The Evolution of Social and Economic Networks[1]

Matthew O. Jackson[2]

Humanities and Social Sciences 228-77, Caltech, Pasadena, California 91125
jacksonm@hss.caltech.edu

and

Alison Watts

Department of Economics, Southern Illinois University, Carbondale, Illinois 62901
wattsa@siu.edu

We examine the dynamic formation and stochastic evolution of networks connecting individuals. The payoff to an individual from an economic or social activity depends on the network of connections among individuals. Over time individuals form and sever links connecting themselves to other individuals based on the improvement that the resulting network offers them relative to the current network. In addition to intended changes in the network there is a small probability of unintended changes or errors. Predictions can be made regarding the likelihood that the stochastic process will lead to any given network at some time, where the stochastic process selects from among the statically stable networks and cycles. We apply these results to examples including the Gale–Shapley marriage problem. Thus the paper achieves two goals. First, it outlines a dynamic solution concept for networks. Second, it applies this concept to matching problems. *Journal of Economic Literature* Classification Numbers: A14, D20, J00.   © 2002 Elsevier Science (USA)

Key Words: networks; evolution; stochastic stability.

1. INTRODUCTION

Network structure is important in determining the outcome of many important social and economic relationships. For example, networks play a fundamental role in determining how information is exchanged. Such

information may be as simple as an invitation to a party, or as consequential as information about job opportunities (e.g., Boorman [6], Montgomery [31], and Topa [37]), literacy (e.g., Basu and Foster [3]), consumer products (e.g., Ellison and Fudenberg [12] and Iacobucci and Hopkins [21]), or even information regarding the returns to crime (e.g., Glaeser *et al.* [18]). Networks also play fundamental roles in the payoffs earned from bargaining with an organization (e.g., Wang and Wen [41]) and in the exchange of goods and services. Exchange examples include trading networks and alliances (e.g., Bell [4], Maxfield [29], Kirman *et al.* [27], Tesfatsion [38, 39], Weisbuch *et al.* [42]), and networks through which financial help or insurance is exchanged in developing countries (e.g., Fafchamps and Lund [14]). Even standard matching problems (e.g., the marriage and college admissions problems studied by Gale and Shapley [17] and Roth and Sotomayor [34]) are special situations where network relationships are important.

Despite the fundamental importance of network structures in many social and economic settings, the development of foundational theoretical models to analyze how the decisions of individuals contribute to network formation is still in its infancy. In this paper, we examine the dynamic formation and stochastic evolution of networks, taking into account the incentives that individuals have to form (or sever) links with each other. We make two contributions in this regard. First, we develop a working model of the dynamic formation of non-directed networks, which can be applied to a wide range of economic and social settings. Second, we apply this model to matching problems.

An Overview of the Model and Results

Our approach is to model network formation as a dynamic process in which individuals form and sever links based on the improvement that the resulting network offers them relative to the current network. This deterministic dynamic process may end at stable networks or in some cases may cycle. To this basic deterministic dynamic we add random perturbations and examine the distribution over networks as the level of random perturbations goes to 0. This stochastic dynamic process refines the prediction of the deterministic process and provides a robustness check to see which networks predicted by the deterministic dynamic are most stable in the face of small perturbations. As one should expect some noise in applications, the stochastic part of the analysis is important.

To be more specific, networks are modeled as graphs, where nodes or vertices represent individuals and links or edges represent connections between the individuals. Links are non-directed and thus reciprocal. A link between two individuals can be formed only if both individuals agree to add the link, while a single individual can sever an existing link. Each

individual receives a payoff or net benefit based on the network configuration that is in place. This payoff can be interpreted as the utility or production that an individual obtains from the social interaction that occurs through the network. As our approach does not require that we make any specific assumptions on how value is generated by networks, the model is very flexible and can be applied to a wide variety of social and economic settings.

The primary tool that we introduce to analyze dynamic network formation is the concept of a sequence of networks that emerge when individuals form or sever links based on the improvement the resulting network offers relative to the current network. Such a sequence, called an "improving path,"[3] has the properties that (i) each network in the sequence differs from the previous network by the addition or deletion of a single link, and (ii) the addition or deletion of the link benefits the individual(s) whose consent is necessary for the change.

The notion of improving path is somewhat myopic in that individuals do not forecast how their decision to add or sever a link might affect future decisions of other individuals or more generally how it might influence the future evolution of the network. Such myopic behavior is natural in the context of large networks where players may have limited information about the incentives of others, and generally provides a useful starting point for the study of the evolution of networks. Another limitation is that only one link is considered at a time. However, this is easily extended and later in the paper, in the context of matching models, we show that this methodology is adaptable to allow for the simultaneous change of a number of links.

The improving paths emanating from any starting network must lead to either a pairwise stable network (where no two players want to form a link, and no individual player wants to sever a link) or a cycle (where a number of networks are repeatedly visited). We show that there always exists either a pairwise stable network or a cycle from which there is no exit. We give a simple trading network example to show that it is possible for cycles to exist while pairwise stable networks fail to exist.

We then use improving paths as the foundation for a stochastic analysis, where in addition to intended changes in the network, unintended mutations or errors are introduced. Such unintended changes may be due to exogenous forces acting on the network, or simply miscalculations or errors on the part of an individual making an assessment or taking an action. Such a process can be described as a Markov chain and we can apply results concerning limiting behavior of Markov processes (Freidlin and

---

[3] This is similar to the concept of improvement paths of Monderer *et al.* [30], except for the bilateral nature of the improvement in our setting.

Wentzell [16], as adapted to discrete settings by Kandori *et al.* [26] and
Young [43]). Specific predictions can be made concerning the relative
amounts of time that the stochastic dynamic process will spend in various
networks. The process naturally gravitates to pairwise stable networks and
cycles, but periodically is bumped away by the random errors or mutations.
The intuition for which networks are visited most often comes from the
idea of resistance of Freidlin and Wentzell [16]. In the network context,
resistance keeps track of how many errors or mutations are needed to get
from some given network to an improving path leading to another
network. Very roughly, networks that are harder to get away from and
easier to get back to, in terms of resistance, are favored by the stochastic
process (although this favoritism depends on the full configuration of
resistance among different networks). We apply these ideas to several
examples to study the set of stochastically stable networks that emerge
under this stochastic dynamic process.

In the last section of the paper, we apply the stochastic dynamic network
formation model to matching problems, such as the Gale–Shapley marriage
problem and the college admissions (hospital-intern) problem (see Roth
and Sotomayor [34]). Such matching problems fit nicely into a network
setting and the methodology outlined above can be used to analyze which
matchings one expects to arise endogenously, in the absence of some coor-
dinating procedure. Theorem 2 shows that, in these problems, the set of
stochastically stable networks coincides with the set of (static) core stable
networks, which are necessarily Pareto efficient. Examples show how this
relationship depends on the definition of improving path that is applied.

The Closely Related Literature

The papers most closely related to this one are Jackson and Wolinsky
[25], Dutta and Mutuswami [10], and Watts [40].

