THE KOREAN ECONOMIC REVIEW
Volume 25, Number 2, Winter 2009

LOCAL INTERACTION, ALTRUISM AND THE EVOLUTION OF NETWORKS*

TACKSEUNG JUN** · JEONG-YOO KIM***

We consider evolving networks on which players interact locally. In an evolutionary environment, players can increase their fitness either by changing their behavior or by changing their neighbors. We propose a static solution concept, what we call Stable Network Configuration (SNC). Roughly speaking, it requires that no player in the population distribution will change his type by imitating his best-performing neighbor, nor change his neighbor by rewiring his links. We characterize the necessary and sufficient conditions for the symmetric SNC in a network formation situation associated with the Prisoner's Dilemma game. Unlike the result by Eshel et al. (1998) that was obtained in a fixed circular network, all altruists and all egoists fare equally well and all altruists have links with some or all egoists in the symmetric SNC.

JEL Classification: D85
Keywords: Altruistic Behavior, Best-Accommodating Strategy, Local Interaction, Network Formation, Stable Network Configuration

---

Received for publication: Aug. 11, 2009. Revision accepted: Dec. 11, 2009.

* This research was begun when the second author visited POSTECH in 2004. We are very much grateful to seminar participants at POSTECH. This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2008-321-B00031).

** Department of Economics, Kyung Hee University, 1 Hoegidong, Dongdaemunku, Seoul 130-701, Korea, Tel: +82-2-961-0964, E-mail: tj32k@khu.ac.kr

*** Corresponding author: Department of Economics, Kyung Hee University, 1 Hoegidong, Dongdaemunku, Seoul 130-701, Korea, Tel: +82-2-961-0986, E-mail: jyookim@khu.ac.kr

I. INTRODUCTION

How can altruistic behavior survive in the long run in an evolutionary environment? This is quite puzzling because selfish behavior, by definition, maximizes an individual's utility or fitness and thus only *homo economicus* appears to be able to survive in the long run.

Eshel (1972) first recognized the possibility that altruistic behavior can survive in an evolutionary environment if agents do not interact with the whole population, but with only part of it. This idea has been recently followed by many authors, for example, Bergstrom and Stark (1993), Nowak and May (1992, 1993), and Eshel, Samuelson and Shaked (1998).[1] This line of literature shows that altruistic behavior can survive in the long run in an environment in which agents interact only with their neighbors in a given network.[2]

In reality, however, networks do not remain fixed. Networks keep evolving either endogenously by decisions of individuals or exogenously. In this paper, we consider networks not as given, but evolving by endogenous choices of players. The rationale for this consideration is as follows. If a certain type of behavior of a player yields low fitness in a situation of local interaction, his fitness may be increased either by changing his behavior or by changing his neighbors. This leads us to take into account the possibility that a player changes his neighbors by changing the network link yielding low fitness, rather than trying to optimize behavior within the existing network. If a player believes that other players using the strategy that would yielding him the highest payoff, what we call best-accommodating strategy, are highly likely to exist elsewhere, it is optimal to sever his link with a poorly

[1] While Nowak and May (1992) showed the long-run survival of altruistic behavior by simulations, Eshel *et al.* proved it analytically.

[2] There have been alternative approaches to the evolution of cooperation. Axelrod (1984) showed how cooperation can be evolutionarily advantageous by computer modeling, more specifically, that the tit-for-tat strategy gets the highest scores in tournaments. Note that the tit-for-tat strategy implies conditional cooperation which is distinguished from altruism. Frank (1988) stressed the important roles of emotions in decision making and personal interactions, and argued that people help others because there is emotional rewards to helping those who deserve the aid. Güth & Yaari (1992) and Güth (1995) developed the indirect evolutionary approach in which preferences are treated as endogenous to an evolutionary process, while actions are still determined by Nash equilibrium.

accommodating neighbor to rewire it to another player.

The objective of this paper is two-fold. First, we are interested in what network will emerge as a result of the evolutionary process when players interact only locally. Second, we are also interested in how the possibility of rewiring links can affect the sustainability of altruistic behavior.

For these purposes, we provide a network formation model in which a linked pair of players interact by playing a two-person game. The equilibrium concept in such a situation has to specify both the network structure and the distribution of behaviors ("types") in the network. We call this pair a network configuration. We propose a solution concept, what we call "Stable Network Configuration (SNC)." Roughly speaking, SNC requires (i) that no player in the population distribution will change his type by imitating his neighbor's and (ii) that no player will change the network structure by rewiring his links.

