(k_1^*)) F_b(k_{j2}(I)) + F_b(k_1^*) F_b(k_{j2}(N)). \end{aligned}$$

$A(k_1^*, k_{j2}(I), k_{j2}(N))$ is the probability that player j chooses N if the population distribution is F_a, and $B(k_1^*, k_{j2}(I), k_{j2}(N))$ is the probability if the population distribution is F_b. The probability that player j has observed I (resp. N) in period 1 is $(1 - F_a(k_1^*))$ (resp. $F_a(k_1^*)$) if the population distribution is F_a, which event can be realized with a probability $\pi_2(\omega_i)$ *from player i's perspective* in period 2. The same argument applies for the event that the population distribution is F_b, which can be realized with a probability $1 - \pi_2(\omega_i)$.

Player i's payoff with a characteristic (θ_i, ω_i) from choosing I given player j's strategy with $(k_{j2}(I), k_{j2}(N))$ is $\theta_i - G(\omega_i, k_1^*, k_{j2}(I), k_{j2}(N))$, and player i's payoff from choosing N is always zero. The following proposition shows the existence of a symmetric equilibrium with a pair $(k_2^*(I), k_2^*(N)) \in (\underline{\theta}, \overline{\theta})^2$, satisfying

$$k_2^*(\omega) = G(\omega, k_1^*, k_2^*(I), k_2^*(N)) \text{ for each } \omega \in \{I, N\}, \tag{5}$$

and $k_2^*(I) < k_2^*(N)$, which confirms our intuition. If player i has observed I before, then the player has a posterior probability $\pi_2(I) > \pi_2(N)$, which in turn implies that a stochastic dominant distribution F_a is more likely than F_b. Hence, he or she anticipates more high types in the population, which results in a lower threshold $k_2^*(I) < k_2^*(N)$, meaning that more types of player i choose to invest.

Proposition 1 *A symmetric equilibrium in period* 2 *exists, and it satisfies* $k_2^*(I) < k_2^*(N)$.

This allows us new interpretations about outcomes of games with incomplete information, which was not previously possible in Harsanyi's model with a common population distribution. Let player i's type θ_i be between $k_2^*(I)$ and $k_2^*(N)$. Then the player chooses to invest given a *positive* outcome I in the past and to not invest given a *negative* outcome N in the past. Hence, the same type, which is determined in the beginning of a game, can make different decisions depending on his or her past experience. We introduce a formal model in the next section.

3 Model

Let time be discrete and infinite, indexed by $t \in \mathbb{N}$. In each period t, every two players in a large population are randomly matched, and play a symmetric normal form game[11] with two actions $\{\alpha, \beta\}$. Each player's payoff is $u : \{\alpha, \beta\}^2 \to \mathbb{R}$ and, as in Carlsson and van Damme (1993), a 2×2 game is parameterized with the pair (d_α, d_β):

$$d_\alpha \equiv u(\alpha, \alpha) - u(\beta, \alpha),$$
$$d_\beta \equiv u(\beta, \beta) - u(\alpha, \beta),$$

where d_α (resp. d_β) denotes the loss that player i incurs when player i unilaterally deviates from the action profile (α, α) (resp. (β, β)).

A 2×2 game can be formulated with incomplete information such that the pair $(d_\alpha(\theta_i), d_\beta(\theta_i))$ depends on the realization of θ_i. Players know that the type θ_i of each player i is i. i. d from a (true) population distribution, and it is either a CDF $F_a : \Theta \to [0,1]$ or a CDF $F_b : \Theta \to [0,1]$, where its support Θ is an interval in $\mathbb{R}$, including a unbounded one, but they do not know which one is the population distribution.

For the class of functions (d_α, d_β), we assume that

(A1) *For each $\gamma = \alpha, \beta$, d_γ is continuous.*

(A2) *For all $\theta_i \in \Theta$, $d_\alpha(\theta_i) + d_\beta(\theta_i) > 0$.*[12]

Define $D : \Theta \to \mathbb{R}$ as

$$D(\theta_i) = \frac{d_\alpha(\theta_i)}{d_\alpha(\theta_i) + d_\beta(\theta_i)}.$$

By (A1)-(A2), the function D is well defined, and continuous.

(A3) *D is a strictly increasing function.*[13]

[11] The framework of this paper is non-cooperative. If it is cooperative so that players share their information about their types, then there is no incentive for them to estimate a population distribution.

[12] If $d_\alpha(\theta_i) + d_\beta(\theta_i) = 0$, player i's payoff does not depend on his opponent's strategies given θ_j and ω_j. See (6) below.

[13] Note that although (A3) can be replaced by a weaker assumption, "D is a strictly monotonic function," we keep this, since it will only reverse the inequality sign of the equilibrium strategy below. The complete proof for the case with "D is a strictly *decreasing* function" is available from the author upon request.

In each period t, after a 2×2 game is played, each player's belief about the population distribution is updated by Bayes' rule using his or her opponent's action. Denote by Ω_t the set of histories up to period $t-1$ and ω_t an element of the set Ω_t. In particular, $\Omega_1 = \varnothing$, and $\Omega_t = \{\omega_t^1, \omega_t^2, ..., \omega_t^{2^{t-1}}\}$ has 2^{t-1} elements for $t = 2, 3, 4....$[14] We assume that for any two players $i \neq j$, they have the same *initial belief*, $\pi_1 \in (0, 1)$ to focus on a symmetric equilibrium in this paper.[15] Let $\pi_t(\omega_{it})$ denote player i's posterior probability that the population distribution is F_a, where the experience ω_{it} is player i's history of observations up to period $t-1$ for $t = 2, 3, 4....$

We call a vector of player i's type and experience, (θ_i, ω_{it}), player i's *characteristic*. For period t, denote by $v_t(\theta_i, \omega_{it}, A_{jt}, B_{jt})$ *the net gain of player i with a characteristic* (θ_i, ω_{it}) *from choosing* α (rather than β) when player i expects that player j chooses β with probability A_{jt} if the population distribution is F_a, and that player j chooses β with probability B_{jt} if the population distribution is F_b.[16] Then, v_t is given as[17]

$$v_t(\theta_i, \omega_{it}, A_{jt}, B_{jt}) \equiv (d_\alpha(\theta_i) + d_\beta(\theta_i))[D(\theta_i) - (\pi_t(\omega_{it})A_{jt} + (1 - \pi_t(\omega_{it}))B_{jt})]. \quad (6)$$

If $v_t(\theta_i, \omega_{it}, A_{jt}, B_{jt}) > 0$ (resp. < 0), it is optimal for player i with a characteristic (θ_i, ω_{it}) to choose α (resp. β). First, we show that player i's equilibrium strategy is

[14] For example, $\Omega_2 = \{\alpha, \beta\}$ and $\Omega_3 = \{(\alpha, \alpha), (\alpha, \beta), (\beta, \alpha), (\beta, \beta)\}$.