The model and notion of pairwise stability that underlies the analysis
conducted here is from Jackson and Wolinsky [25]. Their focus was on
developing a model for the study of (static) stability of networks and using
this model to understand the relationship between stability and efficiency
of networks. Dutta and Mutuswami [10] and Jackson [22] have looked at
this relationship in further detail. As those analyses are static, they leave
open the question of which stable networks will form, if any, as they do not
consider cycles.

Watts [40] analyzes the formation of networks in a dynamic framework.
She extends the Jackson and Wolinsky [25] model to a dynamic process,
but limits attention to the specific context of the "connections model" dis-
cussed by Jackson and Wolinsky and a particular deterministic dynamic.[4]

---

[4] See also Hummon [20] for some simulations of a dynamic in the connections model.

The new contributions here are in terms of both the wider collection of network models that are admitted, as well as the introduction of the concept of improving paths, the stochastic process, and the application to the matching problems.

There are also other papers that provide theoretical models of network formation in strategic contexts. Aumann and Myerson [1] were the first to take an explicit look at network formation in a strategic context where individuals had discretion over their connections; these connections defined a communication structure that was applied to a cooperative game. Slikker and van den Nouweland [36] have extended the Aumann and Myerson [1] model to a one-stage model of link formation and payoff division. However, the analysis in those papers is devoted to issues in cooperative game theory such as the characterization of value allocations (see also, Myerson [32]).[5]

Recent work by Bala and Goyal [2] is close in motivation to our paper, as they are also interested in the dynamics of network formation. However, their approach differs significantly from ours in some basic ways both in terms of modeling and results. First, their focus is directed communication networks, similar to a directed version of the connections model (see also Dutta and Jackson [8]). Directed networks allow one individual to connect to another without the consent of the second individual, and thus applications are to settings such as advertising, the sending of mail, and such. These are fundamentally different applications from the non-directed networks that we consider here where both individuals need to consent to form a relationship, which applies to social relationships such as friendship and marriage, as well as trading relationships, insurance networks, job contacts, bargaining networks, etc. Second, the directed networks end up having different incentive properties as individuals can unilaterally form new links, whereas here we need to consider the incentives of two individuals in forming a link. Third, Bala and Goyal examine a repeated game, and focus on learning as a way to identify equilibria, which is quite different from the stochastic dynamic we examine here.

Finally, recent work on stochastic stability in game theory studies how individuals play games when social structure determines who interacts with whom (see for example, Ellison [11], Ellison and Fudenberg [12], Ely [13], and Young [44]). That literature examines which strategies players play in a game when the set of opponents that a player might face depends on the social structure. The mutations are in the strategy that a given player uses, and the literature examines how strategy choices and rates of convergence depend on social structure. Here instead, our interest is in the formation of the social structure itself. So we abstract away from behavior

---

[5] See Dutta and Jackson [9] for an overview of the network formation literature.

that takes place after the network is in place (modeling that simply via a payoff that each player gets as a function of the network) and so mutations take place on the social structure itself.[6] The aim of this paper is to build a dynamic model of network formation, showing that the Markov tools underlying the game theoretic literature are natural and useful here, and then to illustrate the model through some examples including an application to matching problems.

The remainder of the paper is organized as follows. In Section 2 we provide the definitions comprising the basic model. In Section 3 we analyze improving paths and cycles. Section 4 contains the stochastic analysis and dynamic stability results, including a discussion of dynamic stability and efficiency. In Section 5 we apply the stochastic process to a class of matching problems.

2. A MODEL OF NETWORKS

The model of social and economic networks that we consider is based on that of Jackson and Wolinsky [25], henceforth referred to as JW. The following definitions outline the model.

Players

Let $N = \{1, ..., n\}$ be the finite set of *players*. Depending on the application, a player may be a single individual, a firm, a country, or some other autonomous unit.

Networks and Graphs

The network relations among the players are represented by graphs whose nodes or vertices represent the players and whose links (edges or arcs) capture the pairwise relations. We focus on non-directed networks where links are reciprocal. The complete network, denoted g^N, is the set of all subsets of N of size 2. The set of possible *networks or graphs* on N is $\{g \mid g \subset g^N\}$. The subset of N containing i and j is denoted ij and is referred to as the *link ij*. The interpretation is that if $ij \in g$, then nodes i and j are directly connected, while if $ij \notin g$, then nodes i and j are not directly connected.

Let $g + ij$ denote the network obtained by adding link ij to the existing network g and let $g - ij$ denote the network obtained by deleting link ij from the existing network g (i.e., $g + ij = g \cup \{ij\}$ and $g - ij = g \setminus \{ij\}$).

---

[6] In applications where players strategically interact after a network has formed, these two literatures can ultimately be brought together as it is then natural to have mutations in both the social structure and in the strategy choice. See Jackson and Watts [24], Goyal and Vega-Redondo [19], and Droste *et al.* [7] for such analyses and further discussion.

If $g' = g + ij$ or $g' = g - ij$, then we say that g and g' are *adjacent*.

Let $N(g) = \{i \mid \exists j \text{ s.t. } ij \in g\}$ be the set of players involved in at least one link and $n(g)$ be the cardinality of $N(g)$.

Chains and Components

A *chain*[7] *in g connecting* i_1 *and* i_n is a set of distinct nodes $\{i_1, i_2, ..., i_n\} \subset N(g)$ such that $\{i_1 i_2, i_2 i_3, ..., i_{n-1} i_n\} \subset g$.

A nonempty network $g' \subset g$ is a *component* of g, if for all $i \in N(g')$ and $j \in N(g')$, $i \neq j$, there exists a chain in g' connecting i and j, and for any $i \in N(g')$ and $j \in N(g)$, $ij \in g$ implies $ij \in g'$.

Value Functions and Strong Efficiency

The *value of a network* is represented by $v: \{g \mid g \subset g^N\} \to \mathfrak{R}$, where $v(g)$ represents the total utility or production of the network. The set of all such functions is V. The value function allows for a wide variety of applications and quite general forms of externalities. In some applications the value will be an aggregate of individual utilities or production values, $v(g) = \Sigma_i u_i(g)$, where $u_i: \{g \mid g \subset g^N\} \to \mathfrak{R}$.

A network $g \subset g^N$ is *strongly efficient* if $v(g) \geqslant v(g')$ for all $g' \subset g^N$. Strong efficiency and Pareto efficiency coincide when value is transferable.

Allocation Rules and Pareto Efficiency

An *allocation rule* $Y: \{g \mid g \subset g^N\} \times V \to \mathfrak{R}^N$ describes how the value associated with each network is distributed to the individual players. $Y_i(g, v)$ may be thought of as the payoff to player i from network g under the value function v. For simplicity, if v is fixed, we will simply write $Y_i(g)$.

The allocation rule may represent several things. When considering a purely social network, the allocation rule may represent the utility that each individual receives from the network and this utility might not be transferable. When considering an exchange or production network, the allocation rule may represent either the trades or production accruing to each individual, the outcome of a bargaining process, or some exogenous redistribution.

Pairwise Stability

The following concept describes networks for which no player would benefit by severing an existing link, and no two players would benefit by forming a new link.

