THE KOREAN ECONOMIC REVIEW
Volume 21, Number 2, Winter 2005

THE TRACING PROCEDURE IN A POPULATION GAME MODEL

YONG-GWAN KIM*

The paper provides an evolutionary game theoretic reinterpretation of Harsanyi's (1975) tracing procedure and introduces a new solution concept for a population game model. In our population game theoretic interpretation players' common prior is the initial population state, and the tatonnement process is not a mental process but a gradual change of the population state. The population dynamic guarantees convergence to a Nash equilibrium, since its limit point is the same as the Nash equilibrium in the original tracing procedure. We also use the population tracing procedure repeatedly to define a refinement of Nash equilibria and call the limit set under the iterated population tracing procedure a 'population stable set' (PSS below). It is shown that a PSS always exists and that Swinkels' (1992) equilibrium evolutionarily stable set is a PSS.

JEL Classification: C72, C73
Keywords: population tracing procedure, evolutionary dynamics, population stable set

I. INTRODUCTION

This paper provides an evolutionary game theoretic reinterpretation of Harsanyi's (1975) tracing procedure and introduces a new solution concept for a population game model. The tracing procedure was designed to select a particular Nash equilibrium in a finite noncooperative game as the solution, and was extensively used in Harsanyi and Selten's (1988) general theory of

Received for publication: Oct. 13, 2005. *Revision accepted:* Dec. 16, 2005.

* Department of Economics, Sungkyunkwan University, Seoul 110-745, Korea. I am grateful to Vincent P. Crawford, Joel Sobel, two anonymous referees and the editor for helpful advice and valuable comments, and to Jeong Hun Oh for research assistance. Financial support from the Brain Korea 21 Program of the Ministry of Education of Korea is gratefully acknowledged.

equilibrium selection in games. It can be said that the tracing procedure is a mathematical representation for the mental tatonnement process of fully rational players' convergent expectations. But it is hard to provide an empirical justification of players' common priors in this intellectual process, and it is also difficult to see why all players must follow the same reasoning process until they reach a specific Nash equilibrium.

Following Myerson's (1991) suggestion we reinterpret the tracing procedure in a population game model (see also Binmore (1987, 1988)). Players' common prior is the initial population state, and the tatonnement process is not a mental process but a gradual change of the population state. In our evolutionary reinterpretation we consider two kinds of agents in a population. There are *arational* agents in large populations who do not have any intellectual capabilities and whose strategies can be viewed as a preprogrammed behaviors as in biological game theory. There are also *rational* agents who are intelligent enough to know the present population frequency and choose the best responses against it.

The evolutionary dynamics we suggest are as follows. Initially there is no rational player in the population and the initial population state can be given arbitrarily. Then a small proportion of players, who are called rational players instead of mutants as in biology, change their strategy choices to best respond against the current population state. As them goes by, the proportion of rational players increases and more players best respond to the population state. This process continues until every player becomes rational.

The population dynamic guarantees convergence to a Nash equilibrium, since its limit point is the same as the Nash equilibrium in the original tracing procedure. It can explain how a Nash equilibrium is attained in a population game model. But it does not solve the equilibrium selection problem completely, since every Nash equilibrium can be a limit point of the population tracing procedure. Moreover, as Van Damme (1991) has pointed out, an imperfect Nash equilibrium is also a limit point under the tracing procedure.

Sometimes we have to use the population tracing procedure more than once to find plausible solutions in a game. This implies that the proportion of rational agents increases until they reach a target state, but that agents behave aratinoally again after it is reached. They follow the same steps over and over again whenever they discover a better target state than the present one. Assuming that being rational requires a small amount of costs will justify the repeated application of the population tracing procedure. Once they reach a target state it is not necessary to calculate the present one and to figure out what to do until they discover a better one.

When we use the population tracing procedure repeatedly to define a refinement of Nash equilibria, we call the limit set under the iterated population tracing procedure a 'population stable set' (PSS below). It can be shown that a PSS always exists unlike some other static solution concepts in evolutionary game theory.

We compare the PSS with other evolutionary solution concepts. Our solution concept is related with the linear tracing procedure introduced by Harsanyi (1975), since our population tracing procedure comes from a reinterpretation of it in a population game model. But the prediction of PSS may differ for some games from Harsanyi or Harsanyi and Selten (1988), since we can use the population tracing procedure repeatedly. It is also shown that a PSS is different from the long-run equilibria under Kandori, Mailath, and Rob's (1993) or Young's (1993) evolutionary dynamic, or form a cyclically stable set (CSS below) of Gilboa and Matsui (1991) or Matsui (1992). On the other hand, a PSS is closely related to Swinkels' (1992) equilibrium evolutionarily stable set (EES set below). It is verified that an EES set for a game is always a PSS. This result provides another dynamic justification of the previous evolutionary analyses of communication games, since Kim and Sobel (2003) and Blume, Kim, and Sobel (1993) use the EES set in their analyses.

We also investigate properties of this new solution concept. We show that in some games PSS selects against equilibria which are not subgame perfect, but that it does not eliminate a Nash equilibrium in weakly dominated strategies in general. A PSS does not always satisfy the never-a-weak-best-response property suggested by Kohlberg and Mertens (1986) either. But in some coordination games with an outside option PSS does select the outcomes that forward induction arguments predict. Moreover, the PSS condition eliminates some inefficient equilibria in cheap talk games. PSS also provide justifications for some non-degenerate mixed strategy equilibria.

The rest of this paper is organized as follows. In section Ⅱ, we present our model and the population tracing procedure and provide some properties of it. We define the population stable set using the population tracing procedure in section Ⅲ. We show how it is different from other related solution concepts in section Ⅳ. In particular, comparisons between PSS and EES sets are provided.

