Friendship Networks

by

Jan K. Brueckner
Department of Economics
and
Institute of Government and Public Affairs
University of Illinois at Urbana-Champaign
1206 South Sixth St.
Champaign, IL 61820
e-mail: jbrueckn@uiuc.edu

March 2003

Abstract

This paper develops a model of social networks different from those presented in the recent literature. In contrast to existing models, the level of investment in link formation is a decision variable rather than being exogenous, and links form stochastically rather than deterministically, with the probability depending on the noncooperative investment choices of both parties. Since network structure is then stochastic rather than deterministic, the actual pattern of links cannot be specified, as in previous models, with the analysis focusing instead on which links are most likely to form. The analysis is couched in the context of friendship formation.

Friendship Networks

by

Jan K. Brueckner*

1. Introduction

Building upon a long tradition in sociology, economists have recently turned their attention to the analysis of social networks. Such networks play a crucial role in human interaction. They facilitate the dispersion of economic information, such as the existence of job openings at particular firms, and they provide transmission links for cultural information, such as style trends and opinions on the latest books and movies. In addition, interpersonal networks provide pure enjoyment through individual friendship bonds and broader social interaction at parties and other group gatherings.

Economic analysis of social networks begins with a set of assumptions on the process by which interpersonal links in a network are formed. Generally, link formation requires resource investment on the part of the individuals creating the link. Then, the analysis must specify the gains from network linkages, which include the benefits from direct connections to other individuals as well the gains from indirect connections via a sequence of links. With the costs and benefits of link formation specified, the next step is to analyze the equilibrium structure of the network, characterizing the pattern of links. The question is: who is connected to whom? The final step is to compare the equilibrium network with the one that is socially optimal.

In the recent economic literature, two distinct approaches to the analysis of social networks stand out. The first is developed by Jackson and Wolinsky (1996) and extended in Jackson (2001) and Jackson and Watts (2002a,b). In this model, both individuals must incur a cost in order for a link between them to be formed. If one individual withholds the required investment, the link is severed. This setup is referred to as a "non-directed" network. By contrast, Bala and Goyal (2000a) analyze a "directed" network, where an individual can create a link with another agent, enjoying the resulting benefits, by making a one-sided investment. For benefits to also flow in the opposite direction, the other individual must incur a corresponding cost.

While Dutta and Jackson (2000) provide further analysis of this model, Bala and Goyal (2000b) generalize the setup so that a one-sided investment facilitates a bi-directional flow of benefits.

While link formation differs under these two approaches, network benefits have a similar characterization. In one version of the Jackson-Wolinsky model, the benefits of linkages attenuate with distance, with the gain from association with an individual L links distant reduced by the factor δ^L, where $\delta \leq 1$. Under the Bala-Goyal approach, benefits depend simply on the total number of individuals accessed through direct or indirect linkages, an outcome that can be generated by setting $\delta = 1$ in the Jackson-Wolinsky model.

The analysis of Jackson and Wolinsky shows that the equilibrium network is empty if linkage costs are high, is a star network (with all individuals connected to a single central agent) when linkage costs are moderate, and is a complete network (with everyone linked to everyone else) when costs are low. In addition, they show that the equilibrium network may not be efficient. In Bala and Goyal's model, the efficient network is a wheel, and they show that the equilibrium converges to this form under a dynamic process. For an excellent overview of all of this analysis, as well as references to additional papers, see Dutta and Jackson (2002).[1]

The present paper adds to this emerging literature by proposing a different approach to social-network formation. As in the Jackson-Wolinsky model, formation of a link between two individuals requires two-sided investments in the present framework. But in contrast to their approach, where the required investments are exogenously specified and link formation is deterministic, the level of individual investment is a decision variable in the present model and link formation is stochastic. Thus, the probability that a link is formed between two individuals depends on the "effort" both agents devote to creating the link. These effort levels are chosen noncooperatively via Nash behavior.

Under this approach, the question of network structure takes a different form than in previous work. Since links are stochastic rather than deterministic, investigating the structure of a network involves asking which links are most likely to form, rather than characterizing the actual pattern of links.

Moreover, further assumptions on the technology of link formation militate against the emergence of some network patterns that are familiar from deterministic models. In particular, because of decreasing returns to effort in forming links and the increasing marginal costs of exerting it, individuals will spread their effort over many different potential partners. Thus, unless ex ante asymmetries are added to the model, effort in link formation will be exerted equally across all potential partners, ruling out constructs like the star network from deterministic models.

For the stochastic analog to such a network to emerge, with some links more likely to form than others, ex ante asymmetries must be present. For example, a personally magnetic individual may elicit high effort levels from potential partners. Links with this individual may then be more likely to form than links with non-attractive agents.

Another possible source of asymmetry relates to a further key feature of the model. In particular, the analysis assumes that, for a link to be created between two individuals, they must be "acquainted" ahead of time, a relationship that involves a low level of contact falling short of an actual linkage. Thus, if individuals have different sets of acquaintances within the universe of all agents, effort levels directed toward forming links may not be uniform. By contrast, if everyone is acquainted with everyone else, then all possible links are feasible and symmetric effort levels will emerge.

A final feature of the model concerns the benefits from direct vs. indirect linkages. As in the Jackson-Wolinsky model, indirect links are worth less than direct linkages. But, in contrast to their assumption of a smooth benefit decay as link distance increases, the present framework assumes that benefits are zero when more than two links are involved. Thus, in the symmetric case, an individual i receives a given benefit from a direct linkage to another agent j, and a smaller benefit from being connected indirectly to those individuals with direct links to agent j. However, i receives no benefit from those agents linked indirectly to j.

To motivate the analysis, it is helpful to think of the model as a portrayal of friendship networks. For two individuals to form a friendship, they must be acquainted ahead of time, perhaps as a result of working at the same firm. Each individual must then exert effort, which could involve inviting the other person to dinner at his house, arranging other types of social outings, or buying gifts on special occasions. Friendships form most easily when effort is reciprocated, although bonds may form and be maintained with unbalanced effort levels on the two sides. Effort creates "direct" friendships, and the combination of such links leads to "indirect" friendships. Concretely, a particular individual may invite all of his direct friends to a dinner party at his house, and through socializing with one another, these people enjoy indirect friendships. Because this type of dinner-party interaction is not as intimate as the one-on-one contact between direct friends, it yields lower benefits. For convenience, the model will be developed using this friendship terminology, recognizing that the analysis applies more generally.