[15] The game has to start with certain initial subject beliefs of players in the population. Alternatively, we can assume a common distribution of the initial beliefs. However, this will be awkward. To replace a common distribution of types with a weaker assumption in this paper, we would then have introduced another common distribution, a distribution of the initial beliefs. We would like to emphasize here that only in the first period do players have this common belief, and that after the first period, they will have different beliefs depending on their observations.

[16] Of course, both A_{jt} and B_{jt} depend on player j's strategies, but before we show in Lemma 1 that a cut-off strategy is an equilibrium strategy for the game, we keep those as these "reduced forms." Later, A_{jt} and B_{jt} will be derived precisely in (9).

[17] v_t is from $[1 - (\pi_t(\omega_{it})A_{jt} + (1 - \pi_t(\omega_{it}))B_{jt})]d_\alpha(\theta_i) - (\pi_t(\omega_{it})A_{jt} + (1 - \pi_t(\omega_{it}))B_{jt})d_\beta(\theta_i)$.

a mapping $s_{it} : \Theta \times \Omega_t \rightarrow \{\alpha, \beta\}$ such that

$$s_{it}(\theta_i, \omega_{it}) = \begin{cases} \alpha & \text{if } \theta_i > k_{it}(\omega_{it}), \\ \beta & \text{if } \theta_i < k_{it}(\omega_{it}), \end{cases} \tag{7}$$

where $k_{it}(\omega_{it})$ is player i's threshold type given $\omega_{it} \in \Omega_t$, in which "high" types choose α. It follows from the result that we simply need to examine threshold types for the search of an equilibrium.

Lemma 1 *Under (A1)-(A3), for each i and every $\omega_{it} \in \Omega_t$, any equilibrium strategy s_{it} of player i satisfies (7).*

If the probability that player j chooses β strictly increases, player i's optimal response is to increase his or her threshold, which results in more "low" types choosing β.[18]

For the class of distribution functions (F_a, F_b), we assume that

(A4) *For each $n = a, b$, F_n is continuous.*

(A5) *There exists a subinterval Γ of Θ such that*

$$\Gamma \equiv \{\theta \in \Theta \mid F_a(\theta) < F_b(\theta)\}.$$

(A5) is a *local* stochastic dominance relationship between F_a and F_b. *If F_n were* known as the population distribution, then given player j's threshold k_{jn}, the net gain of player i with a characteristic (θ_i, ω_{it}) from choosing α is $(d_\alpha(\theta_i) + d_\beta(\theta_i))[D(\theta_i) - F_n(k_{jn})]$, and in Lemma 2, we show that any equilibrium is symmetric.

Lemma 2 *Under (A1)-(A5), for each $n = a, b$, if F_n were known as the population distribution, then any equilibrium is symmetric, that is, $k_{in} = k_{jn}$ for $i \neq j$.*

[18] This does *not* mean that in general, the model is related to strategic complementarities. The strategic complementarity of payoffs requires either $d_\alpha(\theta) \geq d_\beta(\theta)$ or $d_\alpha(\theta) \leq d_\beta(\theta)$ for all θ, but none of them is necessary for the condition that D is strictly monotonic.

With assumption (A6) followed by Lemma 2, we show in section 4 that each player's best response, in terms of the player's threshold, has a value between k_a and k_b, which results in the existence of a pure strategy equilibrium in each period.[19]

(A6) *For each* $n = a, b$*, if* F_n *were known as the population distribution, then there exists an equilibrium such that each player's threshold type is an interior point of* Θ*.*

Lemma 2 and (A6) entail that for each $n = a, b$, there exists k_n such that

$$D(k_n) - F_n(k_n) = 0. \tag{8}$$

The last assumption (A7) guarantees an intuitively reasonable outcome $k_a < k_b$ for $k_a, k_b \in \Gamma$. Since D is strictly increasing, it is natural that the stochastic dominant distribution F_a on Γ has a lower threshold if it were known as the population distribution, and their thresholds fall in the area where two distributions are distinct. Figure 1 illustrates that with a strictly increasing D, *both* $k_a > k_b$ *and* $k_a < k_b$ are possible.

(A7) *There exists at least one* $n \in \{a, b\}$ *such that for any pair* $\theta' > \theta$ on Γ,

$$D\left(\theta'\right) - D\left(\theta\right) \geq F_n\left(\theta'\right) - F_n\left(\theta\right).$$

Lemma 3 *Under (A1)-(A7), if* $k_a, k_b \in \Gamma$*,* $k_a < k_b$*.*

Let $\mathbf{k}_{it} = (k_{it}(\omega_t))_{\omega_t \in \Omega_t}$ be a collection of player i's thresholds for each possible experience ω_t, and h_t^* be an equilibrium history up to $t-1$. We assume that each player correctly expects *the same* h_t^* since there can multiple equilibria in a period. Then, the probability that player j chooses β in period t, G_t, is derived as

$$G_t(\omega_{it}, h_t^*, \mathbf{k}_{jt}) \equiv \pi_t(\omega_{it}) A_t(h_t^*, \mathbf{k}_{jt}) + (1 - \pi_t(\omega_{it}))\, B_t(h_t^*, \mathbf{k}_{jt}),$$

where

$$A_t(h_t^*, \mathbf{k}_{jt}) \equiv \textstyle\sum_{\omega_t \in \Omega_t} \Pr\left(\omega_t | F = F_a, h_t^*\right) F_a(k_{jt}(\omega_t)), \tag{9}$$

$$B_t(h_t^*, \mathbf{k}_{jt}) \equiv \textstyle\sum_{\omega_t \in \Omega_t} \Pr\left(\omega_t | F = F_b, h_t^*\right) F_b(k_{jt}(\omega_t)).$$

[19] In the example, a special type of support, a compact interval $\left[\underline{\theta}, \overline{\theta}\right]$ with $\underline{\theta} < 0$ and $1 < \overline{\theta}$, works as a sufficient condition to present the existence of a pure strategy equilibrium in the second period.

