Games and Economic Behavior **34,** 331–341 (2001)
doi:10.1006/game.2000.0803, available online at http://www.idealibrary.com on IDEAL®

A Dynamic Model of Network Formation

Alison Watts*

Department of Economics, Box 1819, Station B, Vanderbilt University, Nashville, Tennessee 37235

Received March 31, 1999

Network structure plays a significant role in determining the outcome of many important economic relationships; therefore it is crucial to know which network configurations will arise. We analyze the process of network formation in a dynamic framework, where self-interested individuals can form and sever links. We determine which network structures the formation process will converge to. This information allows us to determine whether or not the formation process will converge to an efficient network structure. *Journal of Economic Literature* Classification Numbers: A14, C7, D20.

1. INTRODUCTION

Network structure plays a significant role in determining the outcome of many important economic relationships. There is a vast literature which examines how network structure affects economic outcomes. For example, Boorman (1975) and Montgomery (1991) examine the relationship between social network structure and labor market outcomes. Ellison and Fudenberg (1995) show that communication structure can influence a consumer's purchasing decisions. Political party networks can influence election results (see Vazquez-Brage and Garcia-Jurado, 1996). The organization of workers within a firm influences the firm's efficiency (see Keren and Levhari, 1983; Radner, 1993, and Bolton and Dewatripont, 1994). Hendricks *et al*. (1997) show that the structure of airline connections influences competition. Finally, in evolutionary game theory, Ellison (1993), Goyal and Janssen (1997), and Anderlini and Ianni (1996) show that network structure affects whether or not coordination occurs.

Since network structure affects economic outcomes, it is crucial to know which network configurations will arise. We analyze the process of network

*I thank an associate editor, an anonymous referee, Matt Jackson, Hervé Moulin, Anne van den Nouweland, John Weymark, and Giorgio Fagiolo for valuable comments and criticisms. E-mail: alison.watts@vanderbilt.edu.

0899-8256/01 $35.00

formation in a dynamic framework, where self-interested individuals can form and sever links. We determine which network structures the formation process will converge to. This information allows us to determine whether or not the formation process will converge to an efficient network structure. Specifically, we show that the formation process is path dependent, and thus the process often converges to an inefficient network structure. This conclusion contrasts with the results of Qin (1996) and Dutta *et al*. (1998) who find that an efficient network almost always forms.

In our model, there is a group of agents who are initially unconnected to each other. Over time, pairs of agents meet and decide whether or not to form or sever links with each other; a link can be severed unilaterally but agreement by both agents is needed to form a link. Agents are myopic, and thus decide to form or sever links if doing so increases their current payoff. An agent's payoff is determined as in Jackson and Wolinsky's (1996) connections model. (Agents receive a benefit from all direct and indirect connections, where the benefit of an indirect connection is smaller than that of a direct connection. Agents also must pay a cost of maintaining a direct connection, which can be thought of as time spent cultivating the relationship.) We show that if the benefit from maintaining an indirect link is greater than the net benefit from maintaining a direct link, then it is difficult for the efficient network to form. In fact, the efficient network only forms if the order in which the agents meet takes a particular pattern. Proposition 4 shows that as the number of agents increases it becomes less likely that the agents meet in the correct pattern, and thus less likely that the efficient network forms.

There are other papers which also address the idea of network formation. The endogenous formation of coalition structures is examined by Aumann and Myerson (1988), Qin (1996), Dutta *et al*. (1998), and Slikker and van den Nouweland (1997). The most important difference between their work and ours is that we assume that network formation is a dynamic process in which agents are free to sever a direct link if it is no longer beneficial. In contrast, Aumann and Myerson (1988) assume that once a link forms it cannot be severed, while Qin (1996), Dutta *et al*. (1998), and Slikker and van den Nouweland (1997) all consider one-shot games.