There has been growing literature on network formation since Jackson and Wolinsky (1996).[3] Most of the literature, however, abstracts from the details of the interaction following network formation and assigns an exogenous value as a direct benefit from a link. In other words, the literature has ignored the possible conflicts between neighbors that can follow the network formation. There are some exceptions, though. One exception is Berninghaus *et al.* (2004). They consider a network formation model which is followed by neighboring players playing 2×2 normal form game. However, they do not consider an evolutionary approach, but only a sfatic noncooperative game with two neighboring players.

While working on this paper, we found a paper which is very similar in spirit. Hanaki *et al.* (2007) also considers the possibility of changes in the network structure as well as changes in behavior by using simulations. So, the simulations predicted the population's long-run state but the authors did not consider an analytic equilibrium concept similar to the stable network configuration. In addition unlike Hanaki *et al.* (2007) our model, gives players an exclusive choice between changing their behavior *or* changing their neighbors after their performance in the previous period is

[3] To name a few, see Bala and Goyal (2000a, 2000b), Dutta and Mutuswami (1997), and Jackson and Watts (2002).

realized. In other words, players can not change both their and their neighbors at the same time in our model.

We provide a necessary condition for the symmetric SNC in general games, and characterize the necessary and sufficient conditions in the prisoner's dilemma game. In a symmetric SNC, all players must fare equally well regardless of their type, because otherwise one type would find it in its interest to imitate another type. For the prisoner's dilemma game, this leads to the result that an altruist has more egoistic neighbors (friends) than an egoist does. Also, the result of Eshel *et al.* (1998) that altruists will cluster among themselves is modified in our model of flexible networks, since there is an SNC in which all egoists are linked with all altruists. In the prisoner's dilemma game, altruistic behavior is the best-accommodating strategy with an egoist, which makes all egoists willing to sever their link with an egoist (if any) and to rewire it to an altruist.

The paper is organized as follows. Section 2 contains informal discussion for our motivation. In Section 3, we set up a general model of network formation with local interaction and propose our solution concept, Stable Network Configuration. In Section 4, we illustrate our concept in the Prisoner's Dilemma game. Concluding remarks follow in Section 5.

II. MOTIVATION

To motivate our approach, we will follow the model of local interaction by Eshel *et al.* (1998). There are n players arrayed on a circle. Each player is either an Altruist (A) or an Egoist (E). An "Altruist" provides one unit of public good at the net cost $C < 1/2$ so that his neighbors can enjoy the benefit equal to 1. An "Egoist" provides no public goods and bears no costs.

Initially, each player's type (A or E) is determined. In period one, all players choose their strategy (type). At the end of the period, the payoffs are realized and each player observes the payoffs of his own and his neighbors. In period two, each player learns by imitating the action of his neighbor earning the highest average payoff.[4] If the player himself earns

[4] This is called the "imitation dynamics."

the highest payoff, he retains his own action. This procedure is continued infinitely.

A distribution of types is called stationary if, no player changes his type in the next period when the distribution is given,.[5] Eshel *et al.* identify two trivial stationary distributions in which all are Altruists and all are Egoists, and one nontrivial stationary distribution with Egoist strings of length two and Altruist strings of length three or longer, something like $(A, A, A, E, E, A, A, A, A, \cdots)$. In this case, an A in the interior of A strings can survive, and A on the boundary of the string, i.e., between A and E, can survive if an A neighbor has another A neighbor and an E neighbor has an E neighbor. For the payoff of E is 1 and the average payoff of A is $\frac{1-C+2-C}{2} = \frac{3-2C}{2} > 1$ if $C < \frac{1}{2}$. Hence, this distribution is stationary because no player will change his type when the network structure is invariant.

In this story, a player learns how to behave from his total payoff. However, in fact, he can learn more by interacting with each neighbor. Note that both of A and E in the string of (A, E, E) know that they could fare better if their E-neighbor were an A. Thus, if they know the distribution of types, they would tend to sever their link with an E and to form a new link with another (randomly chosen), if the proportion of A's is high in the population. This implies that the nontrivial stationary distribution is not stable, when network changes are allowed, in the sense that both A and E tend to cut their links to E. Thus, we consider the possibility of network reformation.