Finally, in section V we analyze more examples of games to see the implications of the new solution concept.

Ⅱ. THE MODEL AND THE POPULATION TRACING PROCEDURE

In this section we reinterpret the linear tracing procedure in a population game model and introduce the population tracing procedure.

We consider a population game model with a finite number of populations. There are a large number of agents in each population. Players (one player from each population) are randomly and repeatedly matched to take some actions in a game. Every player can choose mixed strategies and each population may be polymorphic in the sense that different players play different strategies. There are arational agents in each population who do not have any intellectual capabilities and whose strategies can be viewed as preprogrammed behaviors as in biology at a given point in time. There are also rational agents in each population who are intelligent enough to change their strategy choices to best respond against the current population frequency. Even though most agents do not behave rationally in the initial state, the fact that rational agents can behave optimally and grow in their proportion over time allows the whole population to select a Nash equilibrium for a game in the end.

Formally, we consider a finite normal form game $G=(I, S, H)$. Let $I=\{1,\dots,i\dots,n\}$ be the set of n populations. There are a large number of agents in each population. The set S_i is the finite pure strategy space for each player in population i for all $i=1,\dots,n$. A pure-strategy profile $s=(s_i)_{i\in I}$, where $s_{i\in}S_i$ for all $i\in I$, is a collection of pure strategies containing one strategy of a player in each population. The pure-strategy space S of the game is the set of all pure strategy profiles. That is, $S=\times_{i\in I}S_i$. $H_i(s)$ is a payoff function for an s_i-strategist in population i for all i.

We allow players to choose mixed strategies. Let $p_i(s_i)$ denote the probability assigned to s_i. It can represent both an individual's mixed strategy and the mean population frequency in population i. In the second case it will be called an "i-population frequency." Mathematically, every p_i is a probability distribution over s_i for all i. Therefore it has the nature of a mixed strategy chosen by players in population i from the viewpoint of players in other populations, even when individuals choose only pure strategies. A collection $p=$

$(p_i)_{i \in I}$ contains a probability distribution p_i over s_i for every population $i \in I$ and will be called a population frequency or a population strategy profile. $p_{-i} = (p_i)_{i \in I/\{i\}}$, which contains one probability distribution for all populations except population i, denotes the average behaviors (or mixed strategy choices) of all other populations from the viewpoint of an agent in population i. It is called either the "i-incomplete population frequency" or "i-incomplete mixed strategy." The set Q_i is the set of all probability distributions over S_i and $Q = \times_{i \in I} Q_i$ is the collection of Q_i's.

We do not model time explicitly in our outcome selection process. Instead we use a parameter $t \in [0, 1]$ to denote the proportion of rational players in a population. At the beginning of the process, there exist only arational agents in all populations and the collection of initial population frequencies is denoted by p. But as time goes by the proportion of rational agents increases and they change their strategy choices to best respond against the current population frequency. Let q_i represent the mixed strategy choice of rational agents in population i, and $q_i(s_i)$ the probability weight given the pure strategy. A collection $q = (q_i)_{i \in I}$ will be called the "rational population frequency." Likewise, $p_{-i} = (p_i)_{i \in I/\{i\}}$ is called an i-incomplete rational population frequency, which implies the average behaviors (or mixed strategy choices) by rational agents in all populations except i.

Given a collection of probability distributions p in Q, the expected payoff function for a s_i-strategist in population I can be written

$$H_i(s_i, p_{-i}) = \sum_{s_{-i} \in S_{-i}} \prod_{j \in I/\{i\}} p_j(s_j) H_i(s_i, s_{-i}) \tag{1}$$

We can extend this payoff function to mixed strategies using the expected utility theorem. A mixed strategy q_i of a player in population i is called the best response to an i-incomplete population frequency p_{-i} and is denoted by $q_i \in BR(p_{-i})$, if

$$H_i(q_i, p_{-i}) \geq H_i(q_i', p_{-i}) \text{ for all } q_i' \in Q_i \tag{2}$$

We call q_i a strong best response if the inequality holds as a strict inequality. The strategy profile $q^* = (q_i^*)_{i \in I}$ is a Nash equilibrium in the game if it satisfies the mutual best response property. That is, for all $i \in I$, $q_i^* \in BR_i(q_{-i}^*)$.

Now we are ready to define a rational dynamic path and a population dynamic path. In our model we allow the proportion of rational agents in all populations to increase gradually over time.

The rational dynamic path is based on a one-parameter family of auxiliary games G^t, in which the proportion of rational agents in every population is t, where $0 \leq t \leq 1$. In any game G^t, rational agents in population i will have the same strategy set Q_i as they would have in the original game G. It is assumed that the population frequency of arational agents do not change over time.[1] Then the expected payoff function H_i^t for a rational agent in population i, who chooses a mixed strategy q_i, will be

$$H_i^t(q_i, q_{-i}) = (1-t)H_t(q_i, p_{-1}^o) + tH_i(q_i, q_{-i}) \tag{3}$$

Each rational agent selects his strategy to maximize his expected payoff function H_i^t. The strategy choice of rational agents in population i who behave this way is denoted by $q_i(t)$. Here we assume that every rational agent chooses the same mixed strategy $q_i(t)$.[2] $q(t)$ is the collection of $q_i(t)$ for all $i \in I$.

