The Korean Economic Review
Volume 28, Number 2, Winter 2012, 137-160.

Ability-based Cooperation in a Prisoner's Dilemma Game*

Hanjoon Michael Jung**

This paper studies the possibility of whole population cooperation based on different abilities of players. Consider the following infinitely repeated game, similar to Ghosh and Ray (1996). At each stage, uncountable numbers of players, who are randomly matched without information about their partners' past actions, play a prisoner's dilemma game. The players have the option to continue their relationship, and they all have the same discount factor. Also, they have two possible types: high ability player (H) or low ability player (L). H can produce better outcomes for his partner as well as for himself than L can. We look for an equilibrium that is robust against both pair-wise deviation and individual deviation, and call such equilibrium a social equilibrium. In this setting, long-term cooperative behavior among the whole population can take place in a social equilibrium because of the players' preference for their partners' ability. In addition, a folk theorem of this model is proposed.

JEL Classification: C72, C78
Keywords: Folk theorem, Random-matching, Social equilibrium, Type-based payoffs

I. Introduction

The motivation of the present paper comes from the studies on the area of Folk Theorem. Classical literature in folk theorem developed by Fudenberg and Maskin (1986), Kandori (1992), and Ellison (1994) showed that a long-term cooperative relationship in a prisoner's dilemma is possible without any legal enforcement, assuming that players'past actions affect their future payoffs. Based on a different assumption that players'past actions might not necessarily affect their future payoffs because they can change their partners in a large population, Ghosh and Ray (1996),

Received: June 25, 2012. Revised: Sept. 28, 2012. Accepted: Oct. 19, 2012.

* The author is most grateful to Kalyan Chatterjee and James Jordan for their helpful comments and encouragement and gratefully acknowledges research support from the National Science Council of the Republic of China (NSC100-2410-H-001-029). The author also wishes to thank Gaurab Aryal and Wei-Torng Juang for their helpful comments.

** The Institute of Economics, Academia Sinica, 128, Sec. 2, Academia Road, Nangang, Taipei 115, Taiwan. E-mail: hanjoon@econ.sinica.edu.tw

hereinafter referred to as GR, maintained that a long-term cooperative relationship is still possible based on the setting of heterogeneous players.

However, GR showed a long-term cooperative relationship among a partial population. In GR's model, there are two types of players; myopic players who have a zero discount factor and non-myopic players who have a positive discount factor. Here, the myopic players will not play any cooperative action because they are not concerned about their future, and thus such a cooperative action is strictly dominated by a non-cooperative action. Then, since players can cooperate only with non-myopic players, the matches with non-myopic players are endowed with a *scarcity value*. This scarcity value is used to sustain cooperation among the non-myopic players. As a result, GR's model can show a long-term cooperative relationship only when there exist a significantly large proportion of the myopic players, because the effectiveness of the scarcity value depends on this proportion of the myopic players. Therefore, GR's model can be viewed as a partial population cooperation model. The present study is motivated to seek a possibility of whole population cooperation in a prisoner's dilemma game assuming that players' past actions might not affect their future payoffs.

Consider a situation of collaboration between two individuals (or two companies). In this situation, each individual exerts his effort to improve the common outcome of their collaboration. His effort, however, brings only himself disutility. Thus, each of them wants the other to exert more effort, while the person himself would be better off exerting as little effort as possible. As a result, this situation can be viewed as the prisoner's dilemma. In addition, suppose that they can break their current relationship and can meet new partners to start a new collaboration. Here, we assume that their new partners do not know how much effort they exerted in the previous matches.

In this collaboration situation, all individuals can play maximum effort levels. This whole population cooperation is based on two assumptions. First, every individual is assumed to be either a high-ability player (H-player) or a low-ability player (L-player). An H-player is defined as a skilled individual so that he can produce better outcomes for his partner as well as for himself than an L-player. This assumption simply reflects the facts that different individuals have different abilities to produce outcomes and that individuals can benefit from the skill of their partners. Second, individuals in a common pair have the option to continue their relationships if they both wish. This assumption is also plausible in that many real life examples can justify it.

In this study, we show that the whole population cooperation happens in a social equilibrium, whose formal definition is presented in Section 4. The logic behind the result is as follows. An H-player wants to match and to play only with another H-player, because a high-ability partner produces better outcomes than a low-ability partner. So, when an H-player meets an L-player, the H-player would break the relationship with the L-player in order to increase the possibility of meeting another H-player. Thus, an H-player would not play any cooperative action with an L-player. Since an L-player is aware of H-players' intention, he realizes that he can only cooperate with another L-player. Consequently, H-players value high-ability partners above low-ability partners because they can get better outcomes when they play with the high-ability partners, and L-players value low-ability partners above high-ability partners because they can have cooperative relationships only with the low-ability partners. As a result, in equilibrium, two kinds of matches, the H-H match and the L-L match, are endowed with a scarcity value. Players can use this scarcity value to sustain their cooperative relationships. Therefore, the result shows that, in equilibrium, long-term cooperative relationships among the whole population are possible based on different abilities of the players. In addition, when the players are sufficiently patient, they play the maximal cooperation level in the proper matches, the H-H matches and the L-L matches. This result is proposed as a folk theorem of this model.

Originally, all individuals want to match and play with H-players because they can get worse outcomes by working with the L-players. In equilibrium, however, only H-players can sustain long-term cooperative relationships with other H-players. This is because individuals in a common pair can continue their relationships only if they both wish and the H-players do not wish to continue their relationships with the L-players. Then, the L-players can play cooperative actions only with other L-players. Therefore, in this situation, the low skill of the L-players gives rise to the scarcity value of the H-H matches, and the discrimination by the H-players gives rise to the scarcity value of the L-L matches. Then, to sustain cooperative relationships in these two kinds of matches, the individuals can use the threat of breaking these relationships. Therefore, all individuals can have long-term cooperative relationships in equilibrium. It is interesting that, if we replace all the L-players with H-players so that there could be only H-players in the new model, then none of them can maintain long-term cooperative relationships in equilibrium.

In addition, note that this mutual generation of the scarcity values distinguishes our model from that of GR. In GR's model, only the existence of the myopic players gives rise to the scarcity value of the non-myopic players' matches. However, by the very definition of the myopic players, the non-myopic players cannot give rise to the scarcity value of any matches including myopic players. Accordingly, in GR's model, only the non-myopic players can play long-term cooperative actions in equilibrium.

The rest of the paper is organized as follows. Section 2 discusses some related literature. Section 3 is devoted to a detailed description of the model. Section 4 introduces the concept of a social equilibrium. Section 5 presents the results of this study, including the folk theorem of this model. Section 6 concludes.

II. Related literature

Datta (1996) and Kranton (1996) studied the possibility of cooperation in settings similar to the present model, repeated prisoner's dilemma games with random matching. They showed that cooperative behavior is possible by means of raising cooperation levels gradually, i.e. building trust. However, these building trust equilibria are not immune against pair-wise deviation as indicated by GR[1] (see also Eeckhout,[9] 2006; Furusawa and Kawakami, 2008; and Fujiwara-Greve and Okuno-Fujiwara, 2009). On the other hand, in the present model, right after the players find their proper matches, they play the highest cooperative actions out of all actions that are robust against individual deviation. Therefore, the equilibria in the current model are immune against pair-wise deviation.

Recently, Fujiwara-Greve (2002) studied a similar issue; the possibility of cooperation in a prisoner's dilemma game with random matching. However, in contrast to the *complete random matching process* in which the probability of meeting a new partner is one, she considered an *incomplete random matching process* in which the probability of meeting a new partner is less than one. She showed that if the probability of meeting a new partner is sufficiently low, then players can play the highest cooperative action from the beginning of their relationships, and thus a folk theorem holds in her model. She explained that under the incomplete random matching process, each match is endowed with a scarcity value, because if a player loses his current partner, then he might not meet a new partner at the next period. As a result, players in her model can use this scarcity value to sustain a long-term cooperative relationship even when personalized punishments are not feasible. Therefore, in her model, the scarcity value is exogenously determined by the assumption about the incomplete random matching process. In the present model, on the other hand, a scarcity value is not given by any assumption because all the players, who have incentives to cooperate, can meet new partners at any time. The scarcity value, however, is endogenously generated by the players' preferences on their partners' ability, and this scarcity value is used by the players when they sustain a cooperative relationship with their partners.[10]

---

[1] Kranton (1996) also extended her model by introducing myopic players into the model, and found a result similar to GR. That is, the result in her extended model is robust against pair-wise deviation, like the result in GR.