The three papers most closely related to the issues considered here are Jackson and Wolinsky (1996), Bala and Goyal (2000), and Jackson and Watts (1999). Jackson and Wolinsky (1996) examine a static model in which self-interested individuals can form and sever links. They determine which networks are stable and which networks are efficient.[1] Thus, they leave open the question of which stable networks will form. Here, we

[1] Dutta and Mutuswami (1997) also examine the tension between stability and efficiency, using an implementation approach.

extend the Jackson and Wolinsky connections model to a dynamic framework. Bala and Goyal (2000) simultaneously examine network formation in a dynamic setting. However, their approach differs significantly from ours in both modeling and results. Bala and Goyal restrict attention to models where links are formed unilaterally (one player does not need another player's permission to form a link with him) in a noncooperative game and focus on learning as a way to identify equilibria. Jackson and Watts (1999) also analyze the formation of networks in a dynamic framework. Jackson and Watts extend the current network formation model to a general network setting where players occasionally form or delete links by mistake; thus, stochastic stability is used as a way to identify limiting networks.

The remainder of the paper proceeds as follows. The model and static results are presented in Section 2, and the dynamic results are presented in Section 3. The conclusion and a discussion of what happens if agents are not myopic are presented in Section 4.

2. MODEL

Static Model and Results[2]

There are n agents, $N = \{1, 2, \ldots, n\}$, who are able to communicate with each other. We represent the communication structure between these agents as a network (graph), where a node represents a player, and a link between two nodes implies that two players are able to directly communicate with each other. Let g^N represent the complete graph, where every player is connected to every other player, and let $\{g \mid g \subseteq g^N\}$ represent the set of all possible graphs. If players i and j are directly linked in graph g, we write $ij \in g$. Henceforth, the phrase "unique network" means unique up to a renaming of the agents.

Each agent $i \in \{1, \ldots, n\}$ receives a payoff, $u_i(g)$, from network g. Specifically, agent i receives a payoff of $1 > \delta > 0$ for each direct link he has with another agent, and agent i pays a cost $c > 0$ of maintaining each direct link he has. Agent i can also be indirectly connected to agent $j \neq i$. Let $t(ij)$ represent the number of direct links in the shortest path between agents i and j. Then $\delta^{t(ij)}$ is the payoff agent i receives from being indirectly connected to agent j, where we adopt the convention that if there is no path between i and j, then $\delta^{t(ij)} = 0$. Since $\delta < 1$, agent i values closer connections more than distant connections. Thus, agent i's

[2] The static model (with the exception of the definition of stability) is identical to Jackson and Wolinsky's (1996) connections model.

payoff, $u_i(g)$, from network g, can be represented by

$$u_i(g) = \sum_{j \neq i} \delta^{t(ij)} - \sum_{j:\ ij \in g} c \tag{1}$$

A network, g, is *stable* if no player i wants to sever a direct link, and no two players, i and j, both want to form link ij and simultaneously sever any of their existing links. Thus, when forming a link agents are allowed to simultaneously sever any of their existing links.[3] Formally, g is stable if (a) $u_i(g) \geq u_i(g - ij)$ for all $ij \in g$ and (b) if $u_i(g + ij - ig - jg) > u_i(g)$, then $u_j(g + ij - ig - jg) < u_j(g)$ for all $ij \notin g$, where ig is defined as follows. If agent i is directly linked only to agents $\{k_l, \ldots, k_m\}$ in graph g, then ig is any subset (including the empty set) of $\{ik_l, \ldots, ik_m\}$.

Notice that the formation of a new link requires the approval of two agents. Thus, this definition of network stability differs from the definition of stability of a Nash equilibrium, which requires that no single agent prefers to deviate.

PROPOSITION 1. *For all N, a stable network exists. Further,*

(i) *if $c < \delta$ and $(\delta - c) > \delta^2$, then g^N is stable,*

(ii) *if $c \geq \delta$, then the empty network is stable,*

(iii) *if $c < \delta$ and $(\delta - c) \leq \delta^2$, then a star[4] network is stable.*

Jackson and Wolinsky (1996) prove Proposition 1 for the case in which agents can either form or sever links but cannot simultaneously form and sever links. However, their proof can easily be adapted to our context and is thus omitted. Note that in case (i), g^N is the unique stable network. However, in the remaining two cases, the stable networks are not usually unique (see Jackson and Wolinsky, 1996).