Definition 2.1: For any p^o, the **rational dynamic path** is defined to be a correspondence q: $[0,1] \rightarrow Q$, whose graph is a path connected set and which satisfies the following condition:
For all $t \in [0,1]$ and for all $i \in I$,

$$q_i(t) \in \arg\max_{q_i \in Q_i} H_i^t(q_i, q_{-i}(t)) \tag{4}$$

By definition, $q(t)$ has the property that each component $q_i(t)$ is a best response to a convex combination of p_{-i} and $q_{-i}(t)$. It is each to see that the strategy profile $q(t)$ is a Nash equilibrium in the auxiliary game G^t.

[1] This assumption implies that all agents choosing different strategies become rational in a proportionate way. If we use a different assumption, the rational dynamic path may reach a different Nash equilibrium. We do not analyze this more general case in the paper.

[2] We are imposing a monomorphic assumption on the behavior of rational agents. We may allow them to choose different strategies subject to the restriction that the average strategy of rational agents is equal to $q_i(t)$.

Observe that when $t=0$ a rational agent's payoff does not depend on rational agents' strategies in other populations, and that when $t=1$ his payoff function is the same as that in the original game G. That is, when $t=0$, G^0 is a game in which the payoff H_i^o of each agent in population i depends only on his own strategy q_i and behaviors of arational agents in other populations. When $t=1$, everybody behaves rationally and the game G^1 is the same as the original game G.

For all $t \in [0,1]$, the population frequency $p(t)$ is given by $(1-t)q^o + tq(t)$. This gives the formal definition of population dynamic path.

Definition 2.2: For any p^o, the **population dynamic path** is defined to be a correspondence p: $[0,1] \rightarrow Q$, which satisfies the following property:
For all $t \in [0,1]$,

$$p(t) = (1-t)p^o + tq(t), \tag{5}$$

where $q(t)$ is a rational dynamic path as defined in Definition 2.1.

$p(t)$ express the population frequency at t when the initial frequency is given as p^o. In contrast, $q(t)$ is the rational agents' strategy choice when their proportion in the population is t. As time goes by, the influence of the rational dynamic path increases gradually until discrepancies between them completely disappear in the end. At the beginning of the dynamical system, the proportion of rational agents is so trivial that the average behaviors of rational agents can be very different from the average behavior of the total population. As time goes by, the proportion of rational agents increases and the dynamic system progresses until both of them finally converge to a specific Nash equilibrium of the game. Thus, the average behavior of the total population will be the same as that of rational agents in the end.

Let X be the graph of the correspondence $t \rightarrow E^t$ for $0 \leq t \leq 1$, where E^t denotes the set all Nash equilibria in the game G^t. Each point x of X will have the mathematical form $x^t = (t, q(t))$. Suppose the graph X contains a path L connecting a point $x^o = (0, q(0))$ with a point $x^1 = (1, q(1))$. Then the path L represents a rational dynamic path. Note that L is the path which Harsanyi describes as a feasible path. The strategy part $q(1)$ of the end point x^1 will be the outcome selected by the rational dynamic path L.

Let Y be the graph of a correspondence $t \rightarrow Q$ for $0 \leq t \leq 1$, where each point y of Y will have the mathematical form $y^t = (t, p(t))$ and $p(t)$ is a point of the population dynamic path as defined in Definition 2.2. Suppose the graph Y contains a path M connecting a point $y^o = (0, p^o)$ with a point $y^1 = (1, p(1))$. Then the path M represents the population dynamic path, and the points y^o and y^1 will be called the starting point and the end point of the path M. The strategy part $p(1)$ of the end point y^1 will be the outcome selected by the population dynamic path M.

Observe that both dynamic paths select the same Nash equilibrium, since they coincide when $t = 1$. That is, $p(1) = q(1)$. We present the existence result of rational and population dynamic paths based on Harsanyi's (1975) proof of existence of a feasible path under the linear tracing procedure.

Proposition 2.1: For all (G, p^o) there exists a rational and population dynamic path.

Now we define the population tracing procedure. We can define it in two ways. In one case we can use the rational dynamic path and in another case we can use the population dynamic path. Both definitions result in the same Nash equilibrium. Below we use the population dynamic path in the definition of population tracing procedure.

Definition 2.3: The population tracing procedure $T(G, p^0)$ is a procedure selecting a strategy profile q^* for a game G by tracing a population dynamic path M from its starting point $y^o = (0, p^o)$ to its end point $y^1 = (1, q^*)$.

In some games there may be multiple dynamic paths for some initial states and the linear tracing procedure is not well-defined in Harsanyi's sense. To overcome this problem Harsanyi introduced a new tracing procedure called the logarithmic tracing procedure. However, we decide not to use the logarithmic tracing procedure in this paper, since it is difficult to provide a realistic interpretation of the logarithmic term in the payoff function of Harsanyi's new procedure. Instead we allow players to take any feasible path when there are multiple rational dynamic paths.

We state a useful lemma for the population tracing procedure below.

Lemma 2.1: Suppose that q^* is a collection of strong best responses at $t=0$ for a given p^0, and is also a collection of best responses, whether strong or weak, at $t=1$. Then q^* is an outcome selected by the population tracing procedure. That is, $q^* = T(G, p^o)$.

Proof: For each population i, q_i^* is a strong best response to p_{-i}^0 at $t=0$, and at least a week best response to q_{-i}^* at $t=1$. Therefore, by (4), q_i^* will be a strong best response to $(1-t)p_{-i}^0 + tq_{-i}^*$ at all t with $0 \leq t < 1$. Thus the graph X contains a rational dynamic path, which is a constant segment for all t with $0 \leq t \leq 1$. This establishes the lemma. Q.E.D.

It is easy to see that Lemma 2.1 implies the following two corollaries for the Nash equilibria in the game G.