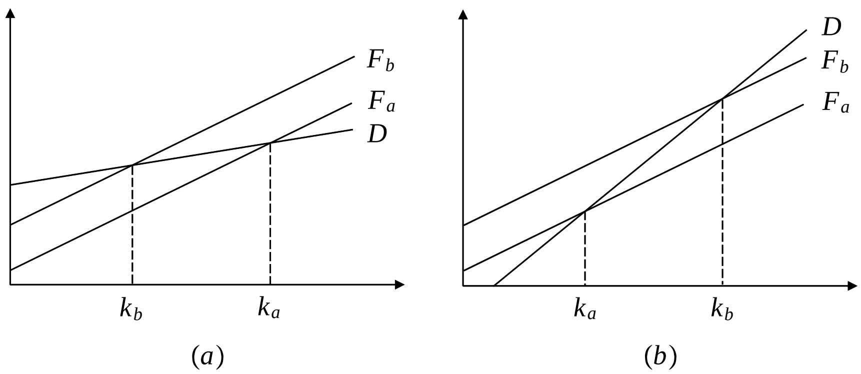

Figure 1: (A7) is violated in (a) and satisfied in (b).

$A_t(h_t^*, \mathbf{k}_{jt})$ is the probability that player j chooses β if the population distribution is F_a, and $B_t(h_t^*, \mathbf{k}_{jt})$ is the probability that player j chooses β if the population distribution is F_b. By (6) and (A2), we define the net gain of player i with a characteristic (θ_i, ω_{it}) from choosing α as V_t:

$$V_t(\theta_i, \omega_{it}, h_t^*, \mathbf{k}_{jt}) \equiv D(\theta_i) - G_t(\omega_{it}, h_t^*, \mathbf{k}_{jt}). \tag{10}$$

Lemma 4 establishes that given each $\omega_t \in \Omega_t$, there exists player i's best response function in terms of a threshold $k_{it}(\omega_t)$ between k_a and k_b.

Lemma 4 *Under (A1)-(A7), given each $\omega_t \in \Omega_t$ and every $\mathbf{k}_{jt} \in [k_a, k_b]^{2^{t-1}}$, player i's best response function exists, which we denote by*

$$k_{it}(\omega_t) = \phi_t(\omega_t, \mathbf{k}_{jt}) \in [k_a, k_b].$$

Then, a symmetric equilibrium in period t is $\mathbf{k}_t^* = (k_t^*(\omega_t))_{\omega_t \in \Omega_t}$ such that for each $\omega_t \in \Omega_t$,

$$V_t(k_t^*(\omega_t), \omega_t, h_t^*, \mathbf{k}_t^*) = 0.$$

By constructing the best response in terms of a threshold, we can transpose a finite action game to a continuous threshold game, and thus we apply a fixed point theorem to find out a pure strategy equilibrium in the typical way.

4 The main results

This section provides three main results of the paper.

4.1 Existence

We present the first main result, the existence of a symmetric equilibrium, by combining two Lemmas below. Lemma 5's existence result holds under $\pi_t(\omega_t) \in (0,1)$.[20]

Lemma 5 *Under (A1)-(A7), for each period t, if $k_a, k_b \in \Gamma$ and $\pi_t(\omega_t) \in (0,1)$, then there exists a symmetric equilibrium such that $k_t^*(\omega_t) \in (k_a, k_b)$ for all $\omega_t \in \Omega_t$.*

If $A_t(h_t^*, \mathbf{k}_t^*)\pi_t + B_t(h_t^*, \mathbf{k}_t^*)(1-\pi_t) \in (0,1)$, for each $\omega_t \in \Omega_t$, the posterior $\pi_{t+1}(\omega_{t+1})$ is derived using Bayes' rule: if α is observed in period t,

$$\pi_{t+1}(\omega_{t+1}) = \frac{(1 - A_t(h_t^*, \mathbf{k}_t^*))\pi_t(\omega_t)}{(1 - A_t(h_t^*, \mathbf{k}_t^*))\pi_t(\omega_t) + (1 - B_t(h_t^*, \mathbf{k}_t^*))(1 - \pi_t(\omega_t))}, \tag{11}$$

and if β is observed in period t,

$$\pi_{t+1}(\omega_{t+1}) = \frac{A_t(h_t^*, \mathbf{k}_t^*)\pi_t(\omega_t)}{A_t(h_t^*, \mathbf{k}_t^*)\pi_t(\omega_t) + B_t(h_t^*, \mathbf{k}_t^*)(1 - \pi_t(\omega_t))}. \tag{12}$$

For the next Lemma, we replace (A4) with (A4′) which requires that F_n's density function is everywhere positive on Γ.

(A4′) *For each $n = a, b$, F_n is continuous and has its density function $f_n(\theta) > 0$ for all θ on Γ.*

Lemma 6 shows that under $k_t^*(\omega_t) \in (k_a, k_b)$, for each $\pi_t(\omega_t) \in (0,1)$ with $\omega_t \in \Omega_t$, we have $\pi_{t+1}(\omega_{t+1}) \in (0,1)$ regardless of whether α is observed or β is observed in period t.

Lemma 6 *Under (A1)-(A3) and (A4′)-(A7), if $k_a, k_b \in \Gamma$ and $k_t^*(\omega_t) \in (k_a, k_b)$, then for each $\pi_t(\omega_t) \in (0,1)$ with $\omega_t \in \Omega_t$, $\pi_{t+1}(\omega_{t+1}) \in (0,1)$ for all $\omega_{t+1} \in \{\alpha, \beta\} \times \{\omega_t\}$.*

[20] Even without $\pi_t(\omega_t) \in (0,1)$, the proof of Lemma 5 establishes the existence of a symmetric equilibrium $k_t^*(\omega_t) \in [k_a, k_b]$ for all $\omega_t \in \Omega_t$, that is, $k_t^*(\omega_t)$ can be k_a or k_b.

Since $\pi_1 \in (0,1)$ in the first period, using a recursive method, it is immediate from Lemmas 5 and 6 that for each t and every $\omega_t \in \Omega_t$, the equilibrium threshold $k_t^*(\omega_t)$ falls within (k_a, k_b) and $\pi_t(\omega_t) \in (0,1)$, which is summarized in the following Theorem.