III. MODEL AND MAIN CONCEPTS

There is a population of players from a set N with $|N| = n$. A network L on N is defined by a set of pairs of distinct players. If $ij \in L$ for a pair of (i, j), we say that player i and player j are *adjacent* or they are neighbors. The *degree* of player i, which will be denoted by $deg(i)$, is defined by the number of his neighbors. The

[5] This is a stronger solution concept than the absorbing set of distributions by Eshel *et al.* (1998) in the sense that the former does not admit cycling among more than one distribution in an absorbing set.

network is *regular* if $deg(i) = deg(j)$ for all $i, j \in N$, in which case all players have the same number of neighbors. The regular network of degree $n-1$ is called a *complete* network. A *path* from i to j in L is defined by a sequence of distinct players $(i, i_1, i_2, \cdots, i_k, j)$ such that $ii_1, i_1i_2, \cdots, i_kj \in L$. We say that a network L is connected if there is a path between any pair of (i, j) for $i, j \in N$.

Players interact only locally, in other words, a pair of players (i, j) who are neighbors play a two-person (normal form) game $\Gamma = \langle S, \Pi \rangle$ where S is the set of strategies available to each player and $\Pi = (\pi_i, \pi_j)$ is payoff functions of the players. Note that we assume that each player has the same strategy set and the payoff function is symmetric across players, i.e., $\pi_i(s_i, s_j) = \pi_j(s_j, s_i)$ for all $i, j \in N$ and for all $s_i, s_j \in S$. A strategy of player j, s, is called the best-accommodating strategy to s_i if $s = \arg\max_{s_j \in S} \pi_i(s_i, s_j)$.

We will call $\langle N, \Gamma, L \rangle$ a game on a network. A game on a network describes a situation in which player i and player j play Γ if and only if $ij \in L$ for $i, j \in N$.

We implicitly assume that the rationality of players is bounded in the sense that they initially choose their strategy dictated by their genes (which will be called their type), but we also assume that they can learn strategies yielding higher payoffs from their neighbors'.

The network structure can be changed by players' choices. Player i can sever his link $ij \in L$ unilaterally and rewire it to a randomly selected player k such that $ik \notin L$. Imagine that players have an electronic-directory with fixed memory capacity and that they can replace someone's number by another. Linking refers to information about interacting with another, for instance, making a call, sending an email or visiting someone's website and it usually does not require the consent of the other player. Also, we assume that the cost of rewiring a link is negligible.[6] Combining this assumption with the limited capacity, severing implies rewiring. Unilateral rewiring and negligible rewiring costs are inessential to our analysis, but we use them for simplicity.[7] By

[6] We make this assumption of zero cost, not because we consider the linking cost as not important, but because we have in mind the situation in which the cost is much lower than the benefit from the interaction.

[7] Our definition (Definition 1) can be easily adapted to the assumptions of positive linking cost

incorporating the network-forming decisions of players into $\langle N,\Gamma\rangle$, we can define a network formation associated with $\langle N,\Gamma\rangle$, denoted by $F(N,\Gamma)$.

Definition 1 *A pair of network and n-tuple strategy profile (L,**s**) is called network configuration.*

A network configuration (L,$\boldsymbol{s}$) tells us who are linked with each other and what type each player is. We say that a network configuration (L,$\boldsymbol{s}$) is connected if and only if L is connected. Also, we will call a network configuration (L,$\boldsymbol{s}$) *homogeneous* if $s_i = s$ for all $i \in N$ in L; otherwise, it will be called *heterogeneous*. Throughout the paper, we will be concentrated only on connected heterogenous network configurations. Let the degree of i with respect to type s be $deg(i;s) = \{k \in N_s \mid ik \in L\}$ where $N_s = \{k \in N \mid s_k = s\}$. We define a network configuration to be symmetric if for all $s \in S$, $deg(i;s) = deg(j;s)$ for every $i, j \in N$ such that $s_i = s_j$.

If players play $\boldsymbol{s}$ in network L, it generates the total payoff of each player. If we denote it by $(\psi_i(\boldsymbol{s},L))_{i\in N}$, we have

$$\psi_i(\boldsymbol{s},L) = \sum_{j\in L_i}\pi_i(s_i,s_j), \tag{1}$$

where $L_i = \{j \in N \mid ij \in L\}$. Also, we can define the average payoff of s-type neighbors to player i by

$$\bar{\psi}_{i,s}(\boldsymbol{s},L) = \frac{\sum_{j\in L_{i,s}}\psi_j(\boldsymbol{s},L)}{|L_{i,s}|}, \tag{2}$$

where $L_{i,s} = \{j \in L_i \cup i \mid s_j = s\}$. We assume that the information of each player is not global but local. That is, each player only knows whom he links with and what type his neighbors are, but he does not know what

---

and requiring mutual consent for a new link, although its application to a specific game as in Section 4 will involve more complicated computations.

type each player in the population is, although we assume that he knows the distribution of players. Thus, esch player can compute the average payoff of s -type neighbors.