Corollary 2.1: Suppose that p^0 itself is a collection of strong best responses at $t=0$. Then p^0 is the only outcome selected by the population tracing procedure. That is, $p^o = T(G, p^o)$.

Corollary 2.2: Suppose that p^0 itself is a collection of at least weak best responses at $t=0$. Then p^0 can be an outcome selected by the population tracing procedure. That is, $p^o \in T(G, p^o)$.

Harsanyi also provides an explanation about the classification of the line segments of linear tracing graphs. There are three types of line segments -- constant segments, jump segments and variable segments. This classification can be used conveniently in finding a rational dynamic path or a population dynamic path for a given game.[3]

[3] One thing that we should note is that we may have a backward-moving variable segment in a rational dynamic path for some games. That is, the value of t may decrease in a segment of the linear tracing graph for a game. Harsanyi and Selten provide an example for a backward-moving variable segment in their book. A behavioral interpretation of this variable segment is that when rational players have difficulty in finding a connected path for the increase of the t-value, they get confused and the proportion of rational agents decreases for a while and increases later. We do not pursue the full implication of this observation in this paper.

Ⅲ. POPULATION STABLE SET

We have reinterpreted the liner tracing procedure in a population game model. In this section, we define the population stability condition to introduce a new solution concept in evolutionary game theory. For this purpose we may use the population tracing procedure more than once. In other to define a new solution concept we need the notion of approachability between population strategy profiles and some lemmas on it.

Definition 3.1: For two population strategy profiles p and q in a game G, q is **directly approachable** from p if there exists a population dynamic path M, where its starting point is $y^o=(0,p)$ and its end point $y^1=(1,q)$.

From the definition of direct approachability and population tracing procedure, the following lemma is obvious. That is, if a population strategy profile is directly approachable from another strategy profile, it must be a Nash equilibrium in the underlying game.

Lemma 3.1: If a population strategy profile q is directly approachable from p in Q, then q is a Nash equilibrium.

Now, let us define a more general approachability relation by using the population tracing procedure more than once.

Definition 3.2: For two population strategy profiles p and q in a game G, q is **indirectly approachable** from p if there exists a finite sequence $\{p_k\}_{k=1}^{K}$ with $K>2$ such that p^{k+1} is directly approachable from p^k for all k with $1\leq k\leq K-1$, $p^1=p$ and $p^K=q$.

Definition 3.3: For two population strategy profiles p and q in a game G, q is **approachable** from p if q is either directly or indirectly approachable from p.

The following two lemmas will be important in the characterization of the new solution concept, which will be defined below using the approachability relation. Lemma 3.2 shows that the approachability relation is a transitive relation, and Lemma 3.3 shows that any two Nash equilibria which belong to the same component are approachable from each other.

Lemma 3.2: If r is approachable from q which is again approachable from p, then r is approachable from p.

Kohlberg and Mertens (1986) provide important characterizations of the Nash equilibrium correspondence for games. One characterization shown by them is that the set of Nash equilibria for a game consists of finitely many closed and connected sets. If we apply our approachability relation to a component (a maximal connected set) of Nash equilibria, we can see that any two Nash equilibria which belong to the same component are approachable from each other. In particular, Corollary 2.2 shows that a Nash equilibrium strategy profile is approachable from itself. If the game G is derived from a generic extensive form game, each connected set of Nash equilibria has a single outcome as shown by Kreps and Wilson (1982) and the population tracing procedure changes players' behaviors only at information sets off the equilibrium path. To move from one Nash equilibrium to another Nash equilibrium of the same component in more general games, we have to pass finitely many different critical Nash equilibria. In this case we can use the population tracing procedure several times. This finding is summarized in Lemma 3.3.

Lemma 3.3: Any two strategy profiles q and r in a maximal connected set of Nash equilibria are approachable from each other.

Now we can present the definition of population stability, which is the new solution concept introduced in this paper. Since we are using the population tracing procedure repeatedly we have a set-valued concept for the definition of our solution. That is, we allow drift of the population frequency at least in the same component of Nash equilibria. We can see the same drift phenomenon in other set-valued solution concepts of evolutionary game theory. Our solution concept is similar to Gilboa and Matsui's (1991) or Matsui's (1992) cyclically stable set, but a different notion of accessibility is used from theirs (see also Gilboa and Samet (1991) for a more general approach).

Definition 3.4: A closed set of strategy profiles $Q^* \subset Q$ in a game is

population stable if no $p \notin Q^*$ is approachable from any $q \in Q^*$ and every $r \in Q^*$ is approachable from all $q \in Q^*$.

As we have emphasized before, the population tracing procedure may be used repeatedly to reach a population stable set. In the repeated applications of population tracing procedure players should become arational and rational repeatedly. Hence, it may take a long time for the population to reach an element of a PSS.

To prove the existence theorem for PSS, let us introduce two sets of strategy profiles approachable from a population strategy profile p in Q. First, let $\bar{A}(p)$ be the set all strategy profiles directly approachable from p. That is, $\bar{A}(p) = \{q \in Q; q$ is directly approachable from $p\}$. Next, let $A(p)$ be the set of all strategy profiles which are approachable from p, whether directly or indirectly. That is, $A(p) = \{q \in Q; q$ is approachable from $p\}$. It is easy to see from Proposition 2.1 that for a finite game that $\bar{A}(p)$ is nonempty. Since $\bar{A}(p) \subset A(p)$, $A(p)$ is also nonempty. It can be also shown that both sets are closed for all $p \in Q$.

Lemma 3.4: The set $\bar{A}(p)$ is closed for all $p \in Q$.