A network, g^*, is *efficient* (see Jackson and Wolinsky, 1996, and Bala and Goyal, 2000) if it maximizes the summation of each agent's utility, thus $g^* = \arg\max_g \sum_{i=1}^{n} u_i(g)$. The proof of the following proposition (on the existence of an efficient network) may be found in Jackson and Wolinsky (1996).

[3] This notion of stability is an extension of Jackson and Wolinsky's (1996) notion of pairwise stability where agents can either form or sever links but cannot simultaneously form and sever links. The current definition of stability is also used in the matching model section of Jackson and Watts (1999).

[4] A network is called a star if there is a central agent, and all links are between that central person and each other person.

PROPOSITION 2 (Jackson and Wolinsky, 1996). *For all N, a unique, efficient network exists. Further,*

(i) *if $(\delta - c) > \delta^2$, then g^N is the efficient network,*

(ii) *if $(\delta - c) < \delta^2$ and $c < \delta + \frac{(n-2)}{2}\delta^2$, then a star network is efficient,*

(iii) *if $(\delta - c) < \delta^2$ and $c > \delta + \frac{(n-2)}{2}\delta^2$, then the empty network is efficient.*

Dynamic Model

Initially the n players are unconnected. The players meet over time and have the opportunity to form links with each other. Time, T, is divided into periods and is modeled as a countable, infinite set, $T = \{1, 2, \ldots, t, \ldots\}$. Let g_t represent the network that exists at the end of period t and let each player i receive payoff $u_i(g_t)$ at the end of period t.

In each period, a link ij is randomly identified to be updated with uniform probability. We represent link ij being identified by $i{:}j$. If the link ij is already in g_{t-1}, then either player i or j can decide to sever the link. If $ij \notin g_{t-1}$, then players i and j can form link ij and simultaneously sever any of their other links if both players agree. Each player is myopic, and so a player decides whether or not to sever a link or form a link (with corresponding severances), based on whether or not severing or forming a link will increase his period t payoff.

If after some time period t, no additional links are formed or broken, then the network formation process has reached a *stable state*. If the process reaches a stable state, the resulting network, by definition, must be a stable (static) network.

3. DYNAMIC NETWORK FORMATION RESULTS

Propositions 3 and 4 tell us what type of networks the formation process converges to. This information allows us to determine whether or not the formation process converges to an efficient network.

PROPOSITION 3. *If $(\delta - c) > \delta^2 > 0$, then every link forms (as soon as possible) and remains (no links are ever broken). If $(\delta - c) < 0$, then no links ever form.*

Proof. Assume $(\delta - c) > \delta^2 > 0$. Since $\delta < 1$, we know that $(\delta - c) > \delta^2 > \delta^3 > \ldots > \delta^{n-1}$. Thus, each agent prefers a direct link to any indirect link. Each period, two agents, say i and j, meet. If players i and j are not directly connected, then they will each gain at least $(\delta - c) - \delta^{t(ij)} > 0$

from forming a direct link, and so the connection will take place. (Agent i's payoff may exceed $(\delta - c) - \delta^{t(ij)}$, since forming a direct connection with agent j may decrease the number of links separating agent i from agent $k \neq j$.) Using the same reasoning as above, if an agent ever breaks a direct link, his payoff will strictly decrease. Therefore, no direct links are ever broken.

Assume $(\delta - c) < 0$ and that initially no agents are linked. In the first time period, two agents, say i and j, meet and have the opportunity to link. If such a link is formed, then each agent will receive a payoff of $(\delta - c) < 0$; since agents are myopic, they will refuse to link. Thus, no links are formed in the first time period. A similar analysis proves that no links are formed in later periods. Q.E.D.

Proposition 3 says that if $(\delta - c) > \delta^2 > 0$, then the network formation process always converges to g^N, which is the unique efficient network according to Proposition 2. This network is also the unique stable network. Therefore, if the formation process reaches a stable state, the network formed must be g^N.

If $(\delta - c) < 0$, then the empty network is always stable (see Proposition 1). However, the empty network is efficient only if $c > \delta + ((n-2)/2)\delta^2$ (see Proposition 2). Thus, the efficient network does not always form. If $c < \delta + ((n-2)/2)\delta^2$, then the star network is the unique efficient network. However, since $c > \delta$, this network is not stable (the center agent would like to break all links), and so the network formation process cannot converge to the star in this case.