Theorem 1 *Under (A1)-(A3) and (A4′)-(A7), if $k_a, k_b \in \Gamma$, then for each t and every $\omega_t \in \Omega_t$, there exists a symmetric equilibrium such that $k_t^*(\omega_t) \in (k_a, k_b)$, and furthermore, $\pi_t(\omega_t) \in (0,1)$.*

For each t and every $\omega_t \in \Omega_t$, we have an "interior belief" $\pi_t(\omega_t) \in (0,1)$, which plays an important role to prove next two main results.

4.2 Monotonicity

The second main result of this paper shows the monotonicity result: if $k_a, k_b \in \Gamma$, then for any two different experiences such that one's belief about a stochastic dominant distribution on Γ is more likely than the other, that is, $\pi_t(\omega_t') > \pi_t(\omega_t)$, the former induces a lower threshold such as $k_t^*(\omega_t') < k_t^*(\omega_t)$.

Theorem 2 *Under (A1)-(A3) and (A4′)-(A7), if $k_a, k_b \in \Gamma$, then for any pair $\omega_t', \omega_t \in \Omega_t$ such that $\pi_t(\omega_t') > \pi_t(\omega_t)$,*

$$k_t^*(\omega_t') < k_t^*(\omega_t).$$

When player i has a posterior probability $\pi_t(\omega_t') > \pi_t(\omega_t)$, player i believes that F_a is more likely than F_b. Lemma 1 shows that high types choose α, and from the stochastic dominance relationship between F_a and F_b, F_a has more high types on Γ. Hence, player i expects that his or her opponent's probability of choosing α is greater. It follows that player i's optimal response is to make the own probability of choosing α strictly increase. As a result, player i has a lower threshold, $k_t^*(\omega_t') < k_t^*(\omega_t)$.

The proof of Theorem 2 also shows that if $k_a, k_b \in \Gamma$, for each t and every $h_t^*, \mathbf{k}_t^*$,

$$A_t(h_t^*, \mathbf{k}_t^*) < B_t(h_t^*, \mathbf{k}_t^*), \tag{13}$$

which entails that in each equilibrium in every period, the probability that each player's opponent plays β if the population distribution is F_b is greater than the probability if the population distribution is F_a.

4.3 Convergence

The last main result of this paper provides the convergence. We show that each player's initial belief about the true distribution converges *almost surely* to a correct belief. Denote by π_t the random variable, the probability that the population distribution is F_a in period t and $\omega(t)$ the observation *in* period t. Then, it follows from Theorem 1, (11) and (12) that the relationship between π_t and π_{t+1} can be written as below:

$$\frac{\pi_{t+1}}{1-\pi_{t+1}} = \left[\frac{(1-A_t(h_t^*, \mathbf{k}_t^*))}{(1-B_t(h_t^*, \mathbf{k}_t^*))}\right]^{\mathbf{1}_{\{\alpha\}}(\omega(t))} \left[\frac{A_t(h_t^*, \mathbf{k}_t^*)}{B_t(h_t^*, \mathbf{k}_t^*)}\right]^{\mathbf{1}_{\{\beta\}}(\omega(t))} \frac{\pi_t}{1-\pi_t}, \tag{14}$$

where $\mathbf{1}_{\{\alpha\}}$ is an indicator function. By the proof of Theorem 2, for each t and every $h_t^*, \mathbf{k}_t^*$, we have $A_t(h_t^*, \mathbf{k}_t^*) < B_t(h_t^*, \mathbf{k}_t^*)$ in (13), which in turn implies that

$$\begin{aligned} \pi_{t+1} > \pi_t \quad &\text{if} \quad \alpha \text{ is observed in period } t, \\ \pi_{t+1} < \pi_t \quad &\text{if} \quad \beta \text{ is observed in period } t, \end{aligned}$$

that is, if α is observed, the probability that the population distribution is F_a is greater. Thus, in order to study the dynamic behavior and the convergence of π_t, the main question is whether α is observed more "frequently" along the equilibrium path if F_a is the population distribution. Note that the random variables in (14) are not i. i. d., so the strong law of large numbers cannot be used to determine the long-run behavior.

We show, using the martingale convergence theorem, in Theorem 3 that each player's initial belief π_1 about the population distribution converges *almost surely* to a correct belief.

Theorem 3 *Under (A1)-(A3) and (A4′)-(A7), if $k_a, k_b \in \Gamma$, then for any $\pi_1 \in (0,1)$, π_t converges almost surely to a correct belief.*

As π_t converges almost surely to a correct belief, each player's sequence of equilibrium strategies converges almost surely to the equilibrium strategy when the population distribution is common knowledge.

5 Concluding remarks

We analyze a dynamic Bayesian game with an unknown population distribution for a 2×2 game. It is shown that each player's initial belief about a population distribution converges almost surely to a correct belief. However, in reality, we can only have a large but *finite* number of matchings, and people are sometimes matched in a non-random manner, so each "sample path" will influence the decisions of players in game-theoretic situations with an unknown population distribution.

In a subsequent work, a model with continuous actions can be introduced to examine conditions under which we can obtain similar results.

Appendix: proofs

Proof of Proposition 1. Recall that for the example, we assume that the support $\left[\underline{\theta}, \overline{\theta}\right]$ has values $\underline{\theta} < 0$ and $1 < \overline{\theta}$, and that either F_a or F_b is uniform. First, we show the existence. Denote

$$\mathcal{G}\left(k_2\left(I\right), k_2\left(N\right)\right) \equiv \left(G\left(I, k_1^*, k_2\left(I\right), k_2\left(N\right)\right), G\left(N, k_1^*, k_2\left(I\right), k_2\left(N\right)\right)\right).$$