Proof: First, suppose $\bar{A}(p)$ is a finite set. Then it is obviously a closed set. Next, suppose $\bar{A}(p)$ is an infinite set. Let $\{q^k\}_{k=1}$ be a sequence in $A(p)$ which converges to q. Then there exist collections of correspondences $\{q^k(t)\}_{t \in [0,1]}$ with $q^k(1) = q^k$ for all $k = 1, 2, \ldots$, whose graphs are path connected sets. This implies that there exists a rational dynamic path $\{q(t)\}_{t \in [0,1]}$ with $q(1) = q$. Here, the existence of $q(t)$ follows from upper hemicontinuity and convex-valuedness property of the best response correspondence and the upper hemicontinuity property of the Nash equilibrium correspondence for the game G^t. Therefore $\bar{A}(p)$ is a closed set. Q.E.D

Lemma 3.5: The set $A(p)$ is closed for all $p \in Q$.

Proof: From Lemma 3.1 the set $A(p)$ must be a set of Nash equilibria. From Lemma 3.3 all the Nash equilibria which belong to the same component must be in the set $A(p)$, if at least one element belongs to $A(p)$. Since the set of

Nash equilibria consists of finitely many closed components, $A(p)$ must also be a union of finitely many closed components of Nash equilibria. Therefore $A(p)$ is a closed set. Q.E.D

Now we can prove the existence theorem of PSS using the above lemmas. Unlike Gilboa and Matsui (1991) and Gilboa and Samet (1991) we do not need a heavy mathematical tool like Zorn's lemma.

Theorem 3.1: There exists a population stable set in a game.

Proof: Pick a component of Nash equilibria in a game. Check whether it is possible to move from an element of it to another Nash equilibrium in another component following the population tracing procedure. If not, the component is a PSS. On the other hand, if there is another component of Nash equilibria that can be reached through the population tracing procedure, check whether there is a third component which is approachable from it. Iterate this procedure until it is found that no other component is approachable from any of these components. This process stops in finitely many steps, since there are only a finite number of components in a game. Then, find the minimal collection of components which are approachable from each other. This is a PSS, since it satisfies two conditions of PSS. Q.E.D.

Since the PSS is defined based upon repeated applications of population tracing procedure and the population tracing procedure guarantees convergence to a Nash equilibrium, all element in a PSS are Nash equilibria, but some Nash equilibria do not belong to any PSS in a game. That is, a PSS is a refinement of Nash equilibria in a game. An example which shows that a Nash equilibrium does not belong to any PSS will be explained below in Figure 1 game. However, a strict Nash equilibrium has a strong stability property under our dynamical system as we have seen in Corollary 2.1 and it is easy to see that it is a PSS as a singleton.[4] This observation is summarized in Lemma 3.6.

Lemma 3.6: A strict Nash equilibrium in a game is a PSS as a singleton.

[4] Considering that an ESS in an asymmetric game is a strict Nash equilibrium, we can say that a PSS is a weakening of Maynard Smith's (1982) ESS.

Because of Lemma 3.6 it is important to note that the outcomes our PSS selects for a game will depend not only on the structure of game G but also on the initial population frequency p^o. We will describe this dependence in a simple class of two-person 2×2 games as shown in Figure 1, where $a>0$.

[Figure 1]

	L	R
T	a, a	0, 0
B	0.0	1, 1

All games of this class have three Nash equilibria. Two pure strategy equilibria are $E_1=(T,L)$ and $E_2=(B,R)$, and one mixed strategy equilibrium is $E_3=((1/(a+1),a/(a+1)),\ (1/(a+1),a/(a+1)))$.

[Figure 1']

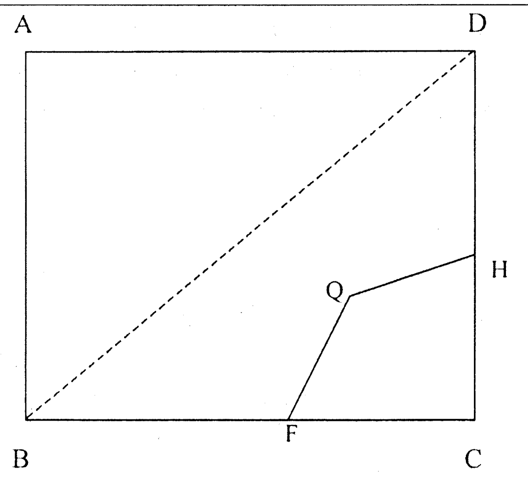


Figure 1' shows the strategy space of a typical game in this class. The three Nash equilibrium points of the game are $\mathrm{A}=E_1$, $\mathrm{C}=E_2$, and $\mathrm{Q}=E_3$. Here the graph is drawn for the case $a>1$. The line segment FQ is an extension of line QD and the line segment QH is an extension of line BQ.

Following the solution concept PSS, we can see that any point p lying above the broken line FQH will yield $\mathrm{A}=E_1$ as the outcome, whereas any point p

lying below that line will yield $C = E_2$ as the outcome. All point p lying on the line will yield either $A = E_1$ or $C = E_2$ as the possible outcome. Observe that the mixed strategy equilibrium is never an element of a PSS, and that two strict Nash equilibria have different basins of attractions.

Ⅳ. COMPARISONS WITH OTHER SOLUTION CONCEPTS

Now we explore the connection between our solution and other related solution concepts including Harsanyi's (1975) tracing procedure, Harsanyi and Selten's (1988) solution, Gilboa and Matsui's (1991) cyclically stable set, the long run equilibria under Kandori, Mailath, and Rob(s (1993) or Young's (1993) evolutionary dynamic, and Swinkel's (1992) EES set.