[9] Eeckhout (2006) studied type-based strategies and showed a cooperative behavior can happen between the players of the same types, which is similar to the outcome in the current study. In his model, however, players' types are irrelevant to their payoffs, thus the cooperative relationship between the same types causes the same payoffs as the cooperative relationship between the different types does. As a result, no cooperative action in his model is possible in equilibrium if the players consider the option of the pair-wise deviation.

[10] To see different approaches on this issue, the possibility of cooperation in a prisoner's dilemma game with random matching, please refer to Boone, Brabander, Carree, Jong, Olffen, and

In the present model, one of the critical assumptions is that the players' types depend on their payoff systems. Watson (1999, 2002) adopted a similar heterogeneity assumption about the players' types and showed that cooperation levels in a prisoner's dilemma game increase gradually. In his models, however, the payoffs to each player do not depend on their partners' types, and heterogeneity in payoff systems induces some of the players to have smaller incentives to cooperate, like the heterogeneity assumption of GR whose model included the myopic players who have no incentive to cooperate. Therefore, the heterogeneity in Watson (1999, 2002) played a role similar to the heterogeneity assumption in GR (see also Rauch and Watson, 2003). In the current model, on the other hand, the payoffs to each player depend on their partners' types, and this heterogeneity in payoff systems gives all the players an incentive to cooperate with specific types. As a result, a long-term cooperative relationship among the whole population is possible based on this heterogeneity assumption.

Finally, long-term economic behavior has been actively studied in an evolutionary frame-work. Based on the work by Foster and Young (1990), Kandori et al. (1993) studied the effect of ongoing mutations. They proved that the players who adopt the myopic best response are inclined to coordinate on the risk dominant equilibrium, which is indeed a stochastically stable equilibrium of Foster and Young (1990). Robson and Vega-Redondo (1996) considered a model similar to that by Kandori et al. (1993). Then, by employing a naive imitation rule, they showed that the Pareto-efficient equilibrium is selected. Recently, Juang and Sabourian (2012) considered a model in which players can revise their rules in playing the game, that is, they adopted the concept of the evolution of rules as well as that of the level of actions. They concluded that a folk theorem on equilibrium can be restored, which implies that the selection power based on mutations can be insignificant.[11] Basically, all these works presumed that the players have limited ability to control their actions or the rules. Hence, they can be viewed as studies of bounded rational behavior. Our model, on the other hand, assumes that the players can fully control their actions and rules. Therefore, we study only, fully, rational behavior.

III. The Model

The following setting of the model comes from GR. A continuum of players are randomly matched in pairs, and bilaterally play an infinitely repeated stage game with an option to break up their relationships. Each stage of the game consists of two substages. At the first substage, players in a common pair play a prisoner's

Witteloostuijn (2002), Bose (1996), Brosig (2002), Outkin (2003), and Yang, Yue, and Yu (2007).

[11] See also Ellison (1993) and Foster and Young (1991).

dilemma game with an action set $[0,\tilde{a}]\subset\mathbb{R}$. At the second substage, after watching the actions chosen before, the players decide whether to break up their relationships. Only when both players in a common pair decide to maintain their relationship, can they play the stage game between themselves at the next stage. If one of the players in a common pair breaks up the relationship, then both in the pair go into *the pool of unmatched players* and are randomly matched with other players in the pool. At the next stage, all players bilaterally repeat this stage game.

The present model introduces new features into the setting of GR. All players have the same discount factor δ, but they have their own types. Each player is either an H-player or an L-player. An H-player has higher abilities to produce an outcome than an L-player does. Based on this ability difference, the present model reflects the situation in which a partner of an H-player can benefit from the high ability of the H-player by sharing the produced outcome. Therefore, it is assumed that a player's payoff depends on his partner's type as well as on his own type, and also depends on his partner's and his actions so that, other things being equal, a player, regardless of his own type, gets a better payoff when he cooperates with an H-player than when he cooperates with an L-player.

The payoff functions of the players are as follows. For any $I,J\in\{H,L\}$, the function $\Pi_{IJ}:[0,\tilde{a}]^2\rightarrow\mathbb{R}$ denotes a payoff function of *I*-type when he works with *J*-type. For example, let $a,a'\in[0,\tilde{a}]$, then $\Pi_{HL}(a,a')$ denotes the payoff to an H-player when he works with an L-player under his action a and his partner's action a'. Here, the players' actions a and a' can be referred to as *cooperation levels*. Then, in order to reflect the prisoner's dilemma setting, it is assumed that for each $a,a'\in[0,\tilde{a}]$, if $a>0$, then $\Pi_{IJ}(0,a')>\Pi_{IJ}(a,a')$. In addition, the payoff under zero actions, $\Pi_{IJ}(0,0)$, is normalized to zero.

In this study, three assumptions about the payoff functions from GR's model are adopted and adapted. First, the payoff function Π_{IJ} is assumed to be continuous, and the function $\Pi_{IJ}(a,a)$ is assumed to be strictly increasing in a. This assumption is used for the sake of simplicity. Second, there exists $a\in(0,\tilde{a}]$ such that $\Pi_{IJ}(a,a)>(1-\delta)\Pi_{IJ}(0,a)$. Third, given any $a_L\in[0,\tilde{a}]$, there exists $a\in(0,\tilde{a}]$ such that $\pi\Pi_{HH}(a,a)+(1-\pi)\Pi_{HL}(a,a_L)>(1-\pi)\Pi_{HL}(0,a_L)$ where π is the proportion of H-players in the pool of unmatched players. If the second or the third assumption does not hold, then players might not have any incentive to play a positive action. Therefore, the latter two assumptions are used to exclude a trivial case in which players have no incentive to cooperate with their partners and prefer to play zero actions.

Regarding information, a player has limited information about types and actions. A player is informed only of his own type. However, if his partner plays a positive action, he can figure out his partner's type by comparing the outcomes drawn from his action and his partner's action. This is because, other things being equal, the cooperative action performed by a high ability partner brings out a better outcome than the action performed by a low ability partner. Note that, because of the normalization of the payoffs, a player cannot figure out his partner's type if his partner plays a zero action. In addition, a player knows only his own actions and his partners' actions from the beginning, but he does not know the actions taken by others. A player's *personal history* is defined as the record of his type, the types of his partners who have played positive actions, and all the actions taken by his partners and by him from the beginning. Therefore, a pure strategy of a player is a possible mapping from his personal histories either to the set of the actions $[0,\tilde{a}]$ for the first substages or to the set of the breakup decisions for the second substages.

IV. Social Equilibrium

In this study, our interest is restricted to *social norms* and *steady states* like in the study of GR. A social norm is a profile of pure strategies such that players of the same type use the same pure strategy. A state is steady if the proportion of H-players in the pool of unmatched players, π, is constant over time.[12] Moreover, our study focuses on the cooperation possibility based on players' preferences for the high ability of an H-player. So, we rule out the cases in which a player prefers betraying an H-player partner rather than cooperating with the H-player partner because of a huge payoff when he betrays the H-player partner. In addition, our equilibrium is required to satisfy two criteria: "Individual incentive constraint" and "Bilateral rationality," which were proposed by GR. These two criteria require an equilibrium to be proof against *individual deviation* and *pair-wise deviation*, respectively.