Section 2 of the paper presents the general version of the model and then analyzes the fully symmetric case. In this case, each individual is acquainted with the entire universe of agents, and everyone is equally attractive as a friend. This section establishes an important efficiency result, which is natural given the structure of the model. In particular, because each individual's effort generates externalities, chosen effort levels are uniformly too low. One externality arises because, in choosing his effort levels, each individual ignores the reciprocal benefits enjoyed by his direct and indirect friends, and a second type of externality is identified in the analysis. Thus, people do not exert enough effort in forming and maintaining friendships.

Section 3 analyzes the effect on the equilibrium network structure of several types of asymmetries, focusing on how asymmetries affect effort levels and the probabilities of link formation. The analysis considers the case where a single individual (a "magnetic agent") offers greater friendship benefits than anyone else, and the case where one agent "knows everyone," being acquainted with the entire universe of agents (the remaining individuals are each acquainted with only a portion of the universe). The analysis is carried out under several alternate assumptions regarding the effect of effort levels on the friendship probability. Under one assumption, this probability depends on the sum of the two effort levels, while under an alternate specification, the minimum of the effort levels is what matters.

Section 4 recasts the model in two-period setting, where indirect friendships formed in the first period can be upgraded to direct friendships by further second-period effort. Section 5 offers conclusions.

2. The General Model

2.1. The setup

Let the universe of individuals be comprised of n agents. Moreover, let $a(i)$ denote the set of agents in this universe with whom individual i is acquainted, so that

$$a(i) \equiv \{j|\ i \text{ and } j \text{ are acquainted}\}. \tag{1}$$

Furthermore, let e_{ij} denote the effort expended by agent i in attempting to establish a friendship link with $j \in a(i)$. The probability that an ij friendship is formed is then given by

$$P(e_{ij}, e_{ji}). \tag{2}$$

The function P, which satisfies $0 \leq P < 1$, is increasing in both arguments, and the second partial derivatives $\partial^2 P/\partial e_{ij}^2$ and $\partial^2 P/\partial e_{ji}^2$ are both negative. In addition, P is a symmetric function, with $P(e_{ij}, e_{ji}) = P(e_{ji}, e_{ij})$. Note while effort levels are specific to individuals, the P function itself does not depend on the identities of individuals attempting to form a friendship link.

This probabilistic approach can be justified by imagining that, conditional on effort levels, friendship formation depends on the realization of a random error term. In particular, suppose that a friendship between i and j is established when $F(e_{ij}, e_{ji}) + \epsilon > 0$, where F is an increasing function and ϵ is an error term that is identically distributed across all potential pairs (i, j) of agents, as well as independent between different pairs. Then, the friendship probability is $\text{Prob}[\epsilon > -F(e_{ij}, e_{ji})] \equiv P(e_{ij}, e_{ji})$.

For agent i, the cost of the effort exerted to establish an ij friendship is $C(e_{ij})$, where C is increasing and strictly convex and where $C(0) = 0$. Consequently, the cost of i's effort across all possible friendship links is $\sum_{j \in a(i)} C(e_{ij})$. Several aspects of this cost formulation should be noted. First, effort is not chosen subject to any kind of resource constraint. Instead, the choice of effort is constrained only by the increasing cost of exerting it. While maintaining this orientiation, an alternate approach would be to define effort e in such a way that its cost is simply e itself (implying that C is the identity function). While this approach is sensible, it is unworkable under a key special case considered below, where the P function is assumed to depend on the sum of the effort levels, being written $P(e_{ij} + e_{ji})$ (see below for a fuller explanation).

A different alternate approach would be to write the total cost of effort as $C(\sum_{k \in a(i)} e_{ik})$. However, this approach does not reflect increasing costs at the individual link level. Intuitively, the current approach can be justified by imagining that effort comes in cardinal units, and that it is increasingly costly to generate extra units on a particular link as the effort level rises. For example, adding extra units of effort in cultivating a particular friend may require increasingly lavish dinner parties, implying $C'' > 0$.

Finally, let u_{ij} denote agent i's benefit from being a direct friend of j, and let $v_{ij} < u_{ij}$ denote i's benefit from indirect friendship with j. As explained above, indirect friendship means that i is a direct friend of an individual k who in turn is a direct friend of j. A crucial assumption in the analysis is that these friendship benefits are cumulative. In other words, individual i can be both a direct and an indirect friend of j, enjoying benefits from both associations. Concretely, i can enjoy direct-friendship benefits from inviting j to dinner at his own house, but he also derives enjoyment from seeing j at a dinner party at k's house, although the latter benefit is lower. Moreover, each time that i sees j at parties held by different hosts, indirect-friendship benefits arise.

An alternative approach would be to assume that the benefits of friendship are not cumulative, reflecting the view that benefits arise purely from knowing another individual, not from the amount of time spent together. Under this approach, once i establishes an indirect friendship with j through a third party k, another indirect-friendship link through agent l would add nothing to benefits. Moreover, a direct friendship would supercede indirect friendships, with i's total benefit from knowing j equal to u_{ij} regardless of the extent of indirect contacts between the two individuals. While some progress can be made in analyzing this alternate model, few results can be derived.

2.2. Equilibrium effort choices

With the above background, expected friendship benefits net of the costs of effort can be computed for individual i. The relevant expression is

$$B_i \;=\; \sum_{j\in a(i)} P(e_{ij}, e_{ji}) \left[u_{ij} \;+\; \sum_{h\in a(j),h\neq i} v_{ih} P(e_{jh}, e_{hj}) \right] \;-\; \sum_{j\in a(i)} C(e_{ij}). \tag{3}$$

The $u_{ij}P(e_{ij}, e_{ji})$ term in (3) gives agent i's expected benefits from formation of a direct friendship with an acquaintance j, and the last summation gives the cost of effort. The middle summation, when combined with $P(e_{ij}, e_{ji})$, captures the expected benefits from indirect friendships formed via agent j. To understand this term, note that $P(e_{ih}, e_{hi})P(e_{hj}, e_{jh})$ gives the probability that i forms an indirect friendship with h through agent j. For such a link to arise, i must be a direct friend of j and j must be a direct friend of h. Multiplying by v_{ih} thus gives expected benefits, and summing across all individuals that are acquaintances of j gives total expected benefits of indirect friendships formed via j.