Our solution concept is closely related with the linear tracing procedure introduced by Harsanyi, since our population dynamic path comes from a reinterpretation of his theory in a population game model. The tracing procedure is developed to model the equilibrium selection process by extending Bayesian decision theory to finite games. An important assumption in this model is that every player assigns the same subjective prior probability distributions to a particular player's strategy choices. Then all players update their expectations about other players' strategy choices and modify their own strategy plans following the same reasoning process until a specific Nash equilibrium in the game is reached. As a positive theory it is difficult to justify the assumption that players have common subjective priors and that they follow the same reasoning processes. We reinterpret and develop the linear tracing procedure in a population game model to overcome those problems. In this context players' common subjective priors play the same role as the initial population frequency. The outcomes are the results from the real population dynamics over time instead of the timeless and a historical mental reasoning process in the original tracing procedure. The differences can be shown in the game in Figure 1. If we follow Harsanyi's linear tracing procedure, the solution will be (T, L), (B, R) or $((1/(a+1), a/(a+1)),\ (1/(a+1), a/(a+1)))$ depending on the initial prior. Our population stability selects only (T, L) or (B, R) as possible solutions.

On the other hand, Harsanyi and Selten use a particular prior called the bicentric prior or the uniform prior to apply the tracing procedure. Depending on the size of payoff a in Figure 1 game, their solution function will select one among these three Nash equilibria. In particular, when $a=1$, the symmetric mixed strategy equilibrium is chosen as the solution.

Gilboa and Matsui's cyclically stable set also selects solutions for a game by defining accessibility under a specific best response dynamic. The main difference is that they specify a different dynamic path from us. We use a dynamic path under the assumption that the proportion of rational agents increase over time, but their dynamic path is defined under the assumption that the proportion of rational agents is always small. This difference results in different limit sets in some games. In particular, a PSS always consists of Nash equilibria, but a CSS may not. Their example shown in Figure 2 tells us the difference.

[Figure 2]

	L	C	R
T	2,2	1.2,1.2	-1.3
M	1.2,1.2	1,1	.2,.2
B	3,-1	.2,.2	0,0

The unique CSS is a triangle which connects three points PQR', where $P=(.4,.5,.1)$, $Q=(.16,.2,.64)$, and $R'=(.04,.8,.16)$. No element of this CSS is a Nash equilibrium. On the other hand, the unique PSS of the game is the unique Nash equilibrium $((1/4,1/2,1/4),\ (1/4,1/2,1/4))$.

Kandori, Mailath, and Rob and Young analyze finite population models with stochastic mutation processes. By imposing the inertia and myopia hypotheses on agents' reaction to their environment and the random mutation hypothesis, they show that there is a unique stationary distribution of the population frequency. They characterize the support of the limiting distribution for a class of coordination games including 2×2 games, as the rate of mutation goes to zero. In the game of Figure 1, there are two strict Nash equilibria. If $a>1$, (T,L) is the limit set, and if $a<1$, (B,R) is the limit set. More generally, they select the equilibrium that satisfies Harsanyi and Selten's risk dominance criterion in 2×2 games. According to PSS, both Nash equilibria are selected as solution in this game since they are strict.

Swinkels proposed a new solution concept of equilibrium evolutionarily stable set as a set-valued generalization of Maynard Smith and Price's (1973) evolutionarily stable strategy by imposing restrictions on the set of possible invading strategies. The formal definition of EES set is as follows.

Definition 4.1: A closed set of Nash equilibria Θ in a game is **equilibrium evolutionarily stable** if it is minimal with respect to the following property: There exists $\varepsilon' > 0$ such that for all $q \in \Theta$ and for all $\varepsilon \in (0, \varepsilon')$, if $r \in BR((1-\varepsilon)q + \varepsilon r)$, then $(1-\varepsilon)q + \varepsilon r \in \Theta$ for all $r \in Q$.

It turns out that there is a close relationship between an ESS set and a PSS. Exploiting characterizations on EES set in Swinkels and Matsui we can prove that an EES set in a game is always a PSS.

Theorem 4.1: If Θ is an EES set in a game, then Θ is a PSS.

The converse of Theorem 4.1 is not true. A PSS always exists in a game, but an ESS set does not. We borrow two examples from Swinkels for the comparison between PSS and EES set.

[Figure 3]

	L	C	R
T	2,2	0,2	5,0
M	2,0	3,3	0,3
B	0,5	3,0	4,4

The game of Figure 3 has three Nash equilibria: (T, L), (M, C), and $((3/7, 5/14, 3/14), (3/7, 5/14, 3/14))$. As Swinkels showed, there is no EES set in this game. On the other hand, there is a unique PSS, which is the singleton (M, C). When everybody in the population chooses the strategy profile (M, C), a new strategy profile (B, R) is also a best response. But when the proportion of (B, R) is high, the best response against the mixture of (M, C) and (B, R) is (T, L). Hence it is impossible to move from (M, C) to (B, R) through the population tracing procedure.