In GR, individual incentive constraint is defined as a social norm under which, given that other players follow the norm, no player has an incentive to deviate from the norm. In addition, bilateral rationality is defined as a social norm under which, given that other players follow the norm, no matched pair of players who have followed the norm can improve their payoffs by making a joint change from the norm. In our model, bilateral rationality, as a result, excludes Pareto-dominated outcomes on each of the on-the-equilibrium paths. Note that bilateral rationality does not mean *renegotiation-proofness*,[13] since renegotiation-proofness is meant to exclude Pareto-dominated outcomes on each of the off-the-equilibrium paths.

---

[12] If we assume that the relationship will be exogenously broken up with a probability $\theta > 0$ regardless of players' breakup decisions, then we can easily show that a constant π is feasible. In addition, given any $\pi > 0$ and any positive number $\varepsilon > 0$, we can find an exogenous breakup probability $\theta > 0$ such that $\varepsilon > 0$ and θ makes π a constant proportion of H-players in the pool over time. Therefore, the steady state in which $\pi > 0$ and $\theta = 0$ can be interpreted as the limit of the exogenous breakup cases. A formal proof of the statement above is provided in the Appendix.

[13] To focus on folk theorem, we do not consider this renegotiation-proofness in our study. Studies related to renegotiation-proof social norms, however, are available from the author.

In the current model, there are two kinds of phases. The first kind happens when two players are newly matched, and thus they do not know their partners' types. GR referred to these phases as S-phases which denote matches between strangers. In this case, the players "test" their partners to find out their partners' types, and naturally they would play relatively small cooperative actions. The other kind of phase happens when two players in a common match know their partners' types. Hence, these phases arise right after the S-phases. GR referred to these phases as F-phases in which players can seek friendship with their partners. Accordingly, players would play the highest possible cooperative actions in the F-phases. Each kind of phases can consist of two H-players, two L-players, or an H-player and an L-player. In S-phases, however, players play the same actions regardless of their partners' types, since they cannot distinguish their partners' types. As a result, we need to consider only five phases, which are 1) an F-phase between two H-players, 2) an S-phase from an H-player's viewpoint, 3) an F-phase between two L-players, 4) an S-phase from an L-player's viewpoint, and 5) an F-phase between an H-player and an L-player.

First, we apply the two criteria, that is, the individual incentive constraint and the bilateral rationality, to the phase in which two H-players are matched into a pair and they are aware of their partners' types.[14] In this phase, let x_H denote a present value to an H-player when he is in the pool of unmatched players. Then, H-players solve the following optimization problem; given $0 \le x_H \le \hat{V}_H^S$ where $\hat{V}_H^S \equiv \frac{1}{\delta}\max_{a\in[0,\bar{a}]}\{\frac{\Pi_{HH}(a,a)}{1-\delta} - \Pi_{HH}(0,a)\}$,

$$\max_{a\in[0,\bar{a}]} \frac{\Pi_{HH}(a,a)}{1-\delta} \equiv V_H^F(x_H) \tag{1}$$

$$s.t. \quad \frac{\Pi_{HH}(a,a)}{1-\delta} \ge \Pi_{HH}(0,a) + \delta x_H . \tag{2}$$

Given a present value to an H-player, this optimization problem yields the highest possible cooperation level, which, therefore, satisfies bilateral rationality, among all the cooperation levels that satisfy individual incentive constraint. In particular, the objective function (1) requires H-players to choose the highest actions, and the constraint (2) shows the range of the possible actions that prevent individual deviation.

Second, based on the optimization problem above, the two criteria are applied to the phase for an H-player when he is newly matched and thus he does not know his partner's type; let a_L^S denote an action of an L-player when he is newly matched,

---

[14] This phase is not necessary for the players to play cooperative actions in equilibrium. In fact, if the players know their partners' types before they play, then they can play optimal cooperative actions from the beginning of their matches.

then given x_H and $a_L^S \in [0,\tilde{a}]$,

$$\max_{a\in[0,\tilde{a}]} \pi\{\Pi_{HH}(a,a)+\delta V_H^F(x_H)\}+(1-\pi)\{\Pi_{HL}(a,a_L^S)+\delta x_H\} \equiv V_H^S(x_H,a_L^S) \tag{3}$$

$$\begin{aligned} s.t.\ & \pi\{\Pi_{HH}(a,a)+\delta V_H^F(x_H)\}+(1-\pi)\{\Pi_{HL}(a,a_L^S)+\delta x_H\} \\ & \geq \pi\{\Pi_{HH}(0,a)+\delta x_H\}+(1-\pi)\{\Pi_{HL}(0,a_L^S)+\delta x_H\}. \end{aligned} \tag{4}$$

Specifically, the objective function (3) demands that H-players choose the highest actions in this phase, and the constraint (4) reveals the range of the possible actions that prevent individual deviation of H-players.

Similarly, the two criteria are applied to the phases for L-players. Third, two L-players who are certain that their partners are L-players solve the following problem; let x_L denote a present value to an L-player when he is in the pool of unmatched players, then given $0 \leq x_L \leq \hat{V}_L^S$ where $\hat{V}_L^S \equiv \frac{1}{\delta}\max_{a\in[0,\tilde{a}]}\{\frac{\Pi_{LL}(a,a)}{1-\delta} - \Pi_{LL}(0,a)\}$,

$$\max_{a\in[0,\tilde{a}]} \frac{\Pi_{LL}(a,a)}{1-\delta} \equiv V_L^F(x_L) \tag{5}$$

$$s.t.\ \frac{\Pi_{LL}(a,a)}{1-\delta} \geq \Pi_{LL}(0,a)+\delta x_L. \tag{6}$$

Like in the case of the H-players, the objective function (5) seeks the highest actions of the L-players and, the constraint (6) presents the range of the possible actions that are proof against individual deviation of L-players.

Fourth, an L-player who is newly matched solves the following problem; let a_H^S denote an action of an H-player when he is newly matched, then given x_L and $a_H^S \in [0,\tilde{a}]$,

$$\max_{a\in[0,\tilde{a}]} \pi\{\Pi_{LH}(a,a_H^S)+\delta x_L\}+(1-\pi)\{\Pi_{LL}(a,a)+\delta V_L^F(x_L)\} \equiv V_L^S(x_L,a_H^S) \tag{7}$$

$$\begin{aligned} s.t.\ & \pi\{\Pi_{LH}(a,a_H^S)+\delta x_L\}+(1-\pi)\{\Pi_{LL}(a,a)+\delta V_L^F(x_L)\} \\ & \geq \pi\{\Pi_{LH}(0,a_H^S)+\delta x_L\}+(1-\pi)\{\Pi_{LL}(0,a)+\delta x_L\}. \end{aligned} \tag{8}$$

Again, like in the case of the H-players, the objective function (7) requests the highest actions in this phase, and the constraint (8) denotes the range of the possible actions that are proof against individual deviation.

Finally, the two criteria are applied to the phase in which an H-player and an L-player are matched into a pair and they are aware of their partners' types. In equilibrium, players could have long-term cooperative relationships in the previous four phases only if they cannot achieve cooperation in this phase. So, given present values in the pool of unmatched players, we need to show that every cooperation level that satisfies the individual incentive constraint does not give an H-player or an L-player a greater payoff than their present values in the pool. This condition is formalized at the condition 9 in Definition 1.

Now, we are ready to define our equilibrium, which we call a "Social Equilibrium." This social equilibrium is adopted and adapted from GR.