Now, we propose a solution concept for $F(N,\Gamma)$ which we will call Stable Network Configuration (SNC).

Definition 2 *A network configuration (L^*,$\boldsymbol{s}^*$) is stable if and only if (i) $s_i^* \in BP_i(L^*)$, for all $i \in N$ where $BP_i(L^*) = \arg\max_{s \in S_i}\{\overline{\Psi}_{i,s}(\boldsymbol{s}^*, L^*)\}$ and $S_i = \{s_j \in S \mid j \in L_i^* \cup i\}$, and (ii) for all $i \in N$, $\psi_i(\boldsymbol{s}^*; L^* - ij + \sigma(i)) \le \psi_i(\boldsymbol{s}^*; L^*)$ for $j \in N$ such that $ij \in L$, where $\sigma(i) = ik$ for a random player $k \in N \setminus L_i \cup \{i\}$.*[8]

In words, condition (i) requires that given network L^*, no player changes his strategy in a way to imitate the strategy of his neighbor's type obtaining the highest average payoff. Condition (ii) implies that given the type distribution $\boldsymbol{s}^*$, no player rewires some or all of his links to others.[9] Following the spirit of Dunbar's law, we permit transformations of a network only by rewiring.[10]

Although SNC is a static solution concept, the following dynamics lies behind it. In the first period, players begin to interact with their neighbors by acting according to their tyoe without knowing the strategies of their neighbors. At the end of the period, they observe their own payoffs (fitness) and the total payoff of their neighbors as well as the strategies that each neighbor chose. In addition, they observe the type distribution in the period. In the next period, each player may imitate his neighbor's strategy yielding the highest average fitness, or change his neighbor(s) by rewiring some or all of his links, or retain his strategy and links. The choice of changing neighbors is made by his expectation of his average fitness. The expectation is myopic on two grounds. First, it is based on his knowledge of the strategy distribution, believing that the distribution in the next period will be the same as the current distribution. Second, they

---

[8] Strictly speaking, $\psi_i(\boldsymbol{s}^*; L^* - ij + \sigma(i))$ is the expected total payoff of player i.

[9] Under condition (ii) requiring no rewiring of one link, no player clearly rewires more than one link, since $\Delta_{jk}(i) \le 0$ for all $j,k \ne i$ implies $\sum_{j \ne k} \Delta_{jk}(i) \le 0$ where $\Delta_{jk}(i)$ is player i's gain from rewiring ($ij \to ik$).

[10] Dunbar's Law says that players in a social network do not maintain more than a certain limited number of links. See Dunbar (1993).

do not take account of the possibility that the network structure will be changed in the future. This game is repeated infinitely.

The criticism based on the myopia can be, in fact, applied to most of static solution concepts including pairwise stability by Jackson and Wolinsky, unless the concept considers all possible credible deviations with perfect foresightedness. Our SNC only considers the deviation of rewiring which does not require the consent from the other. If the other player does not want the new link, the link will be severed soon. So, the idea of mutual consent is already embedded in the definition, but not completely. That is, our solution concept is too strong in the sense that we allow incredible deviations as well, that is, we require a network configuration to be stable to any rewiring which might be severed by the other player immediately. However, it also has the advantage of paring down the set of stable network configurations.

Our solution concept does not rely on players' computing ability when they choose their strategies, but does rely on it when they decide whether to change their neighbors. Inarguably, however, their rationality is still bounded in the sense that they make calculations under the naive assumption that others will not change their strategies. Specifically, if a player, say i, finds his link with the current neighbor j yield him low fitness and expects many nonneighbors to use best-accommodating strategy to s_i, he would change partners. Note that SNC does not allow a player to change both his strategy and his neighbors simultaneously. This can be justified as follows. Once a player changes his neighbors, he cannot imitate his new neighbors immediately, since he had no chance to observe their payoffs. It also makes little sense to change neighbors after learning a new strategy, because the learned strategy is likely to perform well only in his previous local neighborhood.