---

[7] A chain is sometimes referred to as a path in the literature, but we reserve the term path for a sequence of networks and use chain for a sequence of links.

A network g is *pairwise stable* with respect to v and Y if

    (i)   for all $ij \in g$, $Y_i(g, v) \geqslant Y_i(g - ij; v)$ and $Y_j(g, v) \geqslant Y_j(g - ij; v)$, and

    (ii)   for all $ij \notin g$, if $Y_i(g, v) < Y_i(g + ij; v)$ then $Y_j(g, v) > Y_j(g + ij; v)$.

When *a* network g is not pairwise stable it is said to be *defeated* by g' if either $g' = g + ij$ and (ii) is violated for ij, or if $g' = g - ij$ and (i) is violated for ij.

There are variations on the notion of pairwise stability discussed in JW. Dutta and Mutuswami [10] discuss alternative approaches that capture coalitional deviations.

3. IMPROVING PATHS AND CYCLES

Before proceeding to study a stochastic network formation process, we first describe the deterministic sequences that are followed as a network evolves. To avoid confusion, we emphasize that *an improving path represents changes from one network to another*, rather than a chain of links within a given network.

Improving Paths

An improving path is a sequence of networks that can emerge when individuals form or sever links based on the improvement the resulting network offers relative to the current network. Each network in the sequence differs by one link from the previous network. If a link is added, then the two players involved must both agree to its addition, with at least one of the two strictly benefiting from the addition of the link. If a link is deleted, then it must be that at least one of the two players involved in the link strictly benefits from its deletion.

Formally, an *improving path* from a network g to a network g' is a finite sequence of adjacent networks $g_1, ..., g_K$ with $g_1 = g$ and $g_K = g'$ such that for any $k \in \{1, ..., K-1\}$ either:

    (i) $g_{k+1} = g_k - ij$ for some ij such that $Y_i(g_k - ij) > Y_i(g_k)$, or

    (ii) $g_{k+1} = g_k + ij$ for some ij such that $Y_i(g_k + ij) > Y_i(g_k)$ and $Y_j(g_k + ij) \geqslant Y_j(g_k)$.

So, an improving path is a sequence of adjacent networks that might be observed in a dynamic process where players are adding and deleting links, one at a time.

The behavior implicit in an improving path may be myopic in the following sense. A player might delete a link making him or herself better off, but this deletion may lead a second player to delete a second link which

in turn leaves the first player worse off relative to the starting position. If the first player foresaw this, he or she might choose not to sever the link to begin with. This sort of consideration is not taken into account in our analysis, and may be important when there are relatively small numbers of forward-looking players who are well-informed about the value of the network and the motivations of others. However, in larger networks and networks where players' information might be local and limited, or in networks where players significantly discount the future, myopic behavior is a more natural assumption, and a reasonable starting point for our analysis. In our setting all a player needs to know is whether adding or deleting a given link is directly beneficial to him or her in the current circumstances.

In addition to the assumption of myopic behavior, there are other assumptions in the definition of improving path that can be varied. For instance, only a change of a single link is considered at a time,[8] but the definition is easily adapted to allow for the simultaneous addition or deletion of several links at a time. We consider this possibility of simultaneous actions when we discuss matching problems in Section 5.

The improving paths emanating from any starting network lead either to a pairwise stable network or to a cycle (where a number of networks are repeatedly visited in some sequence). In fact, pairwise stability of a network is equivalent to saying that a network has no improving paths emanating from it.

Cycles

A set of networks C, form a *cycle* if for any $g \in C$ and $g' \in C$ there exists an improving path connecting g to g'.

A cycle C is a *maximal cycle* if it is not a proper subset of a cycle.

A cycle C is a *closed cycle* if no network in C lies on an improving path leading to a network that is not in C. A closed cycle is necessarily a maximal cycle.

The concept of improving path provides for an easy proof of the following simple existence result.

LEMMA 1. *For any v and Y there exists at least one pairwise stable network or closed cycle of networks.*

Proof. A network is pairwise stable if and only if it does not lie on an improving path to any other network. So, start at any network. Either it is pairwise stable or it lies on an improving path to another network. In the first case the result is established. So, consider the second case. Follow the

---

[8] So, in a sense the notion of improving path looks at better-than reply rather than best reply dynamics. (See Ritzberger and Weibull [33] for a discussion of better-than reply dynamics.)

improving path. Given the finite number of possible networks, either the improving path ends at some network which has no improving paths leaving it, which then must be pairwise stable, or it can be continued through each network it hits. In the second case, the improving path must form a cycle. Thus, we have established that there always exists either a pairwise stable network or a cycle. Consider the case where there are no pairwise stable networks. Since there must exist a cycle, given the finite number of networks there must exist a maximal cycle. By the definition of maximal cycle and the non-existence of a pairwise stable network, there must be at least one maximal cycle for which there is no improving path leaving the cycle. (There can be improving paths leaving some of the maximal cycles, but these must lead to another maximal cycle. If all maximal cycles had improving paths leaving them, then there would be a larger cycle, contradicting maximality.) Thus, there exists a closed cycle. ∎

It is necessary in Lemma 1 to allow for the existence of either pairwise stable networks or closed cycles. There are cases where only pairwise stable networks exist and no cycles exist. For instance, in a model where adding a link is always beneficial, all networks are on an improving path to the complete network. Also, there are cases where only closed cycles exist and there are no pairwise stable networks. The following example illustrates this point.

Trading Example (Nonexistence of a Pairwise Stable Network)

Players benefit from trading with other players with whom they are linked, and trade can only flow along links. Players begin by forming a network. Subsequently, they receive random endowments and trade along chains of the network. Trade flows without friction along any chain and each connected component trades to a Walrasian equilibrium. The expected utility for a player of being in a given network is calculated by expecting over the Walrasian equilibria that result in the player's connected component as a function of realized endowments.

There are two goods. All players have identical utility functions for the two goods which are symmetric Cobb–Douglas, $U(x, y) = xy$. Each player has a random endowment, which is independently and identically distributed. A player's endowment is either $(1, 0)$ or $(0, 1)$, each with probability $1/2$, realized after the network is in place. For a given network, Walrasian equilibria occur on each connected component, regardless of the configuration of links. For instance, three players in a line have the same trades as three players in a circle (triangle), but with a lower total cost of links. Let the cost of a link be equal to $5/96$ (to each player in the link).

Let us show that if n is a least 4 then there does not exist a pairwise stable network. The utility of being alone is 0. Not accounting for the cost

of links, the expected utility for a player of being connected to one other is $1/8$. This is calculated by noting that there is a $1/2$ probability that the realized endowments will differ, in which case the players will trade to an allocation of $(1/2, 1/2)$ which results in a utility of $1/4$ for each of the two players. There is also a $1/2$ probability that the realized endowments will be identical in which case the utility will be 0 for each player. Similar calculations show that, not accounting for the cost of links, the expected utility for a player of being connected (directly or indirectly) to two other players is $1/6$; and of being connected to three other players is $3/16$. Most importantly, the expected utility of a player is strictly increasing and concave in the number of other players that she is directly or indirectly connected to, ignoring the cost of links.