Note that use of the multiplicative probability expression above to compute the probability of an indirect friendship follows from the previous independence assumption on the error term ϵ. This assumption means that agent h's formation of direct friendships with i and j represent independent events in a probabilitistic sense, so that their joint probability is given by the product expression.

Individual i chooses his effort levels taking the effort choices of others as parametric. The first-order condition for choice of e_{il}, where $l \in a(i)$, is given by

$$\frac{\partial B_i}{\partial e_{il}} \;=\; \frac{\partial P(e_{il}, e_{li})}{\partial e_{il}} \left[u_{il} \;+\; \sum_{h\in a(l),h\neq i} v_{ih} P(e_{lh}, e_{hl}) \right] \;-\; C'(e_{il}) \;=\; 0. \tag{4}$$

The portion of (4) involving u_{il} gives the increase in the expected benefit from a direct friendship with l as e_{il} rises, and the last term gives the marginal cost of the extra effort. An increase in e_{il} also raises the likelihood of all indirect friendships that pass through l by increasing the probability of an il link. The summation in (4), when multiplied by the P derivative, cumulates these incremental benefits across the feasible links through l. Thus, the first-order condition balances the gains from a greater likelihood of both direct and indirect friendships against the cost of additional effort. Finally, it is easily seen that the second-order condition for choice of e_{il} is satisfied given the maintained assumptions on the P and C functions.

Each individual chooses effort levels for the links involving his various acquaintances, satisfying (4) in each case. The overall equilibrium is then found by simultaneous solution of the resulting large collection of first-order conditions (note that each individual contributes his own set of equations). To gain further insight, it is useful to simplify matters by considering the fully symmetric case, where friendship benefits are uniform across individuals and where each person is acquainted with everyone else. In this case, $u_{ij} \equiv u$ and $v_{ij} \equiv v$ hold for all i and j. Since effort choices will be symmetric across individuals in this situation, all the e's in (4) are identical, and the equation can be used to solve for the common value. To write the equation in a compact form, let $P^1(e,e)$ denote the partial derivative of P with respect to its first argument. Then (4) reduces to

$$P^1(e,e)[u \ + \ (n-2)vP(e,e)] \ = \ C'(e). \tag{5}$$

Using this condition, several quick comparative calculations can be carried out. Letting

$$\Phi(e) \ \equiv \ [P^1(u \ + \ (n-2)vP) \ - \ C'], \tag{6}$$

where $\Phi'(e) < 0$, it follows that $\partial e/\partial n = -vPP^1/\Phi' > 0$, with the effects of u and v on e also being positive. Thus, because a larger universe of agents offers more possible indirect-friendship links following formation of a direct friendship, the chosen effort level rises. Similarly, higher friendship benefits, either direct or indirect, raise e. Note that the condition $\Phi'(e) < 0$, used above, is required for local stability of the equilibrium.

As a small extension, consider the effects of population turnover. If people move away in any period with a probability q, then established friendships are severed. Under this modification, the P function must be multiplied by $(1-q)$, so that (5) becomes $P^1[u(1-q) + (n-2)v(1-q)^2P] = C'$. Then, it is easily seen that $\partial e/\partial q < 0$, so that greater population turnover (a higher q) makes people less willing to devote effort to friendship formation, a natural result.

2.3. The social optimum

To compute the socially optimal effort levels, observe that the social welfare function is given simply by $W \equiv \sum_{i=1}^{n} B_i$. Referring to (3), the first-order condition for optimal choice of e_{il} is then

$$\begin{aligned}\frac{\partial W}{\partial e_{il}} \;&=\; \frac{\partial P(e_{il},e_{li})}{\partial e_{il}}\left[(u_{il}+u_{li}) \;+\; \sum_{h\in a(l),h\neq i}(v_{ih}+v_{hi})P(e_{lh},e_{hl})\right] \;+\\ &\quad\; \frac{\partial P(e_{il},e_{li})}{\partial e_{il}}\left[\sum_{k\in a(i),k\neq l}(v_{lk}+v_{kl})P(e_{lk},e_{kl})\right] \;-\; C'(e_{il}) \;=\; 0. \qquad (7)\end{aligned}$$

Satisfaction of the second-order condition for the social optimality problem, which requires that the Hessian matrix of W is negative definite, is not ensured and must be assumed.

To understand (7), start by considering the first line, which differs from the expression in (4) by the inclusion of the u_{li} and v_{hi} terms. The first of these terms gives the benefit enjoyed by individual l from a direct friendship with i. While agent i does not consider this benefit in choosing e_{il}, the planner recognizes that greater effort by i raises expected direct-friendship benefits for *both* i and l. Similarly, while an increase in e_{il} also raises the likelihood of establishing indirect friendships with individuals h who are acquainted with l, agent i considers only his own benefit v_{ih} from such links, ignoring the reciprocal benefits v_{hi} enjoyed by the other individuals. The planner, on the other hand, takes both benefits into account.

While the marginal cost term is seen again in the second line of (7), the first expression captures an effect that does not appear at all in the equilibrium condition (4). This effect arises because greater effort by i in establishing a link with l raises the chances that indirect friendships with l will be formed by other individuals k through agent i. The resulting benefits, which accrue to both l and k, are counted in this first expression.

Thus, two kinds of externalities are not taken into account in individual decision making. First, in choosing effort levels, an individual ignores the reciprocal benefits enjoyed by agents who become his direct or indirect friends. Second, he ignores his role in facilitating indirect friendships for other people, which make use of the direct links that he establishes.

The presence of these externalities naturally suggests that equilibrium effort levels will be too low. To establish this point formally, it is helpful to consider the symmetric case, where both the equilibrium and the optimum are easily characterized. Under symmetry, (7) reduces to

$$2P^1(e^*, e^*)[u \;+\; 2(n-2)vP(e^*, e^*)] \;\;=\;\; C'(e^*), \tag{8}$$

where e^* denotes the common, socially optimal effort level. Note that, compared to (5), a factor of 2 appears twice in (8). The outer 2 captures the planner's focus on the two-way benefits that occur in both direct and indirect friendships, which are not fully recognized in individual decisions. The inner factor of 2 captures indirect-friendship benefits for other individuals that flow through the direct link established by a given agent (as reflected in the second line of (7)).