Definition 1 *A social equilibrium is a collection of actions* ($a_H^F, a_H^S, a_L^F, a_L^S$) *and payoffs* ($V_H^F, V_H^S, V_L^F, V_L^S$) *such that*

1. *given* V_H^S, a_H^F *solves (1) subject to (2);*
2. *given* V_H^S *and* a_L^S, a_H^S *solves (3) subject to (4);*
3. *given* V_L^S, a_L^F *solves (5) subject to (6);*
4. *given* V_L^S *and* a_H^S, a_L^S *solves (7) subject to (8);*
5. *the payoff* V_H^F *equals the maximum value* $V_H^F(V_H^S)$;
6. *the payoff* V_H^S *equals the maximum value* $V_H^S(V_H^S, a_L^S)$;
7. *the payoff* V_L^F *equals the maximum value* $V_L^F(V_L^S)$;
8. *the payoff* V_L^S *equals the maximum value* $V_L^S(V_L^S, a_H^S)$;

and for all a', $a'' \in [0, \tilde{a}]$,

$$9.\quad \text{if } \frac{\Pi_{HL}(a', a'')}{1-\delta} \geq \Pi_{HL}(0, a'') + \delta V_H^S, \text{ then } V_H^S \geq \frac{\Pi_{HL}(a', a'')}{1-\delta} \text{ or} \tag{9}$$

$$\text{if } \frac{\Pi_{LH}(a'', a')}{1-\delta} \geq \Pi_{LH}(0, a') + \delta V_L^S, \text{ then } V_L^S \geq \frac{\Pi_{LH}(a'', a')}{1-\delta}. \tag{10}$$

In Definition 1, the inequality (9) embodies the condition under which an H-player can achieve a higher payoff in the pool of unmatched players than in the long-term cooperative relationship with an L-player. In addition, the inequality (10) provides the condition under which an L-player can achieve a higher payoff in the pool of unmatched players than in the long-term cooperative relationship with an H-player.

V. Results

In this study, the results are similar to GR's in respect to the factors that can influence the level of cooperation in equilibrium. In both studies, cooperation is enhanced when players find their proper matches or when the discount factor goes up. However, while GR's results apply to partial population only, the following results show that a long-term cooperative relationship among the whole population is possible. The first result shows that there exists a social equilibrium. Like in GR, special assumptions on payoff functions are used for the existence of the equilibrium. Note that only Proposition 1 uses these special assumptions.

Assumption 1. For each $J \in \{H, L\}$ and action $a \in [0, \tilde{a}]$, the payoff function $\Pi_{JJ}(a,a)$ is strictly concave in *a*, the function $\Pi_{JJ}(a,0)$ is concave in *a*, and the function $\Pi_{JJ}(0,a)$ is convex in *a*.

Assumption 2. The left-hand partial derivatives of $\Pi_{HL}(a_1,a_2)$ and $\Pi_{LH}(a_1,a_2)$ with respect to the first argument a_1 are continuous in the second argument a_2.

Assumptions 1 and 2 guarantee that the optimization functions $V_H^S(\cdot,\cdot)$ and $V_L^S(\cdot,\cdot)$ and the optimizers in these functions are continuous in their arguments. This property of continuity serves as a stepping stone for the existence of a fixed point in the optimization problems above.

Assumption 3. For each $IJ \in \{HL, LH\}$, the payoff function $\Pi_{IJ}(a_1,a_2)$ is concave in the first argument a_1 and convex in the second argument a_2, and for $a_1 > 0, \Pi_{IJ}(0,a_2) - \Pi_{IJ}(a_1,a_2) \le \Pi_{IJ}(0,a_2') - \Pi_{IJ}(a_1,a_2')$ if $a_2 > a_2'$.

Assumption 3 implies that in the different-type matches, (*i.e.* the H-L matches,) the payoff $\Pi_{IJ}(a_1,a_2)$ decreases with his own action a_1 at an increasing rate and increases with his partner's action a_2 at an increasing rate. In addition, when a_1 is positive, the payoff difference $\Pi_{IJ}(0,a_2) - \Pi_{IJ}(a_1,a\ )$ decreases in a_2, which shows a strategic complementary relationship between two optimal actions a_1 and a_2 in equilibrium. That is, if $\Pi_{IJ}(\cdot,\cdot)$ is twice continuously differentiable, this condition means that the cross partial derivatives of $\Pi_{IJ}(\cdot,\cdot)$ are non-negative. This assumption ensures nice behavior of the players by embodying the situation in which an optimal action by one player would not decrease when the other player in a common match increases his action.

Under Assumptions 1, 2, and 3, Proposition 1 below shows a sufficient condition for the existence of a social equilibrium. The sufficient condition consists of two subconditions. The first subcondition is referred to as Condition E, like in GR, and we present the exact form of this Condition E in the Appendix. This Condition E formulates the situation in which there are appropriate proportions of H-players and L-players in the pool of unmatched players.

In this model, the proportions of H-players and L-players in the pool of unmatched players affect both the scarcity value of the H-H match, V_H^F, and the scarcity value of the L-L match, V_L^F. Accordingly, the possibility of the existence of a social equilibrium depends on these proportions. To illustrate their relationships, suppose that the proportion of H-players in the pool increases. Then, this raises the probability for an H-player to meet another H-player in the pool, and thus the present value to an H-player in the pool, x_H, increases as well. Note that the scarcity value of the H-H match is viewed as the difference between V_H^F and x_H. Hence, the increase in the number of H-players in the pool, x_H, reduces the scarcity value of the H-H match, V_H^F. However, if the scarcity value of the H-H match falls, then the threat of breaking the relationship might not be effective. As a result, H-players cannot sustain their long-term cooperative relationships, which results in the non-existence of a social equilibrium. Therefore, in Condition E, (14) and (15) embody the situation in which H-players sustain their proper numbers in the unmatched pool.

Likewise, a high proportion of L-players in the pool can ruin the possibility of the existence of an equilibrium. Thus, (16) and (17) make L-players maintain their numbers in the pool. Consequently, four conditions in Condition E guarantee appropriate proportions of H-players and L-players in the pool of unmatched players, and so it permits the existence of a social equilibrium. Note that GR provided only two conditions for non-myopic players because only non-myopic players can cooperate in equilibrium. The current model, on the other hand, requires four conditions for both H-players and L-players because all players can cooperate in equilibrium.[15]

The second subcondition is related to the ability difference between an H-player and an L-player. To sustain a social equilibrium, a fixed point in the optimization problems above has to satisfy the condition 9 in Definition 1 in which one of the types has no incentive to cooperate with the other type. If the ability difference between an H-player and an L-player is *wide enough*, then the H-player would have no incentive to cooperate with the L-player, and therefore, the fixed point would satisfy the condition 9 in Definition 1. Definition 2 below provides the level of the ability difference in which an H-player has no incentive to cooperate with an L-player.

Definition 2 *Define* a_H^4 *as the value of a such that* $(1-\delta\pi)\Pi_{LH}(0,a)=\delta(1-\pi)\Pi_{LL}(a_L^1,a_L^1)$. *The ability difference between an H-player and an L-player is said to be* ***wide enough*** *if* $\delta\pi\Pi_{HH}(a_H^1,a_H^1)\geq(1-\delta+\delta\pi)\Pi_{HL}(a_H^4,\tilde{a})$ *whenever* a_H^4 *exists.*

In Definition 2, a_H^4 denotes the value of an H-player's minimum action that endows L-players with an incentive to cooperate with H-players. Thus, the wide enough ability difference means that, in the H-L match, even when an L-player plays the full cooperative action $\tilde{a}$ and an H-player plays the minimum required action a_H^4, the H-player still has an incentive to follow a social norm in which every H-player cooperates only with another H-player. Note that this condition

[15] For the mathematical details of this condition, please refer to GR.

could be regarded as *no-mimicking condition*. In the H-L match, without this condition, an L-player might have an incentive to mimic an H-player by giving his high ability partner the same payoff as in the H-H match. Under this condition, however, an L-player cannot mimic an H-player in the H-L match because of his small contribution to the total payoffs. Therefore, this wide enough ability difference excludes the possibility of an L-player's mimicking an H-player.