Accounting for the cost of a link, if k players are in a connected component of a pairwise stable network, then there must be exactly $k-1$ links. If there are more than $k-1$ links, then there is at least one link that can be severed without changing the component structure of the network and so some player can sever a link and save the cost of the link without losing any expected utility from trading.

Note that if g is pairwise stable, then any component with 3 or more players cannot contain a player who has just one link. This follows from the fact that a player connected to some player who is not connected to anyone else, loses at most $1/6 - 1/8 = 1/24$ in expected utility by severing the link, but saves the cost of $5/96$ and so should sever this link.

From these two observations it follows that if a pairwise stable network existed, then it would have to consist of pairs of connected players (as two completely unconnected players benefit from forming a link), with one unconnected player if n is odd. If n is at least 4, then there must exist at least two pairs. However, such a network is not pairwise stable, since any two players in different pairs gain from forming a link (their utility goes from $1/8 - 5/96$ to $3/16 - 10/96$). Thus, there is no pairwise stable network. From Lemma 1, we know that there exists a closed cycle. An instance of a cycle in this trading example is $\{12, 34\}$ to $\{12, 23, 34\}$ to $\{12, 23\}$ to $\{12\}$ to $\{12, 34\}$. (A closed cycle includes many more networks, as there are a number of alternative improving paths from each of these networks.)

4. A STOCHASTIC PROCESS AND THE EVOLUTION OF NETWORKS

The notion of improving path provides a basis for a dynamic process of network formation, where starting at a given network one can make predictions about which network(s) might result. However, in some simple

examples it makes strange predictions. For instance, consider an example where there are economies of scale and a first link's costs exceed its benefits, but thereafter additional links are valuable. Starting at the empty network, a deterministic formation process based on the idea of improving paths will stay at the empty network, even though all players would be better off with a more connected network. There are at least two ways that one might expect to get around this. First, if players are forward looking instead of myopic they might add a first link anticipating where the process will lead. This is reasonable in situations where players are well informed about the value of overall networks and about the incentives of other players, but not in larger more anonymous systems. We stick with a myopic model. Second, if there are random (exogenous) events that affect network structure, then at some point a link might form and seed the formation process. This is what we examine in some detail. Considering such mutations has been successful in the game theoretic (stochastic stability, in the language of Foster and Young [15]) literature in weeding out tenuous equilibria, that are similar in spirit to the economy of scale example discussed above.

Stochastic mutations in the formation process may arise for a variety of reasons. Mutations may be due simply to exogenous (unmodeled) factors that are beyond the players' control. Mutations may be due instead to the bounded rationality of the players. Players simply make errors in calculating whether adding or severing a link is beneficial. Or it may be that players have limited information, and occasionally experiment to see if adding or severing a link will make them better off. It is quite likely that players may have such limited information in large network settings.

What follows may be also thought of as a check on the robustness of pairwise networks or cycles. Although a number of networks may be pairwise stable, they can differ in how resilient they are to random mutations. For instance, it may be relatively easy to leave and hard to get back to some networks, and vice versa for others.

The discussion that follows takes as given Y and v, although we omit notation indicating this dependence.

A Stochastic Dynamic Process

At a discrete set of times, $\{1, 2, 3, \ldots\}$ decisions to add or sever a link are made. At each time a pair of players ij is randomly identified with probability $p(ij) > 0$. The (potential) link between these two players is the only link that can be altered at that time. (One may think of a random meeting process where players randomly bump into each other and time is identified with the bumping times.) If the link is already in the network, then the decision is whether to sever it, and otherwise the decision is whether to add the link. The players involved act myopically, adding the link if it

makes each at least as well off and one strictly better off, and severing the link if its deletion makes either player better off. After the action is taken, there is some small probability $\varepsilon > 0$ that a mutation (or tremble) occurs and the link is deleted if it is present, and added if it is absent.[9] Thus mutations or trembles operate on link creation and destruction.

The above process naturally defines a Markov chain with states being the network in place at the end of a given period. The Markov chain is irreducible and aperiodic given non-zero trembles, and thus has a unique corresponding stationary distribution. As ε goes to zero, the stationary distribution converges to a unique limiting stationary distribution.

The following definitions are adaptations of ideas of Foster and Young [15], Kandori *et al.* [26], and Young [44] (as based on Freidlin and Wentzell [16]) to the network setting.

Stochastic Stability

A network that is in the support of the limiting (as ε goes to 0) stationary distribution of the above-described Markov process is *stochastically stable*.

For a given network, g, let $\text{im}(g) = \{g' \mid \text{there exists an improving path from } g' \text{ to } g\}$.

Resistance of a Path

A path $p = \{g_1, ..., g_K\}$ is a sequence of adjacent networks. The *resistance* of a path $p = \{g_1, ..., g_K\}$ from g' to g, denoted $r(p)$, is computed by $r(p) = \sum_1^{K-1} I(g_i, g_{i+1})$, where $I(g_i, g_{i+1}) = 0$ if $g_i \in \text{im}(g_{i+1})$ and $I(g_i, g_{i+1}) = 1$ otherwise.

Resistance keeps track of how many mutations must occur along a specific path to follow that path from one network to another. A mutation is necessary to move from one network to an adjacent one whenever it is not in the relevant players' interests to sever or add the link that distinguishes the two adjacent networks. This notion of resistance of a path then provides a natural notion of the resistance from one network to another as the path of least resistance from one to the other.

Let $r(g', g) = \text{minimum}\{r(p) \mid p \text{ is a path from } g' \text{ to } g\}$ and set $r(g, g) = 0$. Note that $r(g', g) = 0$ if and only if g' is in $\text{im}(g)$ or $g' = g$. Thus, if g' and g are in the same cycle, then $r(g', g) = 0$.

g-Trees

Given a network g, a *g-tree* is a directed graph which has as vertices all networks and has a unique directed path leading from each g' to g. Let

---

[9] The probability of a mutation is the same for any link. See Bergin and Lipman [5] for a discussion of the role of such an assumption.

$T(g)$ denote all the g-trees, and represent a t in $T(g)$ as a collection of ordered pairs of networks, so that $g'g'' \in t$ if and only if there is a directed edge connecting g' to g'' in the g-tree t.

Resistance of a Network

The *resistance* of a network g is computed as $r(g) = \min_{t \in T(g)} \sum_{g'g'' \in t} r(g', g'')$.

THEOREM 1. *The set of stochastically stable networks is the set* $\{g \mid r(g) \leqslant r(g') \text{ for all } g'\}$.

The proof of Theorem 1 is in the appendix. It is an application of results of Freidlin and Wentzell [16], as adapted by Young [43] and Kandori *et al.* [26], concerning limiting distributions of aperiodic, irreducible Markov processes. Note that the set of stochastically stable networks is always nonempty as we are taking a minimum over a finite set.

Remark. Theorem 1 (and the results below) hold for any definition of improving path. As illustrated in Section 5, these results are easily adapted to variations on the definition of improving path.

Before illustrating the implications of Theorem 1, we provide a couple of auxiliary results.