The following result provides a comparison the equilibrium and socially optimal effort levels:

Proposition 1. *If the equilibrium in the symmetric case is unique and stable, then the common equilibrium effort level is smaller than the socially optimal level, with $e < e^*$. Thus, people do not expend enough effort in forming friendship links.*

This and all subsequent results are proved in the appendix.

3. Equilibrium Network Structure under Asymmetric Conditions

In the real world, the pattern of interaction in social networks often exhibits striking asymmetries, with certain individuals linked to many other agents while other individuals are poorly connected. To explore this phenomenon in the present model, the analysis now considers the effect of asymmetric conditions on equilibrium effort choices and on the resulting structure of friendship networks, as reflected in link probabilities.

Two related types of asymmetries are considered. In the first case, one individual is a "magnetic agent," offering greater direct and indirect friendship benefits than the other individuals, who remain symmetric. In the second case, one individual "knows everyone," being acquainted with the entire universe of agents, while other agents are each acquainted with only a portion of the universe. The analysis explores the effects of these asymmetries, focusing in particular on the question of whether links to the "attractive" agent are most likely to form. The analysis of each case is carried out assuming a small universe of agents, with $n = 3$ in the magnetic-agent case and $n = 5$ in the knows-everyone case. Somewhat surprisingly, the results do not automatically generalize to larger values of n.

The initial analysis is carried out under the assumption that the P function depends on the sum of the effort levels. The function is thus written $P(e_{ij} + e_{ji})$, with $P' > 0$ and $P'' < 0$. The discussion proceeds by first developing the equilibrium conditions for the magnetic-agent and knows-everyone cases, and the results for both cases are then presented and discussed simultaneously. Following this discussion, the analysis then explores the effect of a different specification for the probability function, where P depends on the minimum of the effort levels, being written $P(\min\{e_{ij}, e_{ji}\})$.

3.1. The magnetic-agent case

Suppose that the universe of agents consists of three individuals, all of whom are acquainted with one another, with individual 1 being the magnetic agent. Individuals 2 and 3 receive benefits of u_1 and v_1 from having the magnetic agent as a direct and indirect friend, respectively. Agents 2 and 3, by contrast, generate direct and indirect benefits of u_x and v_x, respectively, for their partners, where $u_1 > u_x$ and $v_1 > v_x$. A further maintained assumption is $u_1 - u_x > v_1 - v_x$, which says that the magnetic agent's advantage in direct-friendship benefits exceeds his advantage in indirect-friendship benefits. This assumption seems natural given that people will especially benefit from direct links with the magnetic agent.

Effort levels will reflect the partial symmetry of this setup. The efforts expended by agents 2 and 3 in attempting to link with 1 will be equal, with the common level denoted by e_{x1}. Similarly, agent 1 will expend the same efforts in attempting to link with 2 and 3, with the common level denoted e_{1x}. Finally, the efforts expended by agents 2 and 3 in attempting to link with one another will be equal and denoted e_{xx}. Because of these symmetries, the same total effort, denoted $\widetilde{e}$, will be expended on the 1–2 and 1–3 links, and this level will generally be different from the total effort expended on the 2–3 link, denoted $\widehat{e}$. These effort levels satisfy

$$\widetilde{e} \;=\; e_{1x} \;+\; e_{x1} \tag{9}$$

$$\widehat{e} \;=\; 2e_{xx}. \tag{10}$$

Adapting the first-order condition (4) to the present context, the equilibrium is determined by (9) and (10) along with following conditions:

$$P'(\widetilde{e})[u_x \;+ v_x P(\widehat{e})] \;=\; C'(e_{1x}) \tag{11}$$

$$P'(\widetilde{e})[u_1 \;+ v_x P(\widetilde{e})] \;=\; C'(e_{x1}) \tag{12}$$

$$P'(\widehat{e})[u_x \;+ v_1 P(\widetilde{e})] \;=\; C'(e_{xx}) \tag{13}$$

Note in (11) that additional effort by agent 1 raises the likelihood of direct friendships with 2 and 3, yielding benefits of u_x, while also increasing the chances of indirect friendships with these individuals via the 2–3 link, yielding benefits of v_x. In (12), additional effort by agents 2 or 3 in linking with 1 makes a direct friendship with the magnetic agent more likely, yielding a benefit of u_1, while also raising the chance of an indirect friendship with the other non-magnetic individual via agent 1, yielding a benefit of v_x. Finally, in (13), additional effort on the 2–3 link makes a direct friendship with the other non-magnetic agent more likely, yielding a benefit of u_x, while raising the chance of an indirect friendship with 1 via that agent, yielding a benefit of v_1.[2]

3.2. The knows-everyone case

Suppose instead that friendship benefits are symmetric across individuals, but that agents have different sets of acquaintances. In particular, let $n = 5$, and suppose that agent 1 "knows everyone," being acquainted with the four other agents. However, agents 2 and 3 are only acquainted with one another and with agent 1, while agents 4 and 5 are only acquainted with one another and with agent 1. The pattern of acquaintances thus looks like an hour glass, with agent 1 at the narrow point in the center, agents 2 and 3 at the bottom corners and agents 4 and 5 at the upper corners.

Although agent 1 is not magnetic in the above sense, establishing a direct friendship with him gives the other agents potential access to indirect friendships with people they could not otherwise reach. Thus, the kind of asymmetries seen in section 3 will arise, allowing the earlier effort notation to be used without change. Eliminating the subscripts on u and v, the equilibrium conditions for the knows-everyone case are then given by

$$P'(\widetilde{e})[u + vP(\widehat{e})] = C'(e_{1x}) \tag{14}$$

$$P'(\widetilde{e})[u + 3vP(\widetilde{e})] = C'(e_{x1}) \tag{15}$$

$$P'(\widehat{e})[u + vP(\widetilde{e})] = C'(e_{xx}) \tag{16}$$

along with (9)–(10). Observe that agent 1 faces exactly the same incentives for forming links as in the magnetic agent case, with a successful link yielding a direct friendship and one potential indirect friendship. Thus, (14) has the same form as (11). However, for agents 2, 3, 4, and 5 (who are again denoted by x), a direct friendship with agent 1 creates 3 potential indirect friendships, accounting for the factor of 3 in (15). Finally, the incentives for forming links with acquaintances other than individual 1 are the same as before for agents 2, 3, 4, and 5, with a successful link yielding a direct friendship and one potential indirect friendship. Thus, (16) has the same form as (13).