It is worthwhile to compare our model with that of Hanaki *et al.* (2007). Since their model generates some dynamic properties, but these properties rely on many *ad hoc* assumptions. For example, they assume that players choose either to break a link with an existing neighbor or generate a link with a new player, and that if they fail, they try the other option. However, it is not clear why a player who failed to generate a link with some player does not try to create a link with some other player. Also, they assume that interaction with a neighbor incurs some cost, which provides a player the incentive to sever a link. If the interaction cost is zero as in our model, players never sever a link unless playing a normal-form game gives a negative payoff. In our model, a player may terminate his relation (even if the relation yields a positive payoff) in order to interact with a better partner. They also use parameters to represent the speed of learning and forming networks. In other words, not all players can learn the best-performing strategy and not all players can change their links. However, since our model is static, the analysis is robust to such specific assumptions on dynamics. No matter what the dynamics is, an unstable network configuration cannot be sustained, because at some point of time, it will be changed either by learning or by rewiring.

To establish our propositions, the following lemma will be useful.

Lemma 1 *If* $(L,\boldsymbol{s})$ *is symmetric and* $ij \in L$, $\bar{\psi}_{i,A}(\boldsymbol{s},L) = \bar{\psi}_{j,B}(\boldsymbol{s},L)$ *implies that* $\psi_i(\boldsymbol{s},L) = \psi_j(\boldsymbol{s},L)$ *for* $A,B \in S$, *unless the neighbor type ratios of* i *and* j *are the same.*

Proof. See the appendix.

The propositions below give necessary conditions for SNC in the network formation situation associated with Γ.

Proposition 1 *If a symmetric network configuration* $(L^*,\boldsymbol{s}^*)$ *is stable in* $F(N,\Gamma)$, $\psi_i(\boldsymbol{s}^*,L^*) = \psi(\boldsymbol{s}^*,L^*)$ *for all* $i \in N$.

Proof. See the appendix.

This proposition says that the payoffs should be the same not only across the players of the same type but also across types in a symmetric, stable network configuration. If this clid not hold, then one type would imitate on other.

Proposition 2 *Let* Γ *be a* 2×2 *game. If a network configuration* $(L^*,\boldsymbol{s}^*)$ *is stable in* $F(N,\Gamma)$, *the following must hold: For any arbitrary* $i \in N$, *let* s' *be the best-accommodating strategy to* s_i. *If* $s_j \neq s'$ *for some* $j \in L_i^* \cup i$, *then* $ik \in L^*$ *for all* $k \in N$ *such that* $s_k = s'$.

Proof. See the appendix.

In words, if player i has a link with a player who does not use the best-accommodating strategy to s_i, he must be linked with every player using the best-accommodating strategy to s_i in SNC. Otherwise, he would sever his link with the bad neighbor.

IV. PRISONER'S DILEMMA

We devote this section to illustrate SNC. Consider the following 2×2 symmetric game Σ;

$$\begin{array}{ccc} & A & B \\ A & a,a & c,d \\ B & d,c & b,b \end{array}$$

The prisoner's dilemma game requires that $d>a>b>c$.[11] We assume that $c>0$. To represent the claim that the players get benefits from interaction. The two-person version of the public provision game described in Section 2 is a special case of Σ.[12] Thus, in this representation, A type can be interpreted as "Altruist" and B interpreted as "Egoist." Note that A is the best-accommodating strategy to both A and B, since $a>c$ and $d>b$. Proposition 3 provides the necessary and sufficient condition for symmetric SNC in $F(N,\Sigma)$.[13]

Proposition 3 *A symmetric network configuration (L^*,$\boldsymbol{s}^*$) is stable in* $F(N,\Sigma)$ *if and only if (i) for all* $i\in N_A$, $deg(i;A)=n_A-1$ *, (ii) for all* $j\in N_B$, $deg(j;A)=n_A$ *if* $deg(j;B)\neq 0$ *and* $deg(j;A)\leq n_A$ *if*

---

[11] For a prisoner's dilemma game, it is also required that $2a>c+d$, but it is not essential for our analysis.

[12] Specifically, $a=1-C+K$, $b=K$, $c=-C+K$ and $d=1+K$ where $K>0$ is other benefits that could be obtained from the interaction itself.