LEMMA 2. *If* $g' \in \text{im}(g)$ *and* $g \notin \text{im}(g')$, *then* $r(g) \leqslant r(g')$, *with strict inequality if g is pairwise stable or in a closed cycle. Thus, if g is stochastically stable, then either g is pairwise stable or part of a closed cycle. Furthermore, if one network in a closed cycle is stochastically stable then all networks in the closed cycle are stochastically stable.*

Lemma 2 has analogs in the game theory literature (for instance in Young [43]). The proof is routine and appears in the Appendix.

Noting that only pairwise stable networks and closed cycles matter in the dynamic process, and that resistance along improving paths is 0, simplify the calculations of resistance as follows. Given a closed cycle C and a network g, let $r(C, g) = r(g', g)$ where g' is any network in C, and similarly $r(g, C) = r(g, g')$ where g' in C. These are well defined since $r(g', g) = r(g'', g)$ (and similarly $r(g, g') = r(g, g'')$) for any g' and g'' in C since there is a path of zero resistance between g'' and g'.

Given a network g, a *restricted g-tree* is a directed graph which has as its root g, and as other vertices the pairwise stable networks and closed cycles, and has a unique directed path leading from each other vertex to g. Denote the set of restricted g-trees by $RT(g)$. In the following lemma, we let x denote a generic vertex which could be a network g' or a cycle C. The following lemma follows from a result by Young [1993], as shown in the Appendix.

LEMMA 3. $r(g) = \min_{\tau \in RT(g)} \sum_{x'x'' \in \tau} r(x', x'')$.

The implication of Lemma 3 is that the resistance of a network may be calculated by restricting attention to resistance between pairwise stable networks and closed cycles. Any network that is not pairwise stable or in a closed cycle lies on an improving path to a pairwise stable network or closed cycle, and so their additional consideration does not add to the resistance.

Let us illustrate the stochastic dynamic formation of networks in the following examples.

In the first example there are 22 pairwise stable networks, while the only stochastically stable network is the pairwise stable network from this set with the lowest total value and lowest allocation per player.

Co-author Example

Consider the co-author model of JW with 7 players. Each player is a researcher who spends time writing papers. If two players are connected, then they are working on a paper together. The amount of time researcher i spends on a given project is inversely related to the number of projects, n_i, that he is involved in. Formally, player i's payoff is represented as

$$u_i(g) = \sum_{j\,:\,ij \in g} \frac{1}{n_i} + \frac{1}{n_j} + \frac{1}{n_i n_j}$$

for $n_i > 0$. For $n_i = 0$ we assume that $u_i(g) = 0$. Again, $v(g) = \sum_{i \in N} u_i(g)$, and $Y_i(v, g) = u_i(g)$. Here, the interesting tradeoffs from connection come from the benefit of gains from a co-author's time $(1/n_j)$, at the expense of diluting the synergy (interaction) term $1/(n_i n_j)$ with other co-authors.

If $n = 7$ (see Proposition 4 in JW), there are 22 pairwise stable networks. One is the complete network and the 21 others are networks where 5 players are completely interconnected and the 2 remaining players are connected only to each other. There are no cycles. None of the pairwise stable networks are strongly efficient. Moreover, the complete network is strictly Pareto dominated by any of the 21 other pairwise stable networks.

Let us show that the complete network is the unique stochastically stable network. Consider a restricted g-tree for the complete network. Each of the other pairwise stable networks has distance 1 from an improving path to the complete network. By severing the link between the two paired players, one obtains a network on an improving path to the complete network (either player will link with a member of the group of 5 if they have that opportunity, and then would link with each of the others, and so forth). Thus, the complete network has a resistance of $7!/(5!2!) = 21$, which is the minimum possible given the number of pairwise stable networks [22].

However, consider a restricted g-tree for one of the other pairwise stable networks. The complete network lies more than a distance of 1 away from an improving path leading to some other pairwise stable network, since severing only one link leads to a network only on an improving path back to the complete network. So, for any of these pairwise stable networks, the resistance will be greater than 21. Thus, by Theorem 1 and Lemma 3 the unique stochastically stable network is the complete network.

While in the previous example, stochastic stability selected a single network from among many pairwise stable networks; the decisiveness of stochastic stability depends very much on the details of the setting. The next example shows a situation where all pairwise stable networks turn out to be stochastically stable, so that no selection is made.

Trading Example, Part II

Consider the trading example again with $n = 5$ and small $c > 0$. In particular, let c be small enough so that a player in a component of 4 benefits by adding link to an unconnected player. In this case, the three pairwise stable network structures are the star, line, and half-star (example, $\{12, 23, 35, 34\}$ is a half-star with player 3 in the center). All improving paths lead to a pairwise stable network and so there are no cycles. Permuting labels of players, there are 5 possible stars, 60 possible lines, and 60 possible half-stars.

Let us identify the set of stochastically stable networks. First, consider the resistance of the line $g = \{23, 31, 14, 45\}$. The half-star $\{23, 31, 14, 15\}$ has a distance of 1 from $im(g)$ as link 15 can be severed to get a sub-network of g. The star $\{13, 12, 15, 14\}$ has a distance of 1 from the half-star $\{23, 31, 14, 15\}$. The remaining 11 half-stars, with player 1 in the center, each have a distance of 1 from the star $\{12, 13, 14, 15\}$. The 12 half-stars of the form $\{i1, 1j, jk, jn\}$ each have a distance of 1 from a half-star with 1 in the center. The remaining 36 half-stars each have a distance of 1 from a half-star of the form $\{i1, 1j, jk, jn\}$. Each of the remaining 59 lines is a distance of 1 from a half-star. The remaining 4 stars are each a distance of 1 from a half-star. Thus, the total resistance of g is 124 which is the minimum possible resistance given the number of pairwise stable networks. Similarly it can be shown that every star, half-star and line has a resistance of 124, and so it follows from Theorem 1 and Lemma 3 that all pairwise stable networks are stochastically stable.

Remark. As seen in the co-author example, there is no guarantee that the stochastic process will lead to an efficient network. Theorem 1 in JW shows that there are situations where no strongly efficient network is pairwise stable (and where there are no cycles), even if one is free to select the allocation rule Y from a large class (i.e., those which are anonymous

and component balanced, as in JW). Thus, there are examples where no strongly efficient network is stochastically stable even when one is free to select the allocation rule Y. Nevertheless, one would hope that in situations where at least one strongly efficient network is pairwise stable, the stochastic process would select some strongly efficient network as one of the stochastically stable networks.

However, we show in Proposition 1 of the companion working paper (Jackson and Watts [23]) that given settings where at least one strongly efficient network is pairwise stable there does not exist an allocation rule from a large class (i.e., those which are anonymous and component balanced) which always allows at least one strongly efficient network to be stochastically stable. Thus there are settings that have a unique efficient network which is pairwise stable, while that efficient network is not stochastically stable for a wide class of allocation rules. So, *no matter how total value is reallocated* one cannot get the efficient network to be stochastically stable.