If $(\delta - c) < 0$, then multiple stable networks may exist. In this case, the empty network is the most inefficient stable network. For example, if $n = 5$ and $(\delta^2 - \delta^3 - \delta^4) > (c - \delta) > 0$, then the circle network is stable. Each agent receives a strictly positive payoff in the circle network; therefore, the circle is more efficient than the empty network.

PROPOSITION 4. *Assume that* $0 < (\delta - c) < \delta^2$. *For* $3 < n < \infty$, *there is a positive probability*, $0 < p(star) < 1$, *that the formation process will converge to a star. However, as n increases*, $p(star)$ *decreases, and as n goes to infinity*, $p(star)$ *goes to* 0.

The following lemma is used in the proof of Proposition 4.

LEMMA 1. *Assume* $0 < (\delta - c) < \delta^2$. *If a direct link forms between agents i and j and a direct link forms between agents k and m (where agents i, j, k, and m are all distinct), then the star network will never form.*

Proof of Lemma 1. Assume that $0 < (\delta - c) < \delta^2$ and that the star does form. Order the agents so that agent 1 is the center of the star, agent 2 is the first agent to link with agent 1, agent 3 is the second, ..., and agent

n is the last agent to link with agent 1. We show that if the star forms, then every agent $i \neq 1$ meets agent 1 before he meets anyone else.

Assume, at time period t, agent 1 meets agent n and all agents $i \in \{2, \dots, n-1\}$ are already linked to agent 1. Assume agents 1 and n are so far not directly linked. Thus, in order for the star to form, agent 1 must link to agent n. But agent 1 will link to agent n only if agent n is not linked to anyone else. Assume, to the contrary, that agent n is linked to agent 2. If agent 1 links to agent n, agent 1's payoff will change by $(\delta - c) - \delta^2 < 0$ (regardless of whether of not agent n simultaneously severs his tie to agent 2). Therefore agent 1 will not link with agent n. In order for agent n to be unlinked in period t, agent n cannot have met anyone else previously, since a link between two unlinked agents will always form (recall that $\delta > c$), and such a link is never broken unless the two agents have each met someone else and have an indirect connection they like better.

Next consider time period $(t-1)$ in which agent $(n-1)$ joins the star. Again, agent $(n-1)$ must be unlinked to agents $\{2, \dots, n-2\}$, otherwise agent 1 will refuse to link with agent $(n-1)$. Also agent $(n-1)$ cannot be linked to agent n, since agent n must be unlinked in period t. This process can be repeated for all agents. Hence, all agents must meet agent 1 before they meet anyone else. Contradiction. Q.E.D.

Proof of Proposition 4. Lemma 1 states that if two distinct pairs of players get a chance to form a link, then a star cannot form. We show that the probability of this event happening goes to 1 as n becomes large. Fix any pair of players. The probability that a distinct pair of players will be picked to form a link next is $(n-2)(n-3)/n(n-1)$. This expression goes to 1 as n becomes large. Q.E.D.

Lemma 1 states that the dynamic process often does not converge to the star network. When it does not, the process will converge either to another stable network or to a cycle (a number of networks are repeatedly visited); see Jackson and Watts (1999). For certain values of δ and c, no cycles exist, and thus the dynamic process must converge to another stable network. For example, if c is large or if δ is close to 1, then a player only wants to add a link if it is to a player he is not already directly or indirectly connected to. Thus, the dynamic process will converge to a network which has only one (direct or indirect) path connecting every pair of players. For further discussion of cycles and conditions which eliminate cycles, see Jackson and Watts (1999).