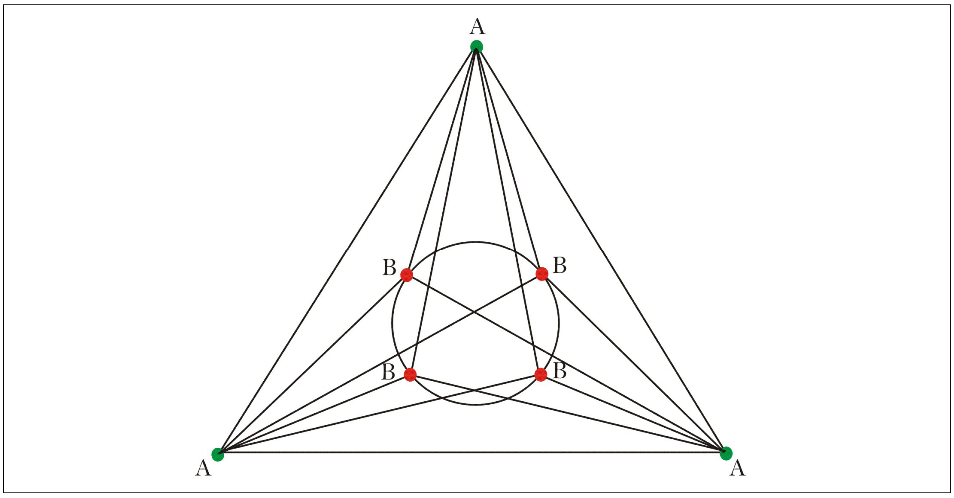
[13] That result depends heavily on the cardinality of the payoffs. A desirable property of a solution concept would be robustness to any affine transformation of payoffs with a positive re-scaling parameter. However, we think that this point is only relevant to a given (normal form) game. In the current model, players decide whether to interact with another player before they play the game. In this situation, it seems more realistic that the linking decision depends on the value from the interaction itself (K), that is, the cardinality of the payoffs.

$deg(j;B)=0$, *and (iii)* $\psi_i(\boldsymbol{s}^*, L^*) = \psi_j(\boldsymbol{s}^*, L^*)$ *for all* $i \in N_A$ *and* $j \in N_B$, *where* $n_s = |N_s|$, $s = A, B$. *The degrees,* $deg(j;A)(\neq n_A)$ *and* $deg(j;B)(\neq 0)$, *are determined from (iii).*

Proof. See the appendix.

We will call player $j \in N_B$ such that $deg(j;B)=0$ an isolated B-type player. This proposition says that the stable network configuration in the network formation game of a prisoner's dilemma game must have the following feature. First, all A types must be connected to one another if they have at least one B-type neighbor, because otherwise a player who is not linked to some A type would sever his link with B (since A is the best-accommodating strategy to A). Second, all non-isolated B types must be connected to all A types, because otherwise B types would want to sever their links with other B types (since A is the best-accommodating strategy to B). Third, the payoff of all B types must be the same as the payoff of A types, because otherwise one type would imitate the other. An example for the symmetric SNC is illustrated in Figure 1.

[Figure 1] An Example of the Symmetric Stable Network Configuration



To complete our discussion, we will check the existence of SNC by finding the equilibrium values for n_A, n_B, $deg(j;A)$ and $deg(j;B)$ for $j \in N_B$. First, consider the case that $deg(j;B) \equiv d_B \neq 0$.

Straightforward calculations lead to

$$\psi_i = (n_A - 1)a + n_B c,$$

$$\psi_j = n_A d + d_B b,$$

for $i \in N_A$ and $j \in N_B$. Thus, (n_A, d_B) must satisfy

$$nc = n_A(d + c - a) + d_B b + a. \tag{3}$$

Since $d > a$ and $b > c$, we can infer that $n_B \gg d_B$ to satisfy (3).[14] To illustrate, if we take $a = 3$, $b = 2$, $c = 1$, $d = 4$ and $n = 11$, equation (3) is reduced to $n_A + d_B = 4$. Thus, $(n_A, n_B, d_B) = (3,8,1), (2,9,2), (1,10,3)$ satisfies equation (3).

Now, let us turn to the case that $deg(j; B) = 0$. Let $deg(j; A) \equiv d_A$. Then, d_A must satisfy

$$(n_A - 1)a + n_B c = d_A d. \tag{4}$$

Thus, for the same parameter values, we have $(n_A, n_B, d_A) = (4,7,4)$.

One interesting implication in this stable network configuration is that the payoffs of an "Altruist" and an "Egoist" should be the same. This is possible, because Altruists have a larger number of neighbors than Egoists. Also, the above numerical illustration suggests that a large proportion of E types in the population can survive, contrary to the local interaction model without network reformation considered by Eshel *et al.* (1998). Their insight that clustering of A types is essential for their survival is still valid in this generalized model and even strengthened, in the sense that all A types must form a complete network among themselves for stability to occur. However, all A type players should be a neighbor of some or all E types in SNC, because E types want to interact with A types. Moreover, due to the additional requirement of payoff equality between the two types, A types have more neighbors of

---

[14] It is not possible that $\psi_i = \psi_j$ if $c \le 0$.