5. AN APPLICATION TO MATCHING MODELS

We now examine a particular application of the network formation model, namely network formation in matching problems, such as the Gale and Shapley [17] marriage problem and the hospital-intern or college admissions problem. (See Roth and Sotomayor [34] for a detailed overview of these problems.) This section is of independent interest as both an application of the dynamic network formation model, as well as an analysis of stochastic dynamics in the Gale and Shapley matching world.

In matching problems, there are restrictions on the set of admissible networks so that only some subset G of all possible networks is feasible. We provide definitions for two of the most extensively studied of these problems.

Marriage Problems

For the marriage problem, the set of players N is divided into a set of men, $M = \{m_1, ..., m_j\}$, and a set of women, $W = \{w_1, ..., w_k\}$. A network, g, is feasible if each woman is linked to at most one man, and each man is linked to at most one woman.

Thus, $G = \{g \mid ij \in g$ implies $ik \notin g$ for $k \neq j$, and for $ij \in g$, $i \in M$ implies $j \in W$ and $i \in W$ implies $j \in M\}$.

Let $m_i(g) = \{j \mid ij \in g\}$ denote the match of player i in the network g.

In a marriage problem, $v(g) = \sum_{i \in N} u_i(g)$, and $Y_i(v,g) = u_i(g)$, where for each i, $u_I: G \to \Re$ depends only on the match of i. That is, for each i u_i is such that $u_i(g) = u_i(g')$ whenever $m_i(g) = m_i(g')$.

Hospital-Intern and College Admissions Problems

For the hospital-intern (or college admissions) problem, the set of players N is divided into a set of hospitals, $H = \{h_1, ..., h_j\}$, and a set of interns, $I = \{i_1, ..., i_k\}$. A network, g, is feasible if each intern is linked to at most one hospital, and each hospital, h, is linked to at most q_h interns, where $q_h > 0$ is the quota for the hospital; thus each hospital has a fixed number of slots.

Thus, $G = \{g \mid \#m_i(g) \leqslant q_i$ for $i \in H$ and $\#m_i(g) \leqslant 1$ for $i \in I$, and for $ij \in g$, $i \in H$ implies $j \in I$ and $i \in I$ implies $j \in H\}$.

Again, $v = \sum_{i \in N} u_i(g)$, and $Y_i(v, g) = u_i(g)$, where for each i, $u_i: G \to \Re$ depends only on the matching of i. That is, for each i, u_i is such that $u_i(g) = u_i(g')$ whenever $m_i(g) = m_i(g')$.

We restrict attention to Hospital-Intern problems where preferences are *responsive* (see Roth and Sotomayor [34]). This is the condition that a hospital (college) has a ranking over interns (students) and an empty slot, such that preferences over subsets are consistent with the hospital's (college's) ranking.

Preferences are *responsive* if for each $h \in H$ there exists $y_h: N \cup \varnothing \to \Re$ such that if

(i) $m_h(g) = m_h(g') \cup i$, then $u_h(g) > u_h(g')$ if and only if $y_h(i) > y_h(\varnothing)$,

(ii) $m_h(g) = m_h(g')/i$, then $u_h(g) > u_h(g')$ if and only if $y_h(i) < y_h(\varnothing)$, and

(iii) $m_h(g) = m_h(g') \cup i/j$, then $u_h(g) > u_h(g')$ if and only if $y_h(i) > y_h(j)$.

Next we give the definition of a core stable network. A network g is core stable if there is no group of players who each prefer network g' to g and who can change the network from g to g' without the cooperation of the remaining players. In the marriage problem, it turns out that a network is core stable if and only if no player wants to sever his/her current link and no two players want to simultaneously sever their existing links and link with each other. This notion of core stability has been explored in great detail in matching models, beginning with Gale and Shapley (see Roth and Sotomayor [34] for an overview).

Core Stability

A network $g \in G$ is *core stable* if there does not exist any set of players A and $g' \in G$ such that:

(i) $Y_i(g') \geqslant Y_i(g)$ for all $i \in A$ (with at least one strict inequality),

(ii) if $ij \in g'$ but $ij \notin g$, then $i \in A$ and $j \in A$, and

(iii) if $ij \notin g'$ but $ij \in g$, then either $i \in A$ and/or $j \in A$.

EXAMPLE (CONTRAST OF CORE STABILITY AND PAIRWISE STABILITY[10]). Consider a marriage market with two men and two women. Preferences are

$$m_1: w_1, w_2$$
$$m_2: w_2, w_1$$
$$w_1: m_1, m_2$$
$$w_2: m_2, m_1.$$

The above table can be read as follows: m_1's first choice for a spouse is w_1 and his second choice is w_2. The remaining preferences can be read in a similar fashion.

The unique core stable matching is $\{m_1w_1, m_2w_2\}$. However, both $\{m_1w_1, m_2w_2\}$ and $\{m_1w_2, m_2w_1\}$ are pairwise stable. Also, both networks are stochastically stable. From either network, two links need to be severed to get to an improving path to the other network.

In the above example, the pairwise stability of $\{m_1w_2, m_2w_1\}$ suggests that the definition of improving path might not be appropriate in the matching setting where there are restrictions on the number of links that a player may have. In particular, if m_1 and w_1 were to meet, they would like to join but need to sever their existing links in order to form a new link. This suggests allowing players to sever links at the same time as they add them. This is permitted under the following variation on the notion of improving path.

Simultaneous Improving Paths

A *simultaneous improving path*, is a sequence of networks $g_0, ..., g_K$ in G such that if g' follows g in the sequence then either

(i) $g' = g - ij$ and either $Y_i(g') > Y_i(g)$ or $Y_j(g') > Y_j(g)$, or

(ii) $g' \in G$ and $g' \in \{g+ij-ik, g+ij-ik-jm, g+ij, g+ij-jm\}$ where $ij \notin g$ and $Y_i(g') \geqslant Y_i(g)$ and $Y_j(g') \geqslant Y_j(g)$ (with one inequality holding strictly).

---

[10] Gale and Shapley have a notion that they call pairwise stability, which corresponds to the definition of core stability when A is restricted to have no more than two members. In these matching problems, that definition of pairwise stability corresponds with core stability. We use the term pairwise stability in the sense of JW, as indicating the lack of improving paths leaving a network.

Improving paths are a subset of simultaneous improving paths. Here the simultaneity refers to the fact that a player may make several changes at once: a player may both sever an existing link and add a new one. Note that the above definition can be altered so that the players, when adding a new link, sever the minimum necessary number of links in order to add the given link.

In the context of the marriage and hospital-intern problems, a core stable network g is any network from which there is no simultaneous improving path leaving g.

As illustrated in the following example, cycles can exist with the notion of simultaneous improving path. (There are no cycles in the marriage problem under the notion of improving path, as a cycle requires some simultaneous changes to be made.)

EXAMPLE (EXISTENCE OF A CYCLE IN A MARRIAGE PROBLEM). Consider a marriage problem with two men and two women, where preferences are as follows:

$$m_1: w_1, w_2$$
$$m_2: w_2, w_1$$
$$w_1: m_2, m_1$$
$$w_2: m_1, m_2.$$

There exists a cycle under the definition of simultaneous improving path: $\{m_1w_1\}$ to $\{m_2w_1\}$ to $\{m_2w_2\}$ to $\{m_1w_2\}$ to $\{m_1w_1\}$.