Lemma 1 can be interpreted as follows. First, recall from Propositions 1 and 2 that if $0 < (\delta - c) < \delta^2$, then a star network is stable, but it is not

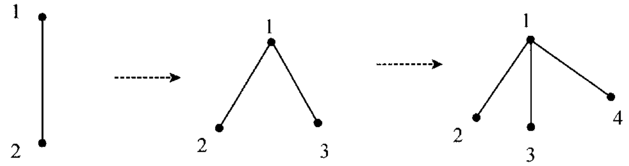

FIGURE 1

necessarily the only stable network. However, the star is the unique efficient network. Therefore, Lemma 1 says that if $0 < (\delta - c) < \delta^2$, then it is difficult for the unique efficient network to form. In fact, the only way for the star to form is if the agents meet in a particular pattern. There must exist an agent j who acts as the center of the star. Every agent $i \neq j$, must meet agent j before he meets any other agent. If, instead, agent k is the first agent player i meets, then players i and k will form a direct link (since $\delta > c$) and, by Lemma 1, a star will never form. These points are illustrated by the following example.

For $n = 4$, a star will form if the players meet in the order (1:2, 1:3, 1:4, 2:3, 2:4, 3:4), but not if the players meet in the order (1:2, 3:4, 1:3, 1:4, 2:3, 2:4). If the players meet in the order (1:2, 1:3, 1:4, 2:3, 2:4, 3:4), then every agent $i \neq 1$ meets agent 1 before he meets any other agent. Since $\delta > c$, every agent $i \neq 1$ will form a direct link with agent 1. Thus, a star forms in three periods, with agent 1 acting as the center (see Fig. 1.).

If the players meet in the order (1:2, 3:4, 1:3, 1:4, 2:3, 2:4), then the network formation process will converge to a circle if $(\delta - c) > \delta^3$, and the formation process will converge to a line if $(\delta - c) < \delta^3$. Next we briefly outline the formation process. Since $\delta > c$, we know that agents 1 and 2 will form a direct link in period 1, agents 3 and 4 will form a direct link in period 2, and agents 1 and 3 will form a direct link in period 3 (see Fig. 2). In period 4, agent 4 would like to delete his link with agent 3 and simultaneously form a link with agent 1; however, agent 1 will refuse to

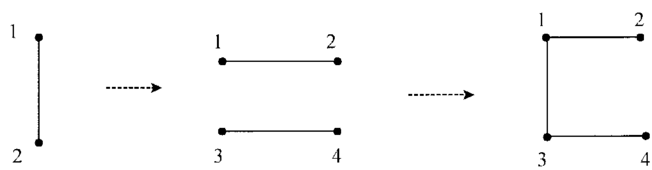

FIGURE 2

link with agent 4 since $\delta^2 > (\delta - c)$. Similarly, in period 5, agent 3 will refuse to link with agent 2. In period 6, agents 2 and 4 will agree to link only if $(\delta - c) > \delta^3$.

Proposition 4 states that as the number of players increases, it becomes less likely that the players will meet in the pattern needed for the star to form. Thus as n increases, the probability of a star forming decreases. For example, if $n = 3$, the probability that the star will form is $p(\text{star}) = 1$. For $n = 4$, the probability that the star forms is $p(\text{star}) = 0.27$, while for $n = 5$, $p(\text{star}) = 0.048$.

The intuition for why it rapidly becomes more difficult for the star to form can be gained by examining the $n = 4$ case. Assume that in period 1, players 1 and 2 meet; players 1 and 2 will form a link since $\delta > c$. Therefore, if a star forms, then we know from Lemma 1 that either agent 1 or agent 2 must act as the center. So, in period 2, the star continues to form as long as agents 3 and 4 do not meet each other. Assume, in period 2, that agents 1 and 3 meet. If the star forms, agent 1 must act as the center. Thus the star can form only if agents 1 and 4 meet before agents 3 and 4 or agents 2 and 4 meet. As the number of agents grows, it becomes less likely that the correct pairs of agents will meet each other early in the game, and thus it becomes less likely that the star forms.

4. CONCLUSION

We show that if agents are myopic and if the benefit from maintaining an indirect link of length two is greater than the net benefit from maintaining a direct link ($\delta^2 > \delta - c > 0$), then it is fairly difficult for the unique efficient network (the star) to form. In fact, the efficient network only forms if the order in which the agents meet takes a particular pattern. One area of future research would be to explore what happens if agents are instead forward looking. The following example gives intuition for what might happen in such a nonmyopic case.