It follows from $\underline{\theta} < 0$ and $1 < \overline{\theta}$ that $\mathcal{G}$ can be defined as a continuous mapping from $\left[\underline{\theta}, \overline{\theta}\right]^2$ to $\left[\underline{\theta}, \overline{\theta}\right]^2$, and that a fixed point $\left(k_2^*\left(I\right), k_2^*\left(N\right)\right) \in \left(\underline{\theta}, \overline{\theta}\right)^2$ exists such that $\mathcal{G}\left(k_2^*\left(I\right), k_2^*\left(N\right)\right) = \left(k_2^*\left(I\right), k_2^*\left(N\right)\right)$. Now, we suppose $k_2^*\left(I\right) \geq k_2^*\left(N\right)$. By (4) and (5),

$$\begin{aligned} & k_2^*\left(I\right) - k_2^*\left(N\right) \qquad (15) \\ = \; & G\left(I, k_1^*, k_2^*\left(I\right), k_2^*\left(N\right)\right) - G\left(N, k_1^*, k_2^*\left(I\right), k_2^*\left(N\right)\right) \\ = \; & \left(\pi_2\left(I\right) - \pi_2\left(N\right)\right)\left[A\left(k_1^*, k_2^*\left(I\right), k_2^*\left(N\right)\right) - B\left(k_1^*, k_2^*\left(I\right), k_2^*\left(N\right)\right)\right] \\ \leq \; & \left(\pi_2\left(I\right) - \pi_2\left(N\right)\right)\left[F_a(k_2^*\left(I\right)) - F_b(k_2^*\left(N\right))\right], \end{aligned}$$

where the last inequality follows from:

$$\begin{aligned}
& A\left(k_1^*, k_2^*(I), k_2^*(N)\right) - B\left(k_1^*, k_2^*(I), k_2^*(N)\right) \\
= \; & \left(1 - F_a\left(k_1^*\right)\right) F_a(k_2^*(I)) + F_a\left(k_1^*\right) F_a(k_2^*(N)) \\
& - \left[\left(1 - F_b\left(k_1^*\right)\right) F_b(k_2^*(I)) + F_b\left(k_1^*\right) F_b(k_2^*(N))\right] \\
\leq \; & \left(1 - F_a\left(k_1^*\right)\right) F_a(k_2^*(I)) + F_a\left(k_1^*\right) F_a(k_2^*(I)) \\
& - \left[\left(1 - F_b\left(k_1^*\right)\right) F_b(k_2^*(N)) + F_b\left(k_1^*\right) F_b(k_2^*(N))\right] \\
= \; & F_a(k_2^*(I)) - F_b(k_2^*(N)).
\end{aligned}$$

Since $\pi_1 \in (0,1)$ and $k_1^* \in (0,1)$, we have $\pi_2(I), \pi_2(N) \in (0,1)$ in (3), which implies $\pi_2(I) - \pi_2(N) \in (0,1)$. If $k_2^*(I) = k_2^*(N)$, then $0 \leq (\pi_2(I) - \pi_2(N))\left[F_a(k_2^*(I)) - F_b(k_2^*(I))\right] < 0$, which is a contradiction. Let $k_2^*(I) > k_2^*(N)$, and WLOG, F_b be uniform. The first-order stochastic dominance relationship entails

$$\begin{aligned}
k_2^*(I) - k_2^*(N) & \leq (\pi_2(I) - \pi_2(N))\left[F_a(k_2^*(I)) - F_b(k_2^*(N))\right] \\
& < (\pi_2(I) - \pi_2(N))[F_b(k_2^*(I)) - F_b(k_2^*(N))].
\end{aligned}$$

From the Mean Value Theorem, there exists $c \in (k_2^*(N), k_2^*(I))$ such that

$$k_2^*(I) - k_2^*(N) < (\pi_2(I) - \pi_2(N))[k_2^*(I) - k_2^*(N)]F_b'(c).$$

Then, $1 < (\pi_2(I) - \pi_2(N))F_b'(c)$. Since F_b is uniform, and $\overline{\theta} - \underline{\theta} > 1$,

$$1 < (\pi_2(I) - \pi_2(N))F_b'(c) < 1,$$

a contradiction. Thus, $k_2^*(I) < k_2^*(N)$. ■

Proof of Lemma 1. First, any equilibrium strategy must be a cut-off strategy for each $\omega_{it} \in \Omega_t$ since D is strictly increasing, and $\pi_t(\omega_{it})A_{jt} + (1 - \pi_t(\omega_{it}))B_{jt}$ does not depend on θ_i. Suppose that there exist player i's equilibrium strategy s_{it} and $\omega_{it} \in \Omega_t$ such that for a pair $\theta_i' > \theta_i$, $s_{it}(\theta_i', \omega_{it}) = \beta$ and $s_{it}(\theta_i, \omega_{it}) = \alpha$. From (A3),

$$(d_\alpha(\theta_i) + d_\beta(\theta_i))\left[D(\theta_i') - (\pi_t(\omega_{it})A_{jt} + (1 - \pi_t(\omega_{it}))B_{jt})\right] > v_t(\theta_i, \omega_{it}, A_{jt}, B_{jt}).$$

$s_{it}(\theta_i, \omega_{it}) = \alpha$ implies $v_t(\theta_i, \omega_{it}, A_{jt}, B_{jt}) \geq 0$. Then,

$$(d_\alpha(\theta_i) + d_\beta(\theta_i)) \left[D(\theta_i') - (\pi_t(\omega_{it}) A_{jt} + (1 - \pi_t(\omega_{it})) B_{jt}) \right] > 0.$$

By (A2), $v_t(\theta_i', \omega_{it}, A_{jt}, B_{jt}) > 0$. $s_{it}(\theta_i', \omega_{it}) = \alpha$ is optimal, and we have a contradiction. ■

Proof of Lemma 2. Suppose there exists an asymmetric equilibrium. WLOG let $k_{in} > k_{jn}$. Hence, k_{jn} is either an interior point or the infimum of Θ, given player i's strategy k_{in}, so player j's payoff satisfies

$$0 \leq (1 - F_n(k_{in})) d_\alpha(k_{jn}) - F_n(k_{in}) d_\beta(k_{jn}) = (d_\alpha(k_{jn}) + d_\beta(k_{jn}))[D(k_{jn}) - F_n(k_{in})].$$