The existence of a marriage cycle was first demonstrated by Knuth [28] and is also discussed in Roth and Vande Vate [35]. While there are cycles in the marriage problem there are no closed cycles. That follows from the main result in Roth and Vande Vate [35], which in the network language says that from each network there is a simultaneous improving path leading to some core stable network. This implies (from Lemma 2 here) that all stochastically stable networks will be core stable networks. So a question arises as to which core stable networks are stochastically stable.

It is known (since Gale and Shapley [17]) that the set of core stable matchings has nice properties. There is a man-optimal core stable matching which is the unanimous favorite of all men out of the set core stable matchings, and similarly there is a women-optimal matching. One core stable matching is preferred by all men to another if and only if all women have the reverse preference ordering over the two matchings. We want to know if the men and women optimal matchings will be stochastically stable, and if so if they will be the only stochastically stable matchings. To

answer these questions, let us describe a stochastic process that corresponds to the notion of simultaneous improving path.

A Simultaneous Stochastic Process

At each time a pair of players is randomly identified (where each pair has a positive probability of being identified that is i.i.d. across time). If the link is already in the network, then the decision is whether to sever it; otherwise, the two players are allowed to form the link and at the same time sever up to one existing link each. Their actions are constrained to lead to a feasible g in G, so in some cases they must sever an existing link in order to add the new link. The players involved act myopically, adding the link (with corresponding severances) if it makes each at least as well off and one strictly better off, and severing the link if its deletion makes either player better off. After the action is taken, there is some small probability that a tremble occurs and the link is deleted if it is present.

There are many possible interpretations for such trembles in the marriage model. Consider two people who should get married but do not, due to a tremble. Such a tremble could represent a fear of change on the part of one of the players. Alternatively, consider a couple who are married and should stay married, but divorce due to a tremble. Here such a tremble could represent a mid-life crisis, a desire to exit a rut, or experimentation by a spouse to see if divorce will put them on a better improving path.

Note that we consider only trembles that delete the given link; thus one does not have to worry about the constraints imposed by feasibility, which might bind in the case of adding a link. This turns out to be irrelevant due to the natural tendency towards the addition of beneficial links, the restrictions on numbers of links, and the absence of externality effects.[11]

S-Stochastically Stable Networks

The set of networks that is the support of the limiting stationary distribution (of the simultaneous improving process) is the set of *S-stochastically stable networks*.

THEOREM 2. *Consider the marriage problem where players' preferences are strict (and players are allowed to prefer staying alone to being in some matches). The set of S-stochastically stable networks coincides with the set of core stable networks.*

---

[11] What is important is that it is impossible to have a situation where someone hesitates to add a link, that once added would lead to changes in incentives for the subsequent networks. In the marriage model, and the college admissions model with responsive preferences, if a link were added that did not benefit the players involved, it would simply be deleted at the next opportunity.

The results of Roth and Vande Vate [35] imply that if one starts at an empty network, then a (mutationless) random meeting process could result in any core stable network, depending on the order in which players happen to meet. Theorem 2 strengthens our understanding of the formation of matchings to also show that every core stable network is actually s-stochastically stable. Thus, there are no core stable networks that are more fragile than others in the face of mutations that might randomly break links.

The proof of Theorem 2 is in the Appendix. An outline is as follows. By Roth and Vande Vate [35], there are no closed cycles in the marriage problem and so every network which is not core stable is on a simultaneous improving path to a core stable network. We show that each core stable network that is not man optimal is only one link away from a simultaneous improving path that leads to a core stable network with strictly more men matched to their man-optimal mates. This fact is utilized to build restricted g-trees for core stable networks. Each such restricted g-tree has a resistance of $K - 1$, where K is the number of core stable networks. Thus, we establish that every core stable network has the same resistance and so all are s-stochastically stable.

The reason that this result is somewhat surprising is that the example below shows that the set of pairwise stable networks does not coincide with the set of stochastically stable networks. So, as we are making parallel changes in these definitions one might not expect the set of core stable and S-stochastically stable networks to always coincide either.

EXAMPLE (CONTRAST WITH PAIRWISE STABLE AND STOCHASTICALLY STABLE NETWORKS). Consider a marriage problem with two men and two women. Preferences are

$$m_1: w_1, w_2$$
$$m_2: w_2, \text{alone}$$
$$w_1: m_1, \text{alone}$$
$$w_2: m_2, m_1$$

Thus m_2 prefers being alone to being matched with w_1. Here $\{m_1 w_1, m_2 w_2\}$ and $\{m_1 w_2\}$ are both pairwise stable, but only $\{m_1 w_1, m_2 w_2\}$ is stochastically stable. This follows since it takes two mutations to get from $\{m_1 w_1, m_2 w_2\}$ to $\{m_1 w_2\}$, but only one mutation to go the other way.

Theorem 2 extends to the college admissions (hospital-intern) problem, as we now state.

COROLLARY 2. *Consider the college admissions (hospital-intern) problem with strict and responsive preferences. The set of S-stochastically stable networks coincides the set of core-stable networks.*

Proof. This follows from the proof Theorem 2, by considering a related marriage market where each college is replicated q_h times. (For more on the relationship between the college admissions and marriage problems see Chapter 5 in Roth and Sotomayor [34].)

Responsiveness is critical to Corollary 2, as shown in the following example.

EXAMPLE (ROLE OF RESPONSIVENESS). Consider a college admissions problem with 1 college and 3 students. The set of feasible networks allows only for links that involve the college and a student. Here the term ''full network'' refers to the network with all three feasible links. Assume that each student prefers being in the college to being out. The college prefers to have all three students to none, but prefers none to having any proper subset of students. Thus, the college's preferences fail to satisfy responsiveness.

There are two pairwise stable networks: one that has each student linked to the college and another where none are linked to the college. The only network that is ''core-stable'' is the full network.

Under stochastic stability the resistance of the full network is 2 (to get from the null network to the full network, at least two links must exist for the third to be added). However, the resistance of the null network is 1 (delete one link from the full network and it lies on an improving path back to the null network). Thus, only the null network (which is Pareto inefficient and out of the core) is stochastically stable. Similarly, only the null network is S-stochastically stable.

6. CONCLUSION

As we have emphasized, there are many modeling choices that have been made with respect to the dynamic process and the definition of improving path. One modeling decision that deserves further attention is the myopic behavior of players in the definition of improving path. It would be natural to have forward-looking players in situations with a small number of players who are well-informed about the incentives of other players, the allocations and valuations, and who care about the future. A interesting problem for future research, is to develop an appropriate definition of improving path for forward looking players and find the set of stochastically stable networks for these forward-looking improving paths.