Proposition 1 *Under Assumptions 1, 2, and 3, a social equilibrium exists if Condition E holds and the ability difference between an H-player and an L-player is wide enough.*

Proof. See Appendix. ▪

Examples with specific payoff functions can be found in GR. The payoff functions from GR, however, have to be adapted for the L-players. In GR, the myopic type, who has the zero discount factor, has no incentive to play any positive action. The zero action by the myopic type lowers a present value to the non-myopic type in the pool of unmatched players, and this lowered present value in turn makes an ongoing cooperative relationship more valuable. As a result, although an one-period payoff from betrayal is high, the non-myopic type players can sustain a long-term cooperative relationship among themselves. In the present model, on the other hand, when H-players are newly matched with L-players, they play positive cooperative actions a_H^S. Since the H-players' actions a_H^S significantly improve present values to L-players in the pool, if one-period payoffs to L-players when they betray other L-players are as high as those in GR, then L-players would prefer betraying their low-ability partners more than cooperating with them. Therefore, the payoff functions from GR need to be modified so that L-players can sustain long-term cooperative relationships among themselves.

The second result describes cooperation levels in each phase in equilibrium. Each type of the players faces two possible phases in which they can play different levels of cooperation. First, each type reaches the first phase right after they confirm that their partners are of the same types as themselves. Next, each type reaches the second phase right after they are newly matched, and thus in this phase, they do not know their partners' types. Proposition 2 below shows that each type plays a higher cooperative action in the former phase than in the latter phase, except that he achieves the same level of cooperation when he plays full cooperative actions in both phases.

According to the interpretation of GR, Proposition 2 characterizes a social equilibrium divided into a "testing phase" and a "cooperation phase." In the testing phase, the players are "cautious," and as a result, they have less to achieve. If they confirm that they are matched with the same type players as themselves, then they move into the cooperation phase where they can play at greater cooperation levels.

Proposition 2 indeed complies with an intuitive feature of a social equilibrium. When the players are newly matched, their actions are confined so that their payoffs in these S-phases are low enough to endow the H-H match and the L-L match in the F-phases with a scarcity value. Then, the players can use this scarcity value to sustain their long-term cooperative relationships.

Proposition 2 *In a social equilibrium,* $a_J^F \geq a_J^S$ *where* $J \in \{H, L\}$ *with strict inequality holding whenever* $a_J^F < \tilde{a}$.

Proof. Consider an H-player case. If $a_H^F = \tilde{a}$, then it is trivial. Let $a_H^F < \tilde{a}$ in equilibrium. By way of contradiction, suppose that $a_H^F \leq a_H^S$. Then, we have that $\Pi_{HH}(0, a_H^S) + \delta V_H^S \geq \frac{\Pi_{HH}(a_H^S, a_H^S)}{1-\delta}$ by the constraint (2). Then,

$$\begin{aligned}
&(1-\delta)\{\Pi_{HH}(0, a_H^S) - \Pi_{HH}(a_H^S, a_H^S)\} \geq \delta\{\Pi_{HH}(a_H^S, a_H^S) - (1-\delta)V_H^S\} \\
&\geq \delta\{\Pi_{HH}(a_H^F, a_H^F) - (1-\delta)V_H^S\} = \delta(1-\delta)\{V_H^F - V_H^S) \\
&> \delta(1-\delta)(V_H^F - V_H^S) + \frac{1-\pi}{\pi}(1-\delta)\{\Pi_{HL}(a_H^S, a_L^S) - \Pi_{HL}(0, a_L^S)\}
\end{aligned}$$

where the fact $\Pi_{HL}(a_H^S, a_L^S) - \Pi_{HL}(0, a_L^S) < 0$ is used at the last inequality. This contradicts (4). Therefore, we have $a_H^F > a_H^S$. Similarly, we can show $a_L^F \geq a_L^S$ with strict inequality holding whenever $a_L^F < \tilde{a}$. ▪

The final result goes one step further from GR's. In their paper, as players become infinitely patient, the cooperation level in equilibrium approaches full cooperation once players find their proper matches. In the present model, Proposition 3 below states that *when players are sufficiently patient, they play the maximal cooperation level in equilibrium right after they check that they are matched with the same type partners as themselves*. The present study proposes Proposition 3 as a folk theorem of this model.

Proposition 3 (Folk Theorem) *There exists a discount factor* $\delta^* < 1$ *such that for any* $\delta \in [\delta^*, 1)$, $a_H^F = a_L^F = \tilde{a}$ *in a social equilibrium under* δ, *whenever the social equilibrium aexists.*

Proof. By way of contradiction, suppose not. Then, for any $\delta < 1$, there exists $1 > \delta' \geq \delta$, such that under δ', there exists a social equilibrium with $a_H^F < \tilde{a}$ or $a_L^F < \tilde{a}$. First, consider the case in which for any $\delta < 1$, there exists $1 > \delta' \geq \delta$ such that under δ', there exists a social equilibrium with $a_H^F < \tilde{a}$. In the social equilibrium under the discount factor δ', let V_H^S be a present value to an H-player in the pool of unmatched players. Then, according to the constraint (2), we have that

$$\frac{\Pi_{HH}(\tilde{a},\tilde{a})}{1-\delta'} < \Pi_{HH}(0,\tilde{a}) + \delta' V_H^S . \tag{11}$$

In addition, we have that

$$V_H^S \leq \frac{1}{\delta'}\{\frac{\Pi_{HH}(a_H^1, a_H^1)}{1-\delta'} - \Pi_{HH}(0, a_H^1)\} \tag{12}$$

where a_H^1 is a maximizer of $\Pi_{HH}(a,a)-(1-\delta)\Pi_{HH}(0,a)$. Note that since $\frac{1}{\delta'}\{\frac{\Pi_{HH}(\tilde{a},\tilde{a})}{1-\delta'} - \Pi_{HH}(0,\tilde{a})\} < V_H^S \leq \frac{1}{\delta'}\{\frac{\Pi_{HH}(a_H^1,a_H^1)}{1-\delta'} - \Pi_{HH}(0,a_H^1)\}$, we have that $a_H^1 < \tilde{a}$. By combining (11) with (12), we have that

$$\frac{\Pi_{HH}(\tilde{a},\tilde{a})}{1-\delta'} < \Pi_{HH}(0,\tilde{a}) + \frac{\Pi_{HH}(a_H^1, a_H^1)}{1-\delta'} - \Pi_{HH}(0, a_H^1)$$

$$\Leftrightarrow \frac{1}{1-\delta'}\{\Pi_{HH}(\tilde{a},\tilde{a}) - \Pi_{HH}(a_H^1, a_H^1)\} < \Pi_{HH}(0,\tilde{a}) - \Pi_{HH}(0, a_H^1) . \tag{13}$$

However, since the inequality (13) holds for any $\delta < 1$ and thus for any $1 > \delta' \geq \delta$, (13) is a contradiction. Similarly, we can show that it is a contradiction that for any $\delta < 1$, there exists $1 > \delta' \geq \delta$ such that under δ', there exists a social equilibrium with $a_L^F < \tilde{a}$. This completes the proof. ▪

The present model is different from GR's model in two aspects. First, in GR, if there exists a social equilibrium, it must be unique, because one type always prefers to play a zero action and the other type has only one best response to the zero action. In the present model, however, there could be multiple social equilibria, because players can play different levels of initial actions (a_H^S, a_L^S) in equilibrium, which in turn affect the present values to each type x_H and x_L in the pool of unmatched players. Since equilibrium actions (a_H^F, a_L^F) and payoffs (V_H^F, V_L^F) are determined by x_H and x_L, various initial actions eventually permit multiple social equilibria. Second, in GR, a change in π, the proportion of non-myopic players of unmatched players, directly influences the payoffs. An increase in π results in an increase in the present values when players are in the unmatched pool, and this also results in a non-increase in the payoffs to non-myopic players when they find non-myopic partners. In the present model, however, due to the possible existence of multiple equilibria, a change in π, the proportion of H-players in the pool of unmatched players, does not have a clear effect on the payoffs $(V_H^F, V_H^S, V_L^F, V_L^S)$. This is because the impact from a change in π could be diluted by the influence from a change in equilibria. For example, an increase in π causes the H-players' payoff V_H^S to increase, but this influence could be canceled out by a change in equilibria from a high value of V_H^S to a low value of it.