From (A2) and (A3),

$$D(k_{in}) - F_n(k_{jn}) \geq D(k_{in}) - F_n(k_{in}) > D(k_{jn}) - F_n(k_{in}) \geq 0,$$

which yields

$$(d_\alpha(k_{in}) + d_\beta(k_{in})) [D(k_{in}) - F_n(k_{jn})] > 0. \tag{16}$$

On the other hand, k_{in} is either an interior point or the supremum of Θ, given player j's strategy k_{jn}, satisfying

$$0 \geq (1 - F_n(k_{jn})) d_\alpha(k_{in}) - F_n(k_{jn}) d_\beta(k_{in}) = (d_\alpha(k_{in}) + d_\beta(k_{in}))[D(k_{in}) - F_n(k_{jn})],$$

which is a contradiction with (16). ■

Proof of Lemma 3. Suppose $k_a \geq k_b$. By Lemma 2, we have

$$D(k_a) - D(k_b) = F_a(k_a) - F_b(k_b).$$

If $k_a = k_b$, $0 = F_a(k_a) - F_b(k_a) < 0$, a contradiction. If $k_a > k_b$, WLOG let F_a satisfy (A7). Then,

$$D(k_a) - D(k_b) = F_a(k_a) - F_b(k_b) < F_a(k_a) - F_a(k_b),$$

which is a contradiction with (A7). ■

Proof of Lemma 4. Let $\mathbf{k}_t^a = \underbrace{(k_a, ..., k_a)}_{2^{t-1}\text{ times}}$ and $\mathbf{k}_t^b = \underbrace{(k_b, ..., k_b)}_{2^{t-1}\text{ times}}$. Fix $\omega_t \in \Omega_t$. For $\mathbf{k}_{jt} \geq \mathbf{k}_t^a$,

$$\begin{aligned} V_t(k_a, \omega_t, h_t^*, \mathbf{k}_{jt}) &\leq V_t(k_a, \omega_t, h_t^*, \mathbf{k}_t^a) = D(k_a) - G_t(\omega_t, h_t^*, \mathbf{k}_t^a) \\ &= D(k_a) - [\pi_t(\omega_t) F_a(k_a) + (1 - \pi_t(\omega_t)) F_b(k_a)] \\ &\leq D(k_a) - F_a(k_a) = 0. \end{aligned}$$

The last inequality follows from the first-order stochastic dominance with $\pi_t(\omega_t) \in [0, 1]$. For $\mathbf{k}_{jt} \leq \mathbf{k}_t^b$,

$$\begin{aligned} V_t(k_b, \omega_t, h_t^*, \mathbf{k}_{jt}) &\geq V_t(k_b, \omega_t, h_t^*, \mathbf{k}_t^b) = D(k_b) - G_t(\omega_t, h_t^*, \mathbf{k}_t^b) \\ &= D(k_b) - [\pi_t(\omega_t) F_a(k_b) + (1 - \pi_t(\omega_t)) F_b(k_b)] \\ &\geq D(k_b) - F_b(k_b) = 0. \end{aligned}$$

The last inequality follows from the first-order stochastic dominance with $\pi_t(\omega_t) \in [0, 1]$. We conclude that for each $\mathbf{k}_{jt} \in [k_a, k_b]^{2^{t-1}}$,

$$V_t(k_a, \omega_t, h_t^*, \mathbf{k}_{jt}) \leq 0 \text{ and } V_t(k_b, \omega_t, h_t^*, \mathbf{k}_{jt}) \geq 0.$$

Since V_t is a continuous and strictly increasing function of θ_i, there exists a unique function $\varphi_t(\omega_t, h_t^*, \mathbf{k}_{jt}) \in [k_a, k_b]$ such that $V_t(\varphi_t(\omega_t, h_t^*, \mathbf{k}_{jt}), \omega_t, h_t^*, \mathbf{k}_{jt}) = 0$. This establishes the result. ■

Proof of Lemma 5. Define $\Phi_t(\mathbf{k}_t) = (\phi_t(\omega_t, \mathbf{k}_t))_{\omega_t \in \Omega_t}$. From the proof of Lemma 4, for each $\omega_t \in \Omega_t$ and every $\mathbf{k}_t \in [k_a, k_b]^{2^{t-1}}$, $\phi_t(\omega_t, \mathbf{k}_t) \in [k_a, k_b]$. Hence, Φ_t is a continuous mapping from $[k_a, k_b]^{2^{t-1}}$ to $[k_a, k_b]^{2^{t-1}}$. By Brouwer's Fixed Point Theorem, there exists $\mathbf{k}_t^*$ such that $\Phi_t(\mathbf{k}_t^*) = \mathbf{k}_t^* \in [k_a, k_b]^{2^{t-1}}$. Suppose there exists $\omega_t \in \Omega_t$ such that either $k_t^*(\omega_t) = k_a$ or $k_t^*(\omega_t) = k_b$. WLOG, let $k_t^*(\omega_t) = k_a$, which implies that $V_t(k_a, \omega_t, h_t^*, \mathbf{k}_t^*) = 0$. Since $\mathbf{k}_t^* \geq \mathbf{k}_t^a$ and $\pi_t(\omega_t) \in (0, 1)$,

$$\begin{aligned} V_t(k_a, \omega_t, h_t^*, \mathbf{k}_t^*) &\leq V_t(k_a, \omega_t, h_t^*, \mathbf{k}_t^a) = D(k_a) - G_t(\omega_t, h_t^*, \mathbf{k}_t^a) \\ &= D(k_a) - [\pi_t(\omega_t) F_a(k_a) + (1 - \pi_t(\omega_t)) F_b(k_a)] \\ &< D(k_a) - [\pi_t(\omega_t) F_a(k_a) + (1 - \pi_t(\omega_t)) F_a(k_a)] \\ &= D(k_a) - F_a(k_a) = 0, \end{aligned}$$

where the strict inequality follows from $F_a(k_a) < F_b(k_a)$, and we have a contradiction. ■

Proof of Lemma 6. Given $k_t^*(\omega_t) \in (k_a, k_b)$, for $n = a, b$,

$$0 \leq F_n(k_a) < F_n\left(k_t^*(\omega_t)\right) < F_n(k_b) \leq 1,$$

which in turn implies that $A_t(h_t^*, \mathbf{k}_t^*) \in (0,1)$ and $B_t(h_t^*, \mathbf{k}_t^*) \in (0,1)$, and the result follows from the condition, $\pi_t(\omega_t) \in (0,1)$. ■

Proof of Theorem 2. In an equilibrium, for any pair $\omega_t', \omega_t \in \Omega_t$, by (10),

$$\begin{aligned} D(k_t(\omega_t')) - D(k_t(\omega_t)) &= G_t(\omega_t', h_t^*, \mathbf{k}_t^*) - G_t(\omega_t, h_t^*, \mathbf{k}_t^*) \qquad (17) \\ &= (\pi_t(\omega_t') - \pi_t(\omega_t))[A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*)]. \end{aligned}$$

Note that $A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*)$ does not depend on ω_t. Suppose that there exists a pair $\omega_t', \omega_t \in \Omega_t$ such that $\pi_t(\omega_t') > \pi_t(\omega_t)$ and $k_t^*(\omega_t') \geq k_t^*(\omega_t)$. We divide the proof into two cases.