[Figure 4]

	L	C	R
A	0,8	15,0	0,15
B	0,7	0,15	15,0
C	5,5	7,1	7,0
D	6,6	6,6	6,6

In the game of Figure 4, there are two maximal connected sets of Nash equilibria: $E_1 = \{((1/2, 1/2, 0, 0), (q, (1-q)/2, (1-q)/2)); 0 \leq q \leq 1/11\}$ and $E_2 = \{(D, (q_1, q_2, 1-q_1-q_2)); 1/2 \leq q_1 \leq 1, 0 \leq q_2 \leq 2/5, 3/5 \leq q_1 + q_2 \leq 1\}$. Both components of Nash equilibria are EES sets, hence PSS(s, in this game. This example shows that neither EES set nor PSS captures the full implication of the forward induction argument suggested by Van Damme (1989). Moreover, unlike an EES set it does not satisfy the never-a-weak-best-response (NWBR below) property suggested by Kohlberg and Mertens in general. An example which shows this property is the game in Figure 3. The unique PSS (M, C) is not a PSS any more in a smaller game, if the never-a-weak-best-responses T and L are eliminated from the original game.

However, as will be shown in the next section, it has an interesting cutting power in communication games, which Kolberg and Mertens' strategic stability does not have.

Ⅴ. APPLICATIONS OF POPULATION STABILITY

In this section we apply the PSS condition to various classes of games.

[Figure 5]

	L	R
T	1,1	1,0
B	1,1	0,0

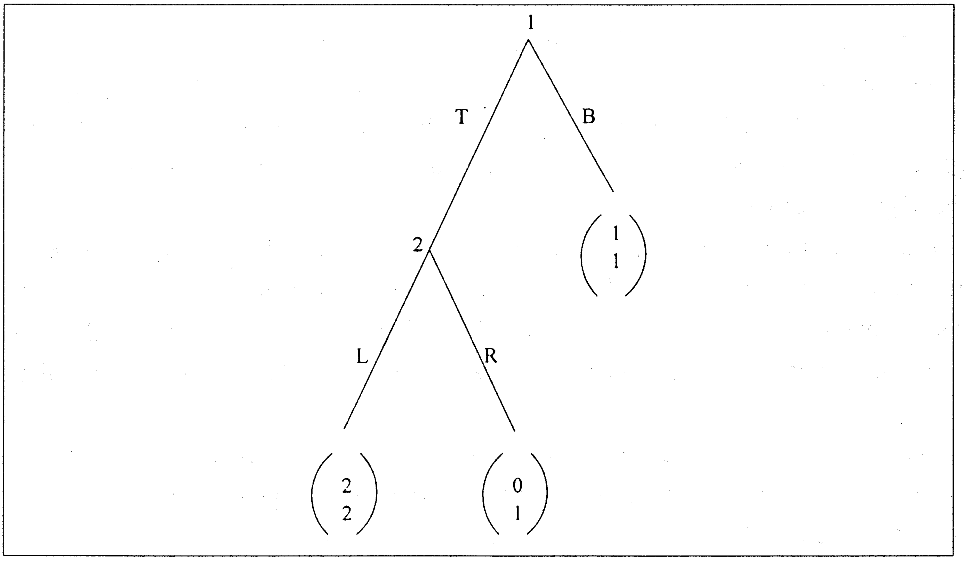
Figure 5 presents a game with dominated strategies for each player. This game has a unique connected set of Nash equilibria, $\{((p, 1-p), L);\ 0 \leq p \leq 1\}$, and PSS supports all these equilibria in which the weakly dominated strategy B is played. It is easy to verify that any two strategy profiles in this set are approachable from each other, and no strategy profile outside the set is approachable from any one inside the set. Note that while R is strongly dominated by L for players in population 2, B is only weakly dominated by T for players in population 1. Since there is no payoff difference between two pure strategies for a player in population 1 when everybody in population 2 plays L, players in population 1 may choose B with positive probability.

In some games, PSS selects against equilibria which are not subgame perfect. Consider the following game in Figure 6 and 6'. There are two connected sets of Nash equilibria: $E_1 = (T, L)$, and $E_2 = \{(B, (q, 1-q));\ 0 \leq q \leq 1/2\}$. It is easy to see that E_1 is approachable from E_2, but not vice versa. The only PSS is E_1, which is also the unique subgame perfect equilibrium.

[Figure 6]

	L	R
T	2,2	0,1
B	1,1	1,1

[Figure 6']



The results obtained in the above two examples are similar to Samuelson and Zhang (1992), Samuelson (1994) and Nöldeke and Samuelson (1993). Using explicit dynamic models in the spirit of replicator dynamics (Taylor and Jonker (1978) and Zeeman (1981)) or Kandori, Mailath, and Rob (1993), these works show that evolutionary game models do not provide a strong support for the backward induction argument unless individuals' mistakes are explicitly introduced into the model (see Selten (1983) for a limit ESS which always satisfies the subgame perfectness condition because tremble in an individual's strategy choice is introduced).

Evolutionary game theory which uses a population game model can provide a justification of mixed strategy equilibrium. Specifically, in a polymorphic population in which only pure strategies are played, the collection of proportions of each pure-strategist in the population provides an alternative to the classical interpretation of mixed strategies. According to this interpretation of PSS, some non-degenerate mixed strategy equilibria are considered as stable while others are not.

[Figure 7]

	L	R
T	1,-1	-1,1
B	-1,1	1,-1

For example, in the Matching Pennies game shown in Figure 7, there is a unique Nash equilibrium which is a non-degenerate mixed strategy equilibrium $((1/2, 1/2), (1/2, 1/2))$. It is easy to see that it is the unique PSS in this game. This is in contrast to the instability of the mixed strategy equilibrium in the Figure 1 game.

Unlike an EES set, the PSS does not satisfy the NWBR property in general. However, it can be shown that it satisfies the forward induction argument in some coordination games with an outside option like CSS. Besides it has an interesting property in communication games just like the other two solution concepts.

Let us consider a complete-information coordination game with pre-play communication first. Suppose two players engage in pre-play communication in the first stage and play the Figure 1 game in the second stage, where $a > 1$.