However, if the players are informed of their partners' types as well as their own types, then there could be at most one social equilibrium in the model. This is because the players in the matches of the same types play a_J^F where $J \in \{H, L\}$ from the very first period of their matches. In this case, therefore, an increase in π results in an increase in the present values for H-players when they are in the unmatched pool, and results in a non-increase in their payoffs when they are matched with H-partners.

VI. Conclusion

A long-term cooperative behavior that is robust to both pair-wise deviation and individual deviation is possible among the whole population in equilibrium. Regarding cooperation levels, after players play a lower cooperation in the testing phase, they move on to higher cooperation in the cooperation phase. If players are patient enough, both H-players and L-players can achieve full cooperation once they find their proper matches. Therefore, based on players' preferences for their partners' ability, a long-term cooperative relationship among the whole population is possible in equilibrium.

We can extend the current model by introducing more types than two and could still maintain a result similar to the ones in this model. That is, the whole population in multiple-type models could show long-term cooperative behavior that is robust to both pair-wise deviation and individual deviation. In the extended model, however, we could find various cooperation patterns based on various payoff systems. For example, suppose there are n number of types in a model. Then, some payoff system might allow long-term cooperative behavior only between the same types of players as in the current model. Another payoff system, on the other hand, could cause i-type players to cooperate with $i-1$-type players and $i+1$-type players as well as with i-type players.

Appendix

1. Condition E.

Like in GR, the notations below are used to simplify the sufficient condition. First, denote by a_H^1 and a a_L^1 the maximizers of the functions $\Pi_{HH}(a,a)-(1-\delta)\Pi_{HH}(0,a)$ and $\Pi_{LL}(a,a)-(1-\delta)\Pi_{LL}(0,a)$, respectively. Next, let a_H^2 and a_L^2 denote the maximum values of a s.t.

$$\pi\{\Pi_{HH}(0,a)-\Pi_{HH}(a,a)\}+(1-\pi)\{\Pi_{HL}(0,\tilde{a})-\Pi_{HL}(a,\tilde{a})\}$$
$$\leq \pi\{\Pi_{HH}(0,a_H^1)-\Pi_{HH}(a_H^1,a_H^1)\} \quad \text{and}$$

$$\pi\{\Pi_{LH}(0,\tilde{a})-\Pi_{LH}(a,\tilde{a})\}+(1-\pi)\{\Pi_{LL}(0,a)-\Pi_{LL}(a,a)\}$$
$$\leq (1-\pi)\{\Pi_{LL}(0,a_L^1)-\Pi_{LL}(a_L^1,a_L^1)\}, \text{ respectively.}$$

Finally, let a_H^3 and a_L^3 denote the maximizers of the strictly concave functions $\pi\Pi_{HH}(a,a)+(1-\pi)\Pi_{HL}(a,\tilde{a})$ and $\pi\Pi_{LH}(a,\tilde{a})+(1-\pi)\Pi_{LL}(a,a)$, respectively. Then, under Assumptions 1, 2, and 3, here is the sufficient condition for the existence of a social equilibrium.

Condition E If $a_H^3 \leq a_H^2$, then

$$\pi\Pi_{HH}(a_H^3,a_H^3)+(1-\pi)\Pi_{HL}(a_H^3,\tilde{a})$$
$$\leq (\pi+\frac{1}{\delta})\Pi_{HH}(a_H^1,a_H^1)+(1-\pi-\frac{1}{\delta})\Pi_{HH}(0,a_H^1). \tag{14}$$

If $a_H^3 > a_H^2$, then

$$\delta\pi\Pi_{HH}(0,a_H^2)+\delta(1-\pi)\Pi_{HL}(0,\tilde{a})\leq \Pi_{HH}(a_H^1,a_H^1)-(1-\delta)\Pi_{HH}(0,a_H^1). \tag{15}$$

If $a_L^3 \leq a_L^2$, then

$$\pi\Pi_{LH}(a_L^3,\tilde{a})+(1-\pi)\Pi_{LL}(a_L^3,a_L^3)$$
$$\leq (1-\pi+\frac{1}{\delta})\Pi_{LL}(a_L^1,a_L^1)+(\pi-\frac{1}{\delta})\Pi_{LL}(0,a_L^1). \tag{16}$$

If $a_L^3 > a_L^2$, then

$$\delta\pi\Pi_{LH}(0,\tilde{a})+\delta(1-\pi)\Pi_{LL}(0,a_L^2)\le\Pi_{LL}(a_L^1,a_L^1)-(1-\delta)\Pi_{LL}(0,a_L^1)\,. \tag{17}$$

2. Proof of Proposition 1.

Proof. Consider (1) subject to (2). Note that $V_H^F(x_H)$ is continuous and non-increasing in x_H. Also, we have that $V_H^F(x_H)-x_H>0$ because of (2). Let

$$\hat{V}_H^S\equiv\frac{1}{\delta}\max_{a\in(0,\tilde{a}]}\{\frac{\Pi_{HH}(a,a)}{1-\delta}-\Pi_{HH}(0,a)\}\quad\text{and}$$
$$\hat{V}_L^S\equiv\frac{1}{\delta}\max_{a\in(0,\tilde{a}]}\{\frac{\Pi_{LL}(a,a)}{1-\delta}-\Pi_{LL}(0,a)\}\,.$$

Then $\hat{V}_H^S$ and $\hat{V}_L^S$ exist and are positive because $\Pi_{JJ}(a,a)>(1-\delta)\Pi_{JJ}(0,a)$ for some $a>0$ where $J\in\{H,L\}$. Given $x_H\in[0,\hat{V}_H^S]$ and $a_L^S\in[0,\tilde{a}]$, define

$$\begin{aligned}
&a_H^S(x_H,a_L^S)\in\arg\max_{a\in[0,\tilde{a}]}\pi\{\Pi_{HH}(a,a)+\delta V_H^F(x_H)\}+(1-\pi)\{\Pi_{HL}(a,a_L^S)+\delta x_H\}\\
&\quad s.t.\ \pi\{\Pi_{HH}(0,a)+\delta x_H\}+(1-\pi)\{\Pi_{HL}(0,a_L^S)+\delta x_H\}\\
&\le\ \pi\{\Pi_{HH}(a,a)+\delta V_H^F(x_H)\}+(1-\pi)\{\Pi_{HL}(a,a_L^S)+\delta x_H\}.
\end{aligned}$$

Let $a_H^2(x_H,a_L^S)$ be the maximum value of *a* such that

$$\begin{aligned}
&\pi\{\Pi_{HH}(0,a)+\delta x_H\}+(1-\pi)\{\Pi_{HL}(0,a_L^S)+\delta x_H\}\\
&\le\ \pi\{\Pi_{HH}(a,a)+\delta V_H^F(x_H)\}+(1-\pi)\{\Pi_{HL}(a,a_L^S)+\delta x_H\}\\
&\Leftrightarrow \pi\{\Pi_{HH}(0,a)-\Pi_{HH}(a,a)\}+(1-\pi)\{\Pi_{HL}(0,a_L^S)+\Pi_{HL}(a,a_L^S)\}\\
&\le\ \delta\pi\{V_H^F(x_H)-x_H\}\,.
\end{aligned}$$