E types than E types do. This enables the population to consist of a relatively high proportion of E types.

V. CAVEAT AND CONCLUSION

In this paper, we showed that altruistic behavior can survive in an evolutionary environment of local interaction on a (endogenously) varying network by using the concept of "Stable Network Configuration."

The prediction of this paper is that the stability of a network configuration does not require the proportion of altruists to be very high. However, for the coexistence of altruists and egoists, altruists must have more egoistic friends than egoists. Egoists do not get along with other egoists very well, so they have only a limited number or even none of egoistic friends, while altruists become a friend of all the players in the population or have at least one egoistic friend as well as all other altruistic friends. On the other hand, altruists who get along with themselves form a complete network.

As we stressed in Section 3, the Stable Network Configuration is a static concept. The dynamic process of population evolution and network evolution will be addressed in a separate paper. Besides, the concept relies on several controversial assumptions. For example, we only consider rewiring deviations, do not require mutual consent for rewiring, assume zero rewiring cost etc. However, as we already argued in Section 3, most of the assumptions are not essential in the sense that our solution concept could be easily redefined by incorporating other possibilities. Then, characterizing the stable network configuration would be much more involving or trivial. For instance, if we consider a positive linking cost, the situation becomes the problem of network formation among heterogeneous players which is extremely complicated.[15] If we allow all kinds of transformation, especially adding a new link without severing an existing link, the complete network tends to emerge if all payoffs are positive as in the prisoner dilemma, since players will form links to all others. Then, only the inefficient strategy ("Egoist") would survive in the long run. This network configuration is indeed stable but homogeneous.

---

[15] See Galeotti, Goyal and Kamphorst (2006).

We admit that our analysis is just a first step in providing a static solution concept for the problem of coevolution of behavior and network. We look forward to the development of more robust and generalized solution concepts in the future.

Appendix

Proof of Lemma 1: In a symmetric network configuration, the average payoff of a player using the same strategy (A or B) must be the same. Let the payoff be ψ^A and ψ^B respectively. We have

$$\bar{\psi}_{i,A} = \frac{m_A\psi^A + m_B\psi^B}{m_A + m_B},$$

$$\bar{\psi}_{j,B} = \frac{m_A'\psi^A + m_B'\psi^B}{m_A' + m_B'},$$

where m_A and m_B are the numbers of i' neighbors whose type is A and B respectively, and m_A' and m_B' are similarly defined for j. Then, simple calculations lead to

$$\bar{\psi}_{i,A} = \bar{\psi}_{j,B} \Rightarrow (\psi^A - \psi^B)(m_A m_B' - m_A' m_B) = 0.$$

Thus, $\psi^A = \psi^B$ if $\frac{m_A}{m_B} \neq \frac{m_A'}{m_B'}$. ||

Proof of Proposition 1: Since we assume that L^* is connected, for any $i, j \in N$ such that $s_i \neq s_j$, there is a path $(i, i_1, \cdots, i_k, j)$. Assuming that $s_i \neq s_{i_1}$, $ii_1 \in L^*$ implies that $\bar{\psi}_{i,s_i}(\boldsymbol{s}^*, L^*) = \bar{\psi}_{i_1,s_{i_1}}(\boldsymbol{s}^*, L^*)$, because otherwise one type would imitate the other, violating the stability of $(L^*, \boldsymbol{s}^*)$. This in turn implies that $\psi_i(\boldsymbol{s}^*, L^*) = \psi_{i_1}(\boldsymbol{s}^*, L^*)$ by the symmetry of $(L^*, \boldsymbol{s}^*)$ and Lemma 1. Similarly, $\psi_{i_1}(\boldsymbol{s}^*, L^*) = \psi_{i_2}(\boldsymbol{s}^*, L^*) = \cdots = \psi_j(\boldsymbol{s}^*, L^*)$. Even if $s_i = s_{i_1}$, there must be (i_j, i_{j+1}) such that $s_j \neq s_{i_{j+1}}$ since $s_i \neq s_j$. Then, the same argument can be applied. Since there is a path between any pair of players with different types, the proof is completed. ||