3.3. Results and discussion

To simplify the presentation of the results, let the term "nonattractive agents" denote the individuals other than agent 1. These individuals are agents 2 and 3 in the magnetic-agent case, and individuals 2, 3, 4, and 5 in the knows-everyone case. Then, the following results can be established:

Proposition 2. *In both the magnetic-agent and knows-everyone cases, the nonattractive agents expend more effort attempting to link with agent 1 than agent 1 expends attempting to link with them. The nonattractive agents expend an intermediate amount of effort in attempting to link with one another. More precisely,*

$$e_{x1} > e_{xx} \geq e_{1x}, \tag{14}$$

with the last inequality holding strictly in the knows-everyone case. The inequality $\widetilde{e} > \widehat{e}$ holds in the knows-everyone case, implying that direct friendships involving agent 1 are *more likely to form than direct friendships involving the nonattractive agents. In the magnetic-agent case, $\widetilde{e} > \widehat{e}$ holds provided that $v_1 - v_x$ is small.*

To understand these conclusions, consider first the magnetic-agent case. Observe that individuals 2 and 3 naturally expend substantial effort in attempting to become friends with the magnetic agent. However, the unattractiveness of individuals 2 and 3 as direct friends leads agent 1 to choose low effort levels, an outcome that is compounded because the indirect friendships generated by direct links with 2 and 3 also yield low benefits. By contrast, even though a direct friendship with the other nonmagnetic individual yields a low benefit for agent 2 or 3, the prospect of a direct friendship with 1 via the other agent leads to an effort level e_{xx} at least as large as e_{1x}.

With the effort levels on the 1–2 and 1–3 links unbalanced, the relationship between total effort on these links and total effort on the 2–3 link is ambiguous. However, a small differential between v_1 and v_x weakens the incentive for agents 2 and 3 to capture 1's indirect-friendship benefits via investments in the 2–3 link. In this situation, total effort on the 2–3 link falls short of effort of the links leading to the magnetic agent, which in turn makes friendships between the nonmagnetic individuals less likely than those involving agent 1.

While a strong incentive to link with agent 1 also exists in the knows-everyone case, this incentive arises because this agent, rather than being innately attractive, provides the path to many indirect friendships through his wide acquaintances. In contrast to the magnetic-agent case, however, an indirect friendship with agent 1 is no better than any other indirect relationship. This fact reduces the incentive for the nonattractive agents to form links with one another, mirroring the situation in the magnetic agent case when $v_1 - v_x$ is small. As a result, $\widetilde{e} > \widehat{e}$ holds unconditionally in the knows-everyone case.

Thus, the above results capture an individual's natural tendency to expend extra effort in forming links with people whose friendship is valuable, where the value arises either for innate reasons or because of the additional opportunities the friendship affords. Because of this extra effort, such friendship links may be more likely to form than links involving less desirable people. Thus, the model generates the natural conclusion that attractive individuals have lots of friends.[3]

It is natural to ask whether the conclusions of Proposition 2 continue to hold when the universe of agents grows in size. Interestingly, clearcut results cannot be derived for $n > 3$ in the magetic-agent case, which indicates that this case is less straightforward than it might appear at first. The source of the ambiguity is that, with more people present, indirect friendships play a larger role in governing effort choices, diluting the effect of the magnetic agent. However, numerical simulations of the model reveal that the conclusions of Proposition 2 are fairly robust to increases in the number of agents.[4]

Similarly, the results for the knows-everyone case do not generalize as the individuals at the top and bottom of the hourglass increase in number. The difficulty is that, since a larger n makes more indirect friendships possible without the need to go through agent 1, links to other nonattractive agents may become relatively more appealing.[5]

3.4. The effect of a minimum-effort specification

Some observers would argue that unbalanced effort levels are not conducive to a creating a friendship. In particular, it could be argued that, when one individual works hard at forming a link while the other is passive, the surplus effort has no effect. This view suggests that, rather than depending on the sum of the effort levels, the P function should depend on the minimum of the levels, being written $P(\min\{e_{ij}, e_{ji}\})$. The purpose of the analysis in this section is to explore the effect of this alternate specification on the equilibria in the magnetic-agent and knows-everyone cases. Note that in a fully symmetric case, the minimum-effort assumption has no impact.

Consider first the magnetic-agent case. Effort levels on the 1–2 and 1–3 links, which were previously asymmetric, will now be symmetric given that an increase in effort above the level chosen by the other individual yields no gain. The effort levels can be written as $e_{1x} = e_{x1} \equiv \widetilde{e}_m$ and $e_{xx} \equiv \widehat{e}_m$, where the notation recognizes that effective total effort on the links, previously equal to the sum, is now just equal to the symmetric individual levels (the m subscript is used to denote the minimum-effort case).

On the links involving the agent 1, where effort levels were previously asymmetric, the nonmagnetic individual's marginal benefit from a higher effort level will be greater than marginal cost. This agent would like a higher common effort level, but he recognizes that a unilateral increase is undesirable. By contrast, the magnetic agent's first-order condition holds as an equality. Because he has a lower incentive to expend effort, agent 1's decisions effectively govern the effort choices on these links. Thus, the following equation system determines the equilibrium for the magnetic-agent case:

$$P'(\widetilde{e}_m)[u_x + v_x P(\widehat{e}_m)] = C'(\widetilde{e}_m) \tag{19}$$

$$P'(\widetilde{e}_m)[u_1 + v_x P(\widetilde{e}_m)] > C'(\widetilde{e}_m) \tag{20}$$

$$P'(\widehat{e}_m)[u_x + v_1 P(\widetilde{e}_m)] = C'(\widehat{e}_m) \tag{21}$$

This system has the same pattern as (11)–(13), with (19) pertaining to the magnetic agent and (20) and (21) applying to agents 2 and 3. Note that each equation has the same LHS as in the previous system, with differences arising only in the C' arguments on the RHS.

The analogous equations for the knows-everyone case are written

$$P'(\widetilde{e}_m)[u + vP(\widehat{e}_m)] = C'(\widetilde{e}_m) \tag{22}$$

$$P'(\widetilde{e}_m)[u + 3vP(\widetilde{e}_m)] > C'(\widetilde{e}_m) \tag{23}$$

$$P'(\widehat{e}_m)[u + vP(\widetilde{e}_m)] = C'(\widehat{e}_m) \tag{24}$$

Again, (22) pertains to agent 1, while (23) shows that agents 2–5 would like a higher effort level on the links leading to 1, but find a unilateral increase undesirable.