Then, $a_H^2(x_H,a_L^S)>0$ since $V_H^F(x_H)-x_H>0$. Also, $a_H^2(x_H,a_L^S)$ is continuous in x_H and a_L^S because 1) $\pi\{\Pi_{HH}(0,a)-\Pi_{HH}(a,a)\}+(1-\pi)\{\Pi_{HL}(0,a_L^S)$ $-\Pi_{HL}(a,a_L^S)\}$ is strictly increasing in a and continuous in a and a_L^S; and 2) $\delta\pi\{V_H^F(x_H)-x_H\}$ is continuous in x_H. In addition, let $a_H^3(a_L^S)$ denote the maximizer of the strictly concave function $\pi\Pi_{HH}(a,a)+(1-\pi)\Pi_{HL}(a,a_L^S)$. Then $a_H^3(a_L^S)>0$ because given any $a_L^S\in[0,\tilde{a}]$, there exists $a>0$ s.t. $\pi\{\Pi_{HH}(a,a)+$ $(1-\pi)\Pi_{HL}(a,a_L^S)>(1-\pi)\Pi_{HL}(0,a_L^S)$. The function $\Pi_{HL}(a_1,a_2)$ is concave in a_1, and therefore, its left-hand partial derivative with respect to a_1, $\frac{\partial\Pi_{HL}(a_1-0,a_2)}{\partial a_1}$, is well-defined on $(0,\tilde{a}]^2$. According to Assumption 2, $\frac{\partial\Pi_{HL}(a_1-0,a_2)}{\partial a_1}$ is continuous in a_2. Also, the function $\Pi_{HH}(a,a)$ is strictly concave in a. Therefore, $a_H^3(a_L^S)$ is continuous in a_L^S. Note that a $a_H^S(x_H,a_L^S)=\min\{a_H^2(x_H,a_L^S),a_H^3(a_L^S)\}$. Since $a_H^2(x_H,a_L^S)$ and $a_H^3(a_L^S)$ are positive and continuous in x_H and a_L^S, so is

$a_H^S(x_H, a_L^S)$. Define

$$\Phi_H(x_H, a_L^S) \equiv \max_{a\in[0,\tilde{a}]} \pi\{\Pi_{HH}(a,a)+\delta V_H^F(x_H)\}+(1-\pi)\{\Pi_{HL}(a,a_L^S)+\delta x_H\}$$
$$s.t.\ \pi\{\Pi_{HH}(0,a)+\Pi_{HH}(a,a)\}+(1-\pi)\{\Pi_{HL}(0,a_L^S)-\Pi_{HL}(a,a_L^S)\}$$
$$\le \delta\pi\{V_H^F(x_H)-x_H\}.$$

Then $\Phi_H(x_H, a_L^S) = \pi\{\Pi_{HH}(a_H^S(x_H,a_L^S), a_H^S(x_H,a_L^S))+\delta V_H^F(x_H)\}+(1-\pi)\{\Pi_{HL}$ $(a_H^S(x_H,a_L^S), a_L^S)+\delta x_H\}$, and $\Phi_H(x_H,a_L^S)$ is continuous in x_H and a_L^S. Similarly, we can define $a_L^S(x_L,a_H^S)$ and $\Phi_L(x_L,a_H^S)$ for an L-player and show that $a_L^S(x_L,a_H^S)$ and $\Phi_L(x_L,a_H^S)$ are continuous in x_L and a_H^S.

If

$$\max_{a_L^S}\{\Phi_H(\hat{V}_H^S, a_L^S)\}\le \hat{V}_H^S \text{ and} \tag{18}$$
$$\max_{a_H^S}\{\Phi_L(\hat{V}_L^S, a_H^S)\}\le \hat{V}_L^S, \tag{19}$$

then by using the values $a_H^S(x_H,a_L^S), a_L^S(x_L,a_H^S)$, $\min\{\Phi_H(x_H,a_L^S),\hat{V}_H^S\}$, and $\min\{\Phi_L(x_L,a_H^S),\hat{V}_L^S\}$, we can construct a continuous function from $[0,\tilde{a}]^2\times$ $[0,\hat{V}_H^S]\times[0,\hat{V}_L^S]$ into $[0,\tilde{a}]^2\times[0,\hat{V}_H^S]\times[0,\hat{V}_L^S]$ such that a fixed point of the function, whose existence is guaranteed by Brouwer's Fixed Point Theorem, solves (3) subject to (4) and solves (7) subject to (8). Therefore, to complete the proof, we must check that (18) and (19) are equivalent to Condition E and also should show that if the ability difference between an H-player and an L-player is wide enough, then the fixed point satisfies the condition 9 in Definition 1.

First, check that (18) and (19) are equivalent to Condition E. Note that

$$\hat{V}_H^S = \frac{1}{\delta}\{\frac{\Pi_{HH}(a_H^1,a_H^1)}{1-\delta}-\Pi_{HH}(0,a_H^1)\}.$$

In addition, note that $\Phi_H(x_H,\tilde{a})\ge\Phi_H(x_H,a_L^S)$ for every $a_L^S\in[0,\tilde{a}]$ since $\Pi_{HL}(0,a_2)-\Pi_{HL}(a_1,a_2)\le\Pi_{HL}(0,a_2')-\Pi_{HL}(a_1,a_2')$ if $a_2>a_2'$. Therefore, if $a_H^3(\tilde{a})$ $\le a_H^2(\hat{V}_H^S,\tilde{a})$, *i.e.* $a_H^3\le a_H^2$, then

$$\max_{a_L^S}\{\Phi_H(\hat{V}_H^S,a_L^S)\}\le\hat{V}_H^S \Leftrightarrow \Phi(\hat{V}_H^S,\tilde{a})\le\hat{V}_H^S$$
$$\Leftrightarrow \pi\Pi_{HH}(a_H^3,a_H^3)+(1-\pi)\Pi_{HL}(a_H^3,\tilde{a})$$
$$+\delta\pi\frac{\Pi_{HH}(a_H^1,a_H^1)}{1-\delta}+\delta\frac{(1-\pi)}{\delta}\{\frac{\Pi_{HH}(a_H^1,a_H^1)}{1-\delta}-\Pi_{HH}(0,a_H^1)\}$$

$$\le \frac{1}{\delta}\{\frac{\Pi_{HH}(a_H^1, a_H^1)}{1-\delta} - \Pi_{HH}(0, a_H^1)\},$$

which is equivalent to (14). If $a_H^3 > a_H^2$, then by the definition of a_H^2,

$$\begin{aligned}
&\Phi_H(\hat{V}_H^S, \tilde{a}) \le \hat{V}_H^S \\
&\Leftrightarrow \pi\Pi_{HH}(0, a_H^2) + (1-\pi)\Pi_{HL}(0, \tilde{a}) \\
&+\frac{1-\pi+\delta\pi}{1-\delta}\Pi_{HH}(a_H^1, a_H^1) - (1-\pi)\Pi_{HH}(0, a_H^1) \\
&\le \frac{1}{\delta}\{\frac{\Pi_{HH}(a_H^1, a_H^1)}{1-\delta} - \Pi_{HH}(0, a_H^1)\},
\end{aligned}$$

which is equivalent to (15). Similarly, we can show that (19) is equivalent to (16) and (17). Therefore, (14), (15), (16), and (17) are a sufficient condition for the existence of a fixed point.

Finally, we need to show that the fixed point satisfies (9) and (10). Let V_H^S and V_L^S be parts of the fixed point such that V_H^S and V_L^S satisfy the respective conditions 6 and 8 in Definition 1. Then, we have $V_H^S(V_H^S, a_L^S) = V_H^S$ and $V_L^S(V_L^S, a_H^S) = V_L^S$ for some a_L^S and a_H^S. Note that $V_H^F(x_H)$ and $V_L^F(x_L)$ are non-increasing in x_H and x_L, respectively, and thus, $V_H^F(x_H') \ge V_H^F(\hat{V}_H^S)$ and $V_L^F(x_L') \ge V_L^F(\hat{V}_L^S)$ for any x_H' and x_L'. By (3) and (7), we have that

$$V_H^S \ge \delta\pi V_H^F(\hat{V}_H^S) + \delta(1-\pi)V_H^S \Leftrightarrow V_H^S \ge \frac{\delta\pi}{1-\delta+\delta\pi}\frac{\Pi_{HH}(a_H^1, a_H^1)}{1-\delta} \tag{20}$$

$$\text{and } V_L^S \ge \delta\pi V_L^S + \delta(1-\pi)V_L^F(\hat{V}_L^S) \Leftrightarrow V_L^S \ge \frac{\delta(1-\pi)}{1-\delta\pi}\frac{\Pi_{LL}(a_L^1, a_L^1)}{1-\delta}. \tag{21}$$