Note that the second-stage game is a common interest game in the sense that there is a unique efficient outcome. Even though our intuition is to select the (a, a) outcome as the unique sensible equilibrium outcome in the communication game, many solution concepts do not provide enough cutting power. In our model the population can drift from an inefficient tracing procedure. The drift and invasion mechanism used in Kim and Sobel (2003) works in the exactly same way (see also Matsui (1991)). That is, using the population tracing procedure we can move the population state from an arbitrary inefficient Nash equilibrium to a Nash equilibrium in which a message is not used and a strategy using that message is not punished. Then, using the population tracing procedure again we can move the population state to an efficient Nash equilibrium in which everybody uses the unused message and attains an efficient Nash equilibrium of the underlying game after the new message. We can summarize Kim and Sobel's efficiency and existence results using PSS instead of EES set as follows. We say that a game is of common interest, if there is a unique weakly Pareto efficient payoff in it. We also say that a game is of equilibrium common interest, if all the Nash equilibria in it are strictly Pareto rankable.

Proposition 5.1: In a complete information game G with pre-play communication, if the underlying game is of equilibrium common interest and if there are at least two messages available to each player, then there is a unique PSS and it contains the set of all Nash equilibria with the most efficient equilibrium payoff in the underlying game.

Proposition 5.2: In a complete information game G with pre-play communication, if the underlying game is of common interest and if there are at least two messages available to each player, then the set of Nash equilibria which support the efficient payoff is a PSS.

Blume, Kim and Sobel's (1993) result for a cheap-talk game with incomplete information can be also rephrased with the new solution concept PSS.

Proposition 5.3: In a cheap talk game G with incomplete information, if G is of common interest and if the message space of the sender is sufficiently large, then there is a unique PSS in the game G and it is the set of Nash equilibria which support the efficient payoff.

REFERENCES

Binmore, K. (1987), "Modeling Rational Players, Part I," *Economics and Philosophy*, 3, 179-214.

__________ (1987), "Modeling Rational Players, Part II," *Economics and Philosophy*, 4, 9-55.

Blume, A., Y.-G. Kim and J. Sobel (1993), "Evolutionary Stability in Games of Communication," *Games and Economic Behavior*, 5, 547-575.

Gilboa, I. and A. Matsui (1991), "Social Stability and Equilibrium," *Econometrica*, 59, 859-867.

__________ and D. Samet (1991), "Absorbent Stable Sets," mimeo, Northwestern University.

Harsanyi, J.C. (1975), "The Tracing Procedure," *International Journal of Game Theory*, 2, 61-94.

__________ and R. Selten (1988), *A General Theory of Equilibrium Selection in Games*. Cambridge, Massachusetts: the MIT Press.

Hofbauer, J. and K. Sigmund (1988), *The Theory of Evolution and Dynamical Systems*. Cambridge: Cambridge University Press.

Kandori, M., G.J. Mailath and R. Rob (1993), "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61, 29-56.

Kim, Y.-G. and J. Sobel (1992), "An Evolutionary Approach to Pre-Play Communication," *Econometrica*, 63, 1191-1193.

__________ (2003), "An Evolutionary Analysis of Games with Pre-Play Communication," *Korean Economic Review*, 19, 6-36.

Kohlberg, E. and J.-F. Mertens (1986), "On the Strategic Stability of Equilibria," *Econometrica*, 54, 1003-1037.

Kreps, D. and R. Wilson (1982), "Sequential Equilibria," *Econometrica*, 50, 863-894.

Matsui, A. (1991), "Cheap-Talk and Cooperation in a Society," *Journal of Economic Theory*, 54, 245-258.

__________ (1992), "Best Response Dynamics and Socially Stable Strategy," *Journal of Economic Theory*, 57, 343-362.

Mynard Smith, J. (1982), *Evolution and the Theory of Games*. New York: Cambridge University Press.

__________ and G.R. Price (1973), "The Logic of Animal Conflict," *Nature*, 246, 15-18.

Myerson, R.B. (1991), *Game Theory.* Cambridge, MA: Harvard University Press.

Nöldeke, G. and L. Samuelson (1993), "An Evolutionary Analysis of Backward and Forward Induction," *Games and Economic Behavior*, 5, 425-454.

Samuelson, L. (1994), "Stochastic Stability in Games with Alternative Best Replies," *Journal of Economic Theory*, 57, 343-362.

__________ and J. Zhang (1992), "Evolutionary Stability in Asymmetric Games," *Journal of Economic Theory*, 57, 363-391.

Selten, R. (1983), "Evolutionary Stability in Extensive Two-Person Games," *Mathematical Social Sciences*, 5, 269-363.

________ (1991), "Evolution, Learning, and Economic Behavior," *Games and Economic Behavior*, 3, 3-24.

Swinkels, J. (1992), "Evolutionary Stability with Equilibrium Entrants," *Journal of Economic Theory*, 57, 306-332.

Taylor, P. and L. Jonker (1978), "Evolutionarily Stable Strategies and Game Dynamics," *Mathematical Biosciences*, 40, 145-156.

Van Damme, E. (1989), "Stable Equilibria and Forward Induction," *Journal of Economic Theory*, 48, 476-496.

___________ (1991), *Stability and Perfection of Nash Equilbria.* Second Edition, Berlin: Springer-Verlag.

Young, P. (1991), "The Evolution of Conventions," *Econometrica*, 61, 57-84.

Zeeman, E. (1981), "Dynamics of the Evolution of Animal Conflicts," *Journal of Theoretical Biology*, 89, 249-270.