Using these conditions, the following results can be established:

Proposition 3. *Under the minimum-effort specification, $\widehat{e}_m > \widetilde{e}_m$ holds in the magnetic-agent case. Thus, direct friendships between nonmagnetic agents are more likely to form than direct friendships involving the magnetic agent. By contrast, $\widehat{e}_m = \widetilde{e}_m$ holds in the knows-everyone case, so that all friendships are equally likely to form*

Remarkably, the minimum-effort specification reverses the expected outcome in the magnetic-agent case, making links with the agent 1 *less likely* than those involving other agents. The reason is that, while effort on the 1–2 and 1–3 links is governed by agent 1's low incentives to form friendships, effort on the 2–3 link takes account of the possible benefits for agents 2 or 3 of an indirect friendship with the magnetic agent via the other individual. As a result, more effort is expended on this link, and 2–3 friendships are more likely to form. By contrast, because an indirect friendship with agent 1 carries no extra benefit for nonattractive agents in the knows-everyone case, no additional effort stimulus is present on the 2–3 and 4–5 links. As a result, effort ends up being the same as on the links involving agent 1.

This analysis shows that, under the minimum-effort specification, it is no longer true that attractive agents have lots of friends. The lower incentives for friendship formation felt by such agents end up governing the effort levels on the links leading to them, making friendships no more likely than on other, less advantageous links. While this result is interesting from a formal standpoint, it obviously lacks realism, indicating that link formation may depend on more than just the minimum effort level of the agents.

4. An Intertemporal Extension

While the analysis up to this point has been based on a single-period model, it is useful to consider friendship formation in a two-period setting. In such a setting, indirect friendships formed in the first period effectively increase an individual's set of acquaintances, allowing direct friendships to be formed with these individuals in the second period.

To analyze effort choices in this two-period setting, imagine that agents are arrayed on an infinite line, with an individual's initial acquaintances consisting of the two adjacent agents. Concretely, this setup can be viewed as an urban neighborhood, where people initially know only their next-door neighbors. An agent will expend effort attempting to link to both his left and right neighbors, but the problem can be analyzed by only considering the choice of effort in the rightward direction. Let i be the individual under consideration, and suppose that his successive neighbors on the right are j, k, and l, with h being his immediate neighbor on the left.

Friendship formation is governed by the following rules. If expenditure of effort in period 1 establishes a direct friendship, then that friendship persists in period 2 without the need for additional effort. However, if the effort fails, with a frienship not forming in period 1, then expenditure of further effort in period 2 is futile. Successful period 1 friendships may generate an indirect friendship, and the resulting link can be developed into a direct friendship by expenditure of further effort in period 2.

Let $\eta < 1$ represent the common discount factor, and let second-period effort choices be denoted by a 2 subscript. Furthermore, suppose that P depends on sum of the effort levels, and let u and v be uniform across agents. Then, focusing on the portion of overall friendship benefits that are affected by e_{ij}, which mostly involve agents to the right of i, the relevant maximand is

$$(1+\eta)\Big[uP(e_{ij}+e_{ji}) \ + \ vP(e_{ij}+e_{ji})P(e_{jk}+e_{jk})\Big] \ - \ C(e_{ij}) \ +$$
$$\eta P(e_{ij}+e_{ji})P(e_{jk}+e_{jk})\Big[uP(e_{2ik}+e_{2ki}) \ + \ vP(e_{2ik}+e_{2ki})(1+P(e_{kl}+e_{lk})) \ - \ C(e_{2ik})\Big]$$
$$+ \ 2\eta P(e_{ij}+e_{ji})P(e_{ih}+e_{hi})P(e_{2jh}+e_{2hj})v. \tag{25}$$

The first line of (25) gives the expected present value of benefits for the direct and indirect friendships established in period 1, minus the cost of the initial effort. Note that individual k, who resides two doors to the right, is i's potential indirect friend in period 1. The second line of (25) gives the expected subsequent benefits that result from building on an indirect friendship with agent k. Note first that the bracketed expression is discounted and multiplied by the probability that k becomes i's indirect friend, with the cost of period-2 effort subtracted off. The first term in the bracketed expression gives the expected benefits of direct friendship with k, and the second term gives potential indirect-friendship benefits via k. These include benefits from indirect friendship with l, k's rightward neighbor, as well as benefits from indirect friendship with j. Note that even though i is already a direct friend of j, he benefits further from seeing j at parties hosted by k, to which he is now invited (being k's direct friend). Note that since the benefits in the second line are conditional on the formation of a link between j and k, the probability that i enjoys the benefits of indirect friendship with j via k is simply the probability of a direct link between i and k, $P(e_{2ik}+e_{2ki})$. This fact explains the appearance of the 1 in the middle bracketed term.

To understand the third line of (25), observe that if i becomes direct friends with both j and his left neighbor h in period 1, then these individuals, being indirect friends, can become direct friends in period 2 by exerting further effort. But i then enjoys additional indirect friendship benefits by seeing h at parties hosted by j and by seeing j at parties hosted by h. These benefits are captured by last line of (25).

Finally, note that (25) reflects the assumption that, in choosing period-2 effort, i is uncertain whether a direct friendship with k will lead to an indirect link with l. In particular, i does not know whether a direct link between k and l was formed in period 1, but instead chooses his period-2 effort based on the probability $P(e_{kl} + e_{lk})$ that such a link exists.

The first-order conditions for choice of e_{ij} and e_{2ik} can be computed and simplified by using symmetry of the equilibrium. Period-1 effort choices will take the same value, denoted e, for all agents, with second period choices also uniform and given by e_2. Let the subscript 2 indicate that the functions P, P', C, and C' are evaluated at e_2, with no subscript indicating evaluation at e. Then, the first order conditions for period-2 and period-1 effort levels are given by

$$P_2'(u \; + \; v(1+P)) \quad = \quad C_2' \tag{26}$$

$$P'\Big[(1+\eta)(u+vP) \; + \; \eta P(uP_2 + vP_2(3+P) - C_2)\Big] \quad = \quad C' \tag{27}$$

While these conditions do not afford any special insights, a comparative static calculation shows that e is increasing in the discount factor η, assuming stability of the equilibrium. Thus, as a higher discount rate puts more weight on future friendship benefits, particularly those from direct links forged in period 2 on the basis of period-1 successes, more period-1 effort is expended.