Suppose that $a', a'' \in [0, \tilde{a}]$ satisfy the premises of (9) and (10). Then,

$$\begin{aligned}
&\frac{\Pi_{LH}(0, a')}{1-\delta} \ge \frac{\Pi_{LH}(a'', a')}{1-\delta} \ge \Pi_{LH}(0, a') + \delta V_L^S \\
&\Rightarrow \delta\frac{\Pi_{LH}(0, a')}{1-\delta} \ge \delta\frac{\delta(1-\pi)}{1-\delta\pi}\frac{\Pi_{LL}(a_L^1, a_L^1)}{1-\delta}
\end{aligned} \tag{22}$$

where (21) is used at the inequality (22). Since the function $\Pi_{LH}(0, a)$ is convex, from (22), we can find that there exists a_H^4 and that $a' \ge a_H^4$. Since $\tilde{a} \ge a''$, we have that $\frac{\Pi_{HL}(a_H^4, \tilde{a})}{1-\delta} \ge \frac{\Pi_{HL}(a', a'')}{1-\delta}$. Since the ability difference between an H-player and an L-player is wide enough,

$$\frac{\delta\pi}{1-\delta+\delta\pi}\Pi_{HH}(a_H^1, a_H^1) \geq \Pi_{HL}(a_H^4, \tilde{a})$$
$$\Rightarrow V_H^S \geq \frac{\Pi_{HL}(a_H^4, \tilde{a})}{1-\delta} \geq \frac{\Pi_{HL}(a', a'')}{1-\delta} \tag{22}$$

where the second inequality follows from (20). Since $a', a'' \in [0, \tilde{a}]$ are arbitrary, the fixed point satisfies the condition 9 in Definition 1. This completes the proof. ▪

3. Feasibility of a steady state

We have studied the state in which π and θ are constant. Such a state is called a steady state. Here, we check the feasibility of the steady state.

Let π_h', π_l' be the proportion of H-players and the proportion of L-players, respectively, who are cooperating with the same type players with respect to the total players. Let π_h^t, π_l^t be the proportion of H-players and the proportion of L-players, respectively, in the pool of unmatched players at stage '*t*.' Then, $\pi_h^t(1-\pi_h'-\pi_l')$ denotes the proportion of unmatched H-players with respect to the total population at stage '*t*' and $(\pi_h^t-(\pi_h^t)^2)(1-\pi_h'-\pi_l')+\theta\pi_h'$ denotes the proportion of unmatched H-players with respect to the total population at stage '$t+1$.' Likewise, $\pi_l^t(1-\pi_h'-\pi_l')$ denotes the proportion of unmatched L-players with respect to the total population at stage '*t*' and $(\pi_l^t-(\pi_l^t)^2)(1-\pi_h'-\pi_l')+\theta\pi_l^*$ denotes the proportion of unmatched L-players with respect to the total population at stage '$t+1$.'

If

$$(\pi_h^t)^2(1-\pi_h'-\pi_l') = \theta\pi_h'$$

and

$$(\pi_l^t)^2(1-\pi_h'-\pi_l') = \theta\pi_l', \tag{24}$$

then they satisfy the sufficient condition for the steady state.

Claim 1 *For any given* $\pi_h^t, \theta \in (0,1)$*, there exist* $\pi_h', \pi_l' \in (0,1)$ *subject to equations (23) and (24).*

Proof. Fix π_h^t and θ. Suppose there exist π_h' and π_l' subject to (23) and (24). Then from (23) and (24),

$$\pi_h' = \pi_l'\frac{(\pi_h^t)^2}{(\pi_l^t)^2}. \tag{25}$$

From equations (23), (24) and (25), we can get

$$\pi_h' = \frac{(\pi_h^t)^2}{\theta + (\pi_h^t)^2 + (\pi_l^t)^2}$$

$$\pi_l' = \frac{(\pi_l^t)^2}{\theta + (\pi_h^t)^2 + (\pi_l^t)^2}.$$

Obviously, we have $\pi_h', \pi_l' \in (0,1)$. Therefore, for any given $\pi_h', \theta \in (0,1)$, if we choose $\frac{(\pi_h^t)^2}{\theta+(\pi_h^t)^2+(\pi_l^t)^2}$ and $\frac{(\pi_l^t)^2}{\theta+(\pi_h^t)^2+(\pi_l^t)^2}$ as π_h' and π_l', respectively, then they satisfy equations (23) and (24) above.

References

Boone, C., B. Brabander, M. Carree, G. Jong, W. Olffen and A. Witteloostuijn (2002), "Locus of Control and Learning to Cooperate in a Prisoner's Dilemma Game," *Personality and Individual Differences*, 32, 929-946.

Bose, G. (1996), "Bargaining Economies with Patient and Impatient Agents: Equilibria and Intermediation," *Games and Economic Behavior*, 14, 149-172.

Brosig, J. (2002), "Identifying Cooperative Behavior: Some Experimental Results in a Prisoner's Dilemma Game," *Journal of Economic Behavior and Organization*, 47, 275-290.

Datta, S. (1996), "Building Trust," STICERD - Theoretical Economics Paper Series 305.

Eeckhout, J. (2006), "Minorities and Endogenous Segregation," *Review of Economic Studies*, 73, 31-53.

Ellison, G. (1993), "Learning, Local Interaction, and Cooperation," *Econometrica*, 61, 1047-1071.

________ (1994), "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," *Review of Economic Studies*, 61, 567-588.

Foster, D. and P. Young (1990), "Stochastic Evolutionary Game Dynamics," *Theoretical Population Biology*, 38, 219-232.

____________________ (1991), "Cooperation in the Short and in the Long Run," *Games and Economic Behavior*, 3, 145-156.

Fudenberg, D. and E. Maskin (1986), "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54, 533-554.

Fujiwara-Greve, T. (2002), "On Voluntary and Repeatable Partnerships under No Information Flow," Proceedings of the 2002 North American Summer Meetings of the Econometric Society, http://www.dklevine.com/proceedings/game-theory.htm.

Fujiwara-Greve, T. and M. Okuno-Fujiwara (2009), "Voluntarily Separable Repeated Prisoner's Dilemma," *Review of Economic Studies*, 76, 993-1021.

Furusawa, T. and T. Kawakami (2008), "Gradual Cooperation in the Existence of Outside Options," *Journal of Economic Behavior & Organization*, 68, 378-389.

Ghosh, P. and D. Ray (1996), "Cooperation in Community Interaction without Information flows," *Review of Economic Studies*, 63, 491-519.

Juang, W. and H. Sabourian (2012), "A Folk Theorem for Stochastic Evolutionary Game Theory," Mimeo.

Kandori, M. (1992), "Social Norms and Community Enforcement," *Review of Economic Studies*, 59, 63-80.

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

Kranton, R. (1996), "The Formation of Cooperative Relationships," *Journal of Law, Economics, and Organization*, 12, 214-233.

Outkin, A. (2003), "Cooperation and Local Interactions in the Prisoners' Dilemma Game," *Journal of Economic Behavior and Organization*, 52, 481-503.

Rauch, J. and J. Watson (2003), "Starting Small in an Unfamiliar Environment," *International Journal of Industrial Organization*, 21, 1021-1042.

Robson, A. and F. Vega-Redondo (1996), "Efficient Equilibrium Selection in Evolutionary Games with Random Matching," *Journal of Economic Theory*, 70, 65-92.

Watson, J. (1999), "Starting Small and Renegotiation," *Journal of Economic Theory*, 85, 52-90.

________ (2002), "Starting Small and Commitment," *Games and Economic Behavior*, 38, 176-199.

Yang, C., C. Yue and I. Yu (2007), "The Rise of Cooperation in Correlated Matching Prisoners Dilemma: An Experiment," *Experimental Economics*, 10, 3-20.