Case 1. Suppose that there exists a pair $\omega_t', \omega_t \in \Omega_t$ such that $\pi_t(\omega_t') > \pi_t(\omega_t)$ and $k_t^*(\omega_t') = k_t^*(\omega_t)$.

It follows from (17) that $A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*) = 0$. Then, for any pair $\omega_t', \omega_t \in \Omega_t$, $D(k_t(\omega_t')) - D\left(k_t(\omega_t)\right) = 0$, which implies that for any pair $\omega_t', \omega_t \in \Omega_t$, $k_t(\omega_t') = k_t(\omega_t) = k$. Theorem 1 shows that $k \in (k_a, k_b)$, so

$$0 = A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*) = F_a(k) - F_b(k) < 0,$$

a contradiction.

Case 2. Suppose that there exists a pair $\omega_t', \omega_t \in \Omega_t$ such that $\pi_t(\omega_t') > \pi_t(\omega_t)$ and $k_t^*(\omega_t') > k_t^*(\omega_t)$.

It follows from (17) that $A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*) > 0$. Denote $k_t^{\min} = \min\{k_t^*(\omega_t^1), k_t^*(\omega_t^2), ..., k_t^*(\omega_t^{2^{t-1}})\}$ and $k_t^{\max} = \max\{k_t^*(\omega_t^1), k_t^*(\omega_t^2), ..., k_t^*(\omega_t^{2^{t-1}})\}$. In addition, denote by $\omega_t^{\max}$ an experience that corresponds to $k_t^{\max}$, and by $\omega_t^{\min}$ an experience that corresponds to $k_t^{\min}$. From (17),

$$D(k_t^{\max}) - D(k_t^{\min}) = (\pi_t(\omega_t^{\max}) - \pi_t(\omega_t^{\min}))[A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*)]. \qquad (18)$$

Since $k_t^{\max} > k_t^{\min}$ and $A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*) > 0$, we conclude $\pi_t(\omega_t^{\max}) - \pi_t(\omega_t^{\min}) > 0$. Let $\mathbf{k}_t^{\max} \equiv \underbrace{(k_t^{\max}, ..., k_t^{\max})}_{2^{t-1} \text{ times}}$ and $\mathbf{k}_t^{\min} \equiv \underbrace{(k_t^{\min}, ..., k_t^{\min})}_{2^{t-1} \text{ times}}$. Then, we derive

$$A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*) \leq A_t(h_t^*, \mathbf{k}_t^{\max}) - B_t(h_t^*, \mathbf{k}_t^{\min}) = F_a(k_t^{\max}) - F_b(k_t^{\min}).$$

From (18),

$$\begin{aligned} D(k_t^{\max}) - D(k_t^{\min}) &= (\pi_t(\omega_t^{\max}) - \pi_t(\omega_t^{\min}))[A_t(h_t^*, \mathbf{k}_t^*) - B_t(h_t^*, \mathbf{k}_t^*)] \\ &\leq (\pi_t(\omega_t^{\max}) - \pi_t(\omega_t^{\min}))[F_a(k_t^{\max}) - F_b(k_t^{\min})]. \end{aligned}$$

By Theorem 1, $\pi_t(\omega_t^{\max}) - \pi_t(\omega_t^{\min}) \in (0, 1)$. WLOG, let F_a satisfy (A7). Then,

$$D(k_t^{\max}) - D(k_t^{\min}) \leq (\pi_t(\omega_t^{\max}) - \pi_t(\omega_t^{\min}))[F_a(k_t^{\max}) - F_b(k_t^{\min})] < [F_a(k_t^{\max}) - F_a(k_t^{\min})],$$

which is a contradiction with (A7). ■

Proof of Theorem 3. $\mathcal{F}_t$ denotes the product σ-field on Ω_t. WLOG, let F_a be the true population distribution. Then, the conditional probability that α is observed given ω_t is $1 - A_t(h_t^*, \mathbf{k}_t^*)$, and the conditional probability that β is observed given ω_t is $A_t(h_t^*, \mathbf{k}_t^*)$. Denote

$$\mu_t \equiv \frac{1 - \pi_t}{\pi_t} \geq 0.$$

Then,

$$\begin{aligned} E\left[\mu_{t+1} | \mathcal{F}_t\right] &= (1 - A_t(h_{t-1}^*, \mathbf{k}_t^*)) \frac{1 - B_t(h_{t-1}^*, \mathbf{k}_t^*)}{1 - A_t(h_{t-1}^*, \mathbf{k}_t^*)} \mu_t + A_t(h_{t-1}^*, \mathbf{k}_t^*) \frac{B_t(h_{t-1}^*, \mathbf{k}_t^*)}{A_t(h_{t-1}^*, \mathbf{k}_t^*)} \mu_t \\ &= \mu_t. \end{aligned}$$

Hence, μ_t is a martingale. Moreover, μ_t is nonnegative. It follows from Corollary (2.11) at p. 236 in Durret (1996) that $\mu_t \to \mu$ a.s. with some μ. Since

$$\pi_t = \frac{1}{1 + \mu_t} \text{ is continuous,}$$

$\pi_t \to \pi$ a.s. From (17), as $t \to \infty$,

$$D(k_t(\omega_t')) - D(k_t(\omega_t)) \to 0 \text{ a.s.,}$$

which results in $k_t \to k$ a.s. This in turn implies

$$A_t(h_t^*, \mathbf{k}_t^*) \to F_a(k) \text{ a.s. and } B_t(h_t^*, \mathbf{k}_t^*) \to F_b(k) \text{ a.s.} \tag{19}$$

Now, suppose $\mu \neq 0$. Since $\mu_t \to \mu$ a.s., and $\log(1/z)$ is a continuous function of z on $\mathbb{R}_{++}$,

$$\log(\frac{1}{\mu_t}) \to \log(\frac{1}{\mu}) \text{ a.s.} \tag{20}$$

Taking the natural log of (14), we have

$$\log(\frac{1}{\mu_{t+1}}) - \log(\frac{1}{\mu_t}) = \mathbf{1}_{\{\alpha\}}(\omega(t)) \log\left(\frac{1 - A_t(h_t^*, \mathbf{k}_t^*)}{1 - B_t(h_t^*, \mathbf{k}_t^*)}\right) + \mathbf{1}_{\{\beta\}}(\omega(t)) \log\left(\frac{A_t(h_t^*, \mathbf{k}_t^*)}{B_t(h_t^*, \mathbf{k}_t^*)}\right). \tag{21}$$