The main lesson of this analysis is that current investment in friendships pays future dividends by allowing new direct friendships to be built on the indirect links created today. This lesson would obviously be strenghtened in a setup with multiple future periods, where the span of direct friendship links would gradually expand over time to include an individual's entire neighborhood.

5. Conclusion

This paper has developed a model of social networks different from those presented in the recent literature. In contrast to existing models, the level of investment in link formation is a decision variable rather than being exogenous, and links form stochastically rather than deterministically, with the probability depending on the noncooperative investment choices of both parties. Since network structure is then stochastic rather than deterministic, the actual pattern of links cannot be specified, as in previous models, with the analysis focusing instead on which links are most likely to form. This alternate approach leads to a much simpler mathematical structure than in previous work.

The analysis is couched in the context of friendship networks, and its first lesson is that individual investment in friendship formation is too low. This result arises in part because an individual does not consider the gain to the other agent in deciding whether to increase his investment in a friendship link, but another externality is also involved. The discussion then explores the effect of several asymmetries on the equilibrium structure of friendship networks. It is shown that friendship links are likely to form when they involve an attractive agent, whose appeal arises either because of personal magnetism or a broad group of acquaintances. The analysis is also extended to an intertemporal setting, showing that current investment in friendship links can improve future options by allowing an individual's set of friends to grow over time.

Given the intense current interest in social networks, researchers are likely to benefit from availability of a greater variety of modeling approaches, especially ones that place fewer mathematical demands on the analyst. As a result, the present framework could help advance the state of knowledge in this important area.

Appendix

Proof of Proposition 1: Let (5) be rewritten (using (6)) as $\Phi(e) = 0$ and let (8) be written as $\Gamma(e^*) = 0$, where $\Gamma(e^*) = 2P^1[u + 2(n-2)vP)] - C'$. Then, since the factors of 2 in Γ imply $\Gamma(e^*) > \Phi(e^*)$, the fact that $\Gamma(e^*) = 0$ yields $\Phi(e^*) < 0$. But with stability of the equilibrium implying $\Phi'(e) < 0$ and uniqueness implying that $\Phi(e)$ has a single solution, it follows from $\Phi(e^*) < 0$ that the solution must satisfy $e < e^*$.

Proof of Proposition 2: Consider first the magnetic-agent case. To establish $e_{x1} > e_{1x}$, suppose that the contrary is true, with $e_{1x} \geq e_{x1}$. Then, given $C'' > 0$, the RHS of (11) must exceed or equal the RHS of (12), with the same relationship holding for the left-hand sides of the two equations. But given $u_1 > u_x$, the only way the latter relationship can hold is for $\widehat{e} > \widetilde{e}$ to be satisfied. Next note that the maintained assumption $u_1 - u_x > v_1 - v_x$ implies $u_1 + v_x P(\widetilde{e}) > u_x + v_1 P(\widetilde{e})$, given $P(\widetilde{e}) < 1$. Using $\widehat{e} > \widetilde{e}$ and the latter inequality, (12) and (13) then imply $e_{xx} < e_{x1}$. But, given (9) and (10), the only way that this last inequality can be consistent with $\widehat{e} > \widetilde{e}$ is for $e_{1x} < e_{x1}$ to hold. This inequality, however, violates the initial assumption that $e_{1x} \geq e_{x1}$, establishing its impossibility.

To establish $e_{xx} \geq e_{1x}$, assume the contrary, with $e_{xx} < e_{1x}$ holding. Given $C'' > 0$, $v_x \geq v_1$, and $P'' < 0$, the only way that (11) and (13) can be satisfied along with the previous assumption is for $\widehat{e} > \widetilde{e}$ to hold. But since $e_{xx} < e_{1x}$ together with $e_{x1} > e_{1x}$ (established above) yield $\widehat{e} = 2e_{xx} < 2e_{1x} < e_{1x} + e_{x1} = \widetilde{e}$, a contradiction arises, ruling out the initial assumption.

To establish $e_{x1} > e_{xx}$, again assume the contrary, with $e_{x1} \leq e_{xx}$ holding. Under this assumption, $\widetilde{e} > \widehat{e}$ must hold for (12) and (13) to be satisfied. But with $e_{xx} \geq e_{x1} > e_{1x}$ implied by the above assumption, it follows using (9) and (10) that $\widehat{e} > \widetilde{e}$, a contradiction that rules out the assumption.

To establish that $\widetilde{e} > \widehat{e}$ holds when $v_1 - v_x$ is small, suppose $v_1 = v_x$, an assumption that has no effect on the preceding arguments. Since $e_{xx} \geq e_{1x}$ continues to hold, it follows from inspection of (11) and (13) that $\widetilde{e} \geq \widehat{e}$ must now be satisfied, with $\widetilde{e} = \widehat{e}$ implying $e_{xx} = e_{1x}$. However, the case where $\widetilde{e} = \widehat{e}$ holds can be ruled out because $e_{xx} = e_{1x} < e_{x1}$ implies $\widehat{e} < \widetilde{e}$, a contradiction. Thus, $e_{xx} > e_{1x}$ and $\widetilde{e} > \widehat{e}$ must be satisfied, a conclusion that also holds by continuity when $v_1 - v_x$ is small.

For the knows everyone case, a similar proof applies. To establish $e_{x1} > e_{1x}$, suppose that the contrary is true, with $e_{1x} \geq e_{x1}$. Then, given the 3 factor in (15), the only way this relationship can hold is for $\widehat{e} > \widetilde{e}$ to be satisfied. Using (15) and (16), this inequality implies $e_{xx} < e_{x1}$. But, given (9) and (10), the only way that this last inequality can be consistent with $\widehat{e} > \widetilde{e}$ is for $e_{1x} < e_{x1}$ to hold. This inequality, however, violates the initial assumption that $e_{1x} \geq e_{x1}$, establishing its impossibility.

To establish $e_{xx} > e_{1x}$, assume the contrary, with $e_{xx} \leq e_{1x}$ holding. Then, (14) and (16) imply $\widehat{e} \geq \widetilde{e}$. But since $e_{xx} \leq e_{1x}$ together with $e_{x1} > e_{1x}$ yield $\widehat{e} = 2e_{xx} \leq 2e_{1x} < e_{1x} + e_{x1} = \widetilde{e}$, a contradiction arises, ruling out the initial assumption. Thus, $e_{xx} > e_{1x}$ holds, and (14) and (16) yield $\widetilde{e} > \widehat{e}$.