Proof of Proposition 2: If $ik \notin L^*$ for some $k \in N_{s'}$, player i would sever the link with player j and rewire it to some $k \in N \setminus (L_i^* \cup i)$, since

$N\setminus(L_i^*\cup i)$ contains at least one s' type of player, k. This contradicts to the stability of $(L^*, \boldsymbol{s}^*)$. ||

Proof of Proposition 3: $(\Rightarrow)$ (i) Since $(L^*, \boldsymbol{s}^*)$ is heterogeneous and connected, there must be one B type of player, say j, such that $i_0 j \in L^*$ for some $i_0 \in N_A$. Thus, by Proposition 2, $i_0 i \in L^*$ for all $i \in N_A \setminus i_0$, i.e., $deg(i_0; A) = n_A - 1$. By symmetry of $(L^*, \boldsymbol{s}^*)$, $deg(i; A) = n_A - 1$ for all $i \in N_A$. (ii) If $deg(j; B) \neq 0$ for $j \in N_B$, $ij \in L^*$ for all $i \in N_A$ by Proposition 2, which implies that $deg(j; A) = n_A$. (iii) If $\psi_i(\boldsymbol{s}^*, L^*) \neq \psi_j(\boldsymbol{s}^*, L^*)$ for $i \in N_A$ and $j \in N_B$, then $\overline{\psi}_{i,A}(\boldsymbol{s}^*, L^*) \neq \overline{\psi}_{j,B}(\boldsymbol{s}^*, L^*)$ by symmetry. This implies that one type imitates the other, violating the stability of $(L^*, \boldsymbol{s}^*)$.

$(\Leftarrow)$ is trivial. ||

References

Axelrod, R. (1984), The Evolution of Cooperation, New York: Basic Books.

Bala, V. and S. Goyal (2000a), "A Noncooperative Model of Network Formation," *Econometrica*, Vol. 68, pp. 1181-1229.

Bala, V. and S. Goyal (2000b), "A Strategic Analysis of Network Reliability," *Review of Economic Design*, Vol. 5, pp. 205-228.

Bergstrom, T. and O. Stark (1993), "How Altruism can Prevail in an Evolutionary Environment," *American Economic Review (Papers and Proceedings)*, Vol. 83, pp. 149-155.

Berninghaus, S. and B. Vogt (2004), Network Formation in Symmetric 2×2 Games, Mimeo

Dunbar, R. I. M. (1993), "Coevolution of Neocortical Size, Group Size and Language in Humans," *Behavioral and Brain Sciences*, Vol. 16, pp. 681-735.

Dutta, B. and S. Mutuswami (1997), "Stable Networks," *Journal of Economic Theory*, Vol. 76, pp. 322-344.

Eshel, I. (1972), "On the Neighbor Effect and the Evolution of Altruistic Traits," *Theoretical Population Biology*, Vol. 3, pp. 258-77.

Eshel, I., L. Samuelson and A. Shaked (1998), "Altruists, Egoists, and Hooligans in a Local Interaction Model," *American Economic Review*, Vol. 88, pp. 157-179.

Frank, R. (1988), Passions within Reason, New York: Norton

Galeotti, A., S. Goyal and J. Kamphorst (2006), "Network Formation with Heterogeneous Players," *Games and Economic Behavior*, Vol. 54, pp. 353-372.

Güth, W. (1995), "An Evolutionary Approach to Explaining Cooperative Behavior by Reciprocal Incentives," *International Journal of Game Theory*, Vol. 24, pp. 323-344.

Güth, W. and M. Yaari (1992), An Evolutionary Approach to Explaining Reciprocal Behavior in a Simple Strategic Game, In: "Explaining Process and Change - Approaches to Evolutionary Economics" (ed. Witt, U.), University of Michigan Press, Ann Arbor, pp. 23-34.

Hanaki, N., A. Peterhansl, P. Dodds and D. Watts (2007), "Cooperation in Evolving Social Networks," *Management Science*, Vol. 53, pp. 1036-1050.

Jackson, M. and A. Watts (2002), "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, Vol. 106, 265-295.

Jackson, M. and A. Wolinsky (1996), "A Strategic Model of Social and

Economic Networks," *Journal of Economic Theory*, Vol. 71, pp. 44-74.

Nowak, M. and R. May (1992), "Evolutionary Games and Spatial Chaos," *Nature*, Vol. 359, pp. 826-829.

Nowak, M. and R. May (1993), "The Spatial Dilemmas of Evolution," *International Journal of Bifurcation and Chaos*, Vol. 3, pp. 35-78.