By (20),

$$\log(\frac{1}{\mu_{t+1}}) - \log(\frac{1}{\mu_t}) \to 0 \text{ a.s.}$$

However, by (19),

$$\begin{aligned} \log\left[\frac{(1 - A_t(h_{t-1}^*, \mathbf{k}_t^*))}{(1 - B_t(h_{t-1}^*, \mathbf{k}_t^*))}\right] &\to \log\left[\frac{(1 - F_a(k))}{(1 - F_b(k))}\right] \text{ a.s.,} \\ \log\left[\frac{A_t(h_{t-1}^*, \mathbf{k}_t^*)}{B_t(h_{t-1}^*, \mathbf{k}_t^*)}\right] &\to \log\left[\frac{F_a(k)}{F_b(k)}\right] \text{ a.s.} \end{aligned}$$

It follows from the proof of Theorem 1 that $k \in [k_a, k_b]$ in the limit, and thus $F_a(k) < F_b(k)$, which implies that the right-hand side of (21) is almost surely non-zero. We have a contradiction. This implies that $\mu_t \to 0$ a.s. Since π_t is bounded by 1, $(1 - \pi_t) \to 0$ a.s., which leads to $\pi_t \to 1$ a.s. ■

References

Athey, S. (2001), Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information, *Econometrica* **69**, 861–889.

Carlsson, H., and van Damme, E. (1993), Global Games and Equilibrium Selection, *Econometrica,* 61, 989–1018.

Chen, Y., Katuscak, P., and Ozdenoren, E. (2007), Sealed bid auctions with ambiguity: theory and experiments, *Journal of Economic Theory,* 136, 513–535.

Dekel, E., D. Fudenberg, and D. Levine (2004), Learning to Play Bayesian Games, *Games and Economic Behavior*, 46, 282–303.

Durret, R. (1996), *Probability: Theory and Examples* (2nd ed.), Wadsworth Publishing Company, U.S.A.

Epstein, L.G., and Wang, T. (1996), "Beliefs about beliefs" without probabilities, *Econometrica,* 64, 1343–1373.

Fudenberg, D., and D. Levine (1993a), Self-Confirming Equilibrium, *Econometrica*, 61, 523–545.

Fudenberg, D., and D. Levine (1993b), Steady State Learning and Nash Equilibrium, *Econometrica*, 61, 547–574.

Fudenberg, D., and D. Kreps (1993b), Learning in Extensive Form Games, I. Self-Confirming Equilibria, *Games and Economic Behavior*, 8, 20–55.

Fudenberg, D., and D. Levine (1998): *Theory of Learning in Games.* MIT Press, Cambridge, MA.

Gilboa, I. and Schmeidler, D. (1989), Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics*, 18, 141–153.

Harsanyi, J.C. (1967), Games with Incomplete Information Played by Bayesian Players, *Management Sciences*, 14, 159–182, 320–334, 486–502.

Jackson, M.O., and E. Kalai (1997), Social Learning in Recurring Games, *Games and Economic Behavior*, 21, 102–134.

Jordan, J.S. (1991), Bayesian Learning in Normal Form Games, *Games and Economic Behavior*, 3, 60–81.

Jordan, J.S. (1995), Bayesian Learning in Repeated Games, *Games and Economic Behavior*, 9, 8–20.

Kalai, E., and E. Lehrer (1993a), Rational Learning Leads to Nash Equilibrium, *Econometrica*, 61, 1019–1045.

Kalai, E., and E. Lehrer (1993b), Subjective Equilibrium in Repeated Games, *Econometrica*, 61, 1231–1240.

Lo, K.C. (1998), Sealed bid auctions with uncertainty averse bidders, *Economic Theory*, 12, 1–20.

Marimon, R. (1997), Learning from Learning in Economics, in *Advances in Economics and Econometrics,* ed. by D. Kreps, and K. Wallis, vol. I, chap. 9. Cambridge University Press, Cambridge, U.K.

McAdams, D. (2003), Istone equilibrium in games of incomplete information, *Econometrica,* **69**, 861–889.

Mertens, J.F. and Zamir, S. (1985), Formulation of bayesian analysis for games with incomplete information, *International Journal of Game Theory,* **14**, 1–29.

Morris, S., and H.S. Shin (2003), Global Games—Theory and Applications, in *Advances in Economics and Econometrics, 8thWorld Congress of the Econometric Society*, ed. by M. Dewatripont, L. Hansen, and S. Turnovsky. Cambridge University Press, Cambridge, U.K. 56–114.

Nachbar, J.H. (1997), Prediction, Optimization, and Learning in Repeated Games, *Econometrica*, 65, 275–309.

Nachbar, J.H. (2004), Learning to Play Equilibria: The Bayesian, Repeated Games Approach, mimeo.

Nyarko, Y. (1994), Bayesian Learning Leads to Correlated Equilibria in Normal Form Games, *Economic Theory*, 4, 821–841.

Nyarko, Y. (1998), Bayesian Learning and Convergence to Nash Equilibria Without Common Priors, *Economic Theory*, 11, 643–655.

Rubinstein, A. (1989), The electronic mail game: strategic behavior under "almost common knowledge", *American Economic Review,* 79, 385–391.

Salo, A. and Weber, M. (1995), Ambiguity aversion in first-price sealed-bid auctions, *Journal of Risk and Uncertainty*, 11, 123–137.

Schmeidler, D. (1986), Subjective utility and expected utility without additivity, *Econometrica*, 57, 571–587.

van Huyck, J.B., R.C. Battalio, and R.O. Beil (1990), Tacit coordination games, strategic uncertainty, and coordination failure, *American Economic Review,* 80, 234–248.

Vega-Redondo, F. (2003), *Economics and the Theory of Games*, Cambridge University Press, Cambridge, U.K.

Vives, X. (1990), Nash equilibrium with strategic complementaries, *Journal of Mathematical Economics* **19**, 305–321.

Van Zandt, T., and Vives, X. (2007), Monotone equilibria in Bayesian games of strategic complementarities, *Journal of Economic Theory* **134**, 339–360.