To establish $e_{x1} > e_{xx}$, again assume the contrary, with $e_{x1} \leq e_{xx}$ holding. Under this assumption, $\widetilde{e} > \widehat{e}$ must hold for (15) and (16) to be satisfied. But with $e_{xx} \geq e_{x1} > e_{1x}$, it follows using (9) and (10) that $\widehat{e} > \widetilde{e}$, a contradiction that rules out the assumption.

Proof of Proposition 3: In the magnetic-agent case, the first step is to show inequality must hold in (20) and equality must hold in (19) rather than the reverse. For the reverse relationships to hold, $\widehat{e}_m > \widetilde{e}_m$ must also be satisfied. But since $u_1 + v_x P(\widetilde{e}_m) > u_x + v_1 P(\widetilde{e}_m)$ is satisfied, (20) and (21) then imply that $C'(\widetilde{e}_m) > C'(\widehat{e}_m)$ holds, a contradiction. Equations (19)–(21) are then relevant, and if $\widetilde{e}_m \geq \widehat{e}_m$, then the LHS of (19) is less than the LHS of (21), implying that the same relationship holds for the RHS expressions. The latter conclusion, however, constitutes a contradiction, implying $\widetilde{e}_m < \widehat{e}_m$. For the knows-everyone case, an argument like that above establishes that inequality must hold in (23). The conclusion that $\widetilde{e}_m = \widehat{e}_m$ then follows from inspection of (22) and (24).

Note that this last conclusion presumes that the knows-everyone solution is unique. In other words, it is assumed that a solution to (22) and (24) other than the one where $\widetilde{e}_m = \widehat{e}_m$ does not exist.

References

Bala, V., Goyal, S., 2000a. A noncooperative model of network formation. *Econometrica* 68, 1181-1229.

Bala, V., Goyal, S., 2000b. A strategic analysis of network reliability. *Review of Economic Design* 5, 205-228.

Boorman, S.A., 1975. A combinatorial optimization model for transmission of job information through contact networks. *Bell Journal of Economics* 6, 216-249.

Calvó-Armegnol, A., Zenou, Y., 2001. Job matching, social networks and word-of-mouth communication. Unpublished paper, University of Southampton.

Dutta, B., Jackson, M.O., 2000. The stability and efficiency of directed communication networks. *Review of Economic Design* 5, 251-272.

Dutta, B., Jackson, M.O., 2002, On the formation of networks and groups. In: Dutta, B., Jackson, M.O. (Eds.), *Models of Strategic Formation of Networks and Groups*. Springer-Verlag, Berlin (forthcoming).

Ioannides, Y., Loury, L.D., 2002. Job information networks, neighborhood effects and inequality. Unpublished paper, Tufts University.

Jackson, M.O., Watts, A., 2002. The evolution of social and economic networks. *Journal of Economic Theory* 106, 265-295.

Jackson, M.O., Watts, A., 2002. On the formation of interaction networks in social coordination games. *Games and Economic Behavior* 41, 265-291.

Jackson, M.O., Wolinsky, A., 1996. A strategic model of social and economic networks. *Journal of Economic Theory* 71, 44-74.

Montgomery, J.D., 1991. Social networks and labor market outcomes: Towards an economic analysis. *American Economic Review* 81, 1408-1418.

Footnotes

*I thank Kangoh Lee for helpful comments. However, he is not responsible for any shortcomings in the paper.

[1]A related literature analyzes the role of social networks in providing information about job openings, which helps to facilitate the matching process in the labor market. See Boorman (1975) for an early study and Montgomery (1991) and Calvó-Armegnol and Zenou (2001) for more recent contributions. Ioannides and Loury (2002) provide a helpful survey of this literature.

[2]The earlier discussion mentioned that when P depends on the sum of the effort levels, the assumption that $C(e) \equiv e$ or any other linear form is not tenable. To understand this point, note that, under such an assumption, the right-hand sides of (11)–(13) are all constants. But since the system then contains three equations to solve for only two unknowns ($\widetilde{e}$ and $\widehat{e}$), it is overdetermined.

[3]If agent 1 is unattractive instead of magnetic, so that $u_1 < u_x$ and $v_1 < v_x$ hold, then it can be shown that all the inequalities in Proposition 2 are reversed. Agent 1 now expends the higher effort level on the 1–2 and 1–3 links, with agents 2 and 3 expending an intermediate effort level in linking with one another. Total effort is now higher on the 2–3 link when $v_1 - v_x$ is small, implying that agents 2 and 3 are more likely to be direct friends with one another than with the unattractive agent.

[4]To see the sources of this ambiguity, note that when $n > 3$, v_x in (11) and (12) is replaced by $(n-2)v_x$, while $v_1P(\widetilde{e})$ in (13) is replaced by $(n-3)v_xP(\widehat{e}) + v_1P(\widetilde{e})$. To understand the changes in (11) and (12), observe that by linking to one of the nonmagnetic individuals, agent 1 gains potential access to $n-2$ indirect friendships with other nonmagnetic agents. The same indirect benefits arise when a nonmagnetic agent links to agent 1. To understand the change in (13), note that when one nonmagnetic agent links to another, he gains potential access to $n-3$ indirect friendships with other nonmagnetic individuals as well as indirect access to the magnetic agent. The obstacle to generalizing Proposition 2 is that the inequality $u_1 + v_xP(\widetilde{e}) > u_x + v_1P(\widetilde{e})$, which follows from $u_1 - u_x > v_1 - v_x$ and plays a crucial role in the appendix proof, does not generalize to the case of a larger n. The equivalent inequality for this case is written $u_1 + (n-3)v_xP(\widehat{e}) + (n-2)v_1P(\widetilde{e}) > u_x + v_1P(\widetilde{e})$, and no simple condition ensures that it holds.

[5]To see the relevant changes in (14)–(16), let m denote the number of individuals on each end of the hourglass, so that $n = 2m+1$. Then v in (14) is replaced by $(m-1)v$, $3v$ in (15) is replaced by $(2m-1)v$, and $vP(\widetilde{e})$ in (16) is replaced by $(m-2)vP(\widehat{e}) + vP(\widetilde{e})$. With these substitutions, difficulties similar to those outlined in footnote 4 arise in the appendix proof

of Proposition 2.