First, consider a myopic four-player example where $\delta^2 > \delta - c > 0$. Suppose that agents have already formed the line graph where 1 and 2 are linked, 2 and 3 are linked, and 3 and 4 are linked. If agents 1 and 3 now have a chance to link, then agent 1 would like to simultaneously delete his link with 2 and link with 3. However, agent 3 will refuse such an offer since he prefers being in the middle of the line to being the center agent of the star. This example raises the question: will player 1 delete his link with agent 2 and wait for a chance to link with 3 in a model with foresight?

To answer this question, we first observe that even though the star is the unique efficient network, the payoff from being the center agent is $3\delta - 3c$, which is much smaller than the payoff from being a noncenter agent (which equals $(\delta - c) + 2\delta^2$). Thus, in a model with foresight, player 1 may delete his link with agent 2 and wait for a chance to link with agent 3. However, agent 3 would rather that someone else be the center of the star; thus, when 3 is offered a chance to link with 1, he has an incentive to refuse this link in the hope that agent 1 will relink with agent 2 and that agent 2 will then become the center of the star. However, agent 2 will also have incentive not to become the center of the star. Thus, it is unlikely that forward-looking behavior will increase the chances of the star forming.

REFERENCES

Anderlini, L., and Ianni, A. (1996). "Path Dependence and Learning from Neighbors," *Games Econom. Behav.* **13**, 141–177.

Aumann, R. J., and Myerson, R. B. (1988). "Endogenous Formation of Links between Players and of Coalitions: An Application of the Shapley Value" in *The Shapley Value* (A. Roth, Ed.). New York: Cambridge Univ. Press.

Bala, V., and Goyal, S. (2000). "A Non-Cooperative Model of Network Formation," forthcoming in *Econometrica*.

Bolton, P., and Dewatripont, M. (1994). "The Firm as a Communication Network," *Q. J. Econom.* **109**, 809–839.

Boorman, S. (1975). "A Combinatorial Optimization Model for Transmission of Job Information through Contact Networks," *Bell J. Econom.* **6**, 216–249.

Dutta, B., and Mutuswami, S. (1997). "Stable Networks," *J. Econom. Theory* **76**, 322–344.

Dutta, B., van den Nouweland, A., and Tijs, S. (1998). "Link Formation in Cooperative Situations," *Int. J. Game Theory* **27**, 245–256.

Ellison, G. (1993). "Learning, Local Interaction and Coordination," *Econometrica* **61**, 1047–1072.

Ellison, G., and Fudenberg, D. (1995). "Word-of-Mouth Communication and Social Learning," *Q. J. Econom.* **110**, 93–126.

Goyal, S., and Janssen, M. (1997). "Non-Exclusive Conventions and Social Coordination," *J. Econom. Theory* **77**, 34–57.

Hendricks, K., Piccione, M., and Tan, G. (1997). "Entry and Exit in Hub-Spoke Networks," *Rand J. Econom.* **28**, 291–303.

Jackson, M. O., and Watts, A. (1999). "The Evolution of Social and Economic Networks," Mimeo, Vanderbilt University.

Jackson, M. O., and Wolinsky, A. (1996). "A Strategic Model of Social and Economic Networks," *J. Econom. Theory* **71**, 44–74.

Keren, M., and Levhari, D. (1983). "The Internal Organization of the Firm and the Shape of Average Costs," *Bell J. Econom.* **14**, 474–486.

Montgomery, J. (1991). "Social Networks and Labor Market Outcomes," *Amer. Econom. Rev.* **81**, 1408–1418.

Qin, C. Z. (1996). "Endogenous Formation of Cooperation Structures," *J. Econom. Theory* **69**, 218–226.

Radner, R. (1993). "The Organization of Decentralized Information Processing," *Econometrica* **61**, 1109–1146.

Slikker, M., and van den Nouweland, A. (1997). "A One-Stage Model of Link Formation and Payoff Division," CentER Discussion Paper No. 9723.

Vazquez-Brage, M., and Garcia-Jurado, I. (1996). "The Owen Value Applied to Games with Graph-Restricted Communication," *Games Econom. Behav.* **12**, 42–53.