To get an idea of what types of differences might emerge from the myopic model, consider the following three-player example. Let players have payoffs which meet the following inequalities: $Y_1(12) < 0 < Y_1(12, 13) < Y_1(13) < Y_1(12, 13, 23)$; $Y_2(23) < 0 < Y_2(12, 23) < Y_2(12) < Y_2(12, 13, 23)$; $Y_3(13) < 0 < Y_3(13, 23) < Y_3(23) < Y_3(12, 13, 23)$; $Y_i(\{\}) = 0$ and $0 < Y_i(ij, jk) < Y_i(12, 13, 23)$, for all i. Thus $\{12, 13, 23\}$ is the unique efficient network. If players are myopic, then there are two pairwise stable networks, $\{12, 13, 23\}$ and $\{\}$. From the network $\{12, 13, 23\}$ one link must be severed to get to an improving path to $\{\}$. From the network $\{\}$ one link must be added to get to an improving path to $\{12, 13, 23\}$. Thus, by Theorem 1, both networks are stochastically stable.

Now assume that players are non-myopic and care about future payoffs. Assume also that each player knows every other player's payoff function. Suppose the players are currently in the network $\{12\}$ and suppose that the link $\{12\}$ is identified and can thus be severed by either player 1 or 2. If player 1 were myopic then he would sever the link. However, if player 1 values the future enough, then he will decide not to sever the link given that he should have the opportunity to add the link 13 in the future which will then eventually lead to the full network (as then the remaining link will form when the opportunity arises), or else players 2 and 3 will add the link 23 which will eventually lead to the full network. Such reasoning relies on j in the network $\{ij, jk\}$ having the incentive to maintain both links with the expectation that the link ik will soon be formed, which is better for j than severing the link given sufficient patience. This reasoning provides forward looking improving paths from any $\{ij\}$ should lead to $\{12, 13, 23\}$, and similarly from any $\{ij, jk\}$, and so $\{12, 13, 23\}$ emerges as the unique stochastically stable network. Our reasoning here is only suggestive and heuristic as one needs to carefully define forward looking improving path to capture these ideas. It becomes clear even in this simple example that the definition of such improving paths depends on what players expect improving paths from other networks to look like, and for instance relies on the belief that from some network $\{ij, jk\}$ the only path will lead to the fully efficient network and not back to some single link network. It appears that such a definition needs to be inductive or self-referential, and so carefully developing such a notion presents a challenge for future research.

APPENDIX

Proof of Theorem 1. The proof is an application of a theorem from Young [43].

Before stating Young's [43] theorem, we give the following definitions. Consider a stationary Markov process on a finite state space X with transition matrix P.

A set of *mutations of P* is a range $(0, a]$ and a stationary Markov process on X with transition matrix $P(\varepsilon)$ for each ε in $(0, a]$, such that (i) $P(\varepsilon)$ is aperiodic and irreducible for each ε in $(0, a]$, (ii) $P(\varepsilon) \to P$, and (iii) $P(\varepsilon)_{xy} > 0$ implies that there exists $r \geqslant 0$ such that $0 < \lim \varepsilon^r P(\varepsilon)_{xy} < \infty$.

The number r in (iii) above is the *resistance* of the transition from state x to y. There is a path from x to z of *zero resistance* if there is a sequence of states starting with x and ending with z such that the transition from each state to the next state in the sequence is of zero resistance. Note that from (ii) and (iii), this implies that if there is a path from x to z of zero resistance, then the nth order transition probability associated with P of x to z is positive for some n.

The *recurrent communication classes* of P, denoted $X_1, ..., X_J$, are disjoint subsets of states such that (i) from each state there exists a path of zero resistance leading to a state in at least one recurrent communication class, (ii) any two states in the same recurrent communication class are connected by a path of zero resistance (in both directions), and (iii) for any recurrent communication class X_j and states x in X_j and y not in X_j such that $P(\varepsilon)_{xy} > 0$, the resistance of the transition from x to y is positive.

For two communication classes X_i and X_j, since each $P(\varepsilon)$ is irreducible, it follows that there is a sequence of states $x_1, ..., x_k$ with x_1 in X_i and x_k in X_j such that the resistance of transition from x_k to x_{k+1} is defined by (iii) and finite. Denote this by $r(x_k, x_{k+1})$. Let the resistance of transition from X_i to X_j be the minimum over all such sequences of $\sum_1^{K-1} r(x_k, x_{k+1})$, and denote it by $r(X_i, X_j)$.

Given a recurrent communication class X_i, an *i-tree* is a directed graph with a vertex for each communication class and a unique directed path leading from each class j ($\neq i$) to i. The *stochastic potential* of a recurrent communication class X_j is then defined by finding an i-tree that minimizes the summed resistance over directed edges, and setting the stochastic potential equal to that summed resistance.

Also, given any state x, an *x-tree* is a directed graph with a vertex for each state and a unique directed path leading from each state $y(\neq x)$ to x. The resistance of x is then defined by finding an x-tree that minimizes the summed resistance over directed edges.

The following theorem is a combination of Theorem 4 and Lemmas 1 and 2 in Young:

THEOREM (Young [43]).   *Let P be the transition matrix associated with a stationary Markov process on a finite state space with a set of mutations $\{P(\varepsilon)\}$ and with corresponding (unique) stationary distributions $\{m(\varepsilon)\}$. Then $m(\varepsilon)$ converges to a stationary distribution m of P, and a state x has $m_x > 0$ if and only if x is in a recurrent communication class of P which has a minimal stochastic potential. This is equivalent to x having minimum resistance.*

To apply this to our setting, note that under the dynamic process, each pairwise stable network and closed cycle is a recurrent communication class of the corresponding process P (and these are exactly the recurrent communication classes of P). Next, the transition from any network g to an adjacent one g' has probability on the order of ε if g is not in $\operatorname{im}(g')$ (and thus has resistance 1), and is of the order of 1 otherwise (and thus has resistance 0). ∎

Proof of Lemma 2. Let us start by showing that if $g' \in \operatorname{im}(g)$ and $g \notin \operatorname{im}(g')$, then $r(g) \leqslant r(g')$. Consider the g'-tree relative to which $r(g')$ is calculated. Construct a g-tree by starting with the g'-tree and directing an edge from g' to g, and erasing the edge that led away from g. Since $g' \in \operatorname{im}(g)$, it follows that the added edge has 0 resistance, so this g-tree has a resistance of no more than $r(g')$. Thus, $r(g) \leqslant r(g')$.

Moving out of the order of the statement of Lemma 2, let us next verify that if one network, g, in a closed cycle is stochastically stable then any other network, g', in the same closed cycle is also stochastically stable. To see this simply start with a g-tree relative to which $r(g)$ is calculated, and switch the places of g' and g. Since g and g' are both connected to each other by improving paths, $r(g'', g) = r(g'', g')$ and $r(g', g'') = r(g, g'')$ for any g''. Thus, the resistance will be unchanged, and so g' must also be stochastically stable.

Next, let us show that if $g' \in \operatorname{im}(g)$ and $g \notin \operatorname{im}(g')$ and g is pairwise stable, then $r(g) < r(g')$. Again, construct a g-tree by starting with the g'-tree relative to which $r(g')$ is calculated, directing an edge from g' to g, and erasing the edge that led away from g. Note that if g is pairwise stable, then erasing the edge that led away from g saved at least 1 unit of resistance. Since the added edge has 0 resistance, it follows that $r(g) < r(g')$. This argument is extended to the case where g is in a closed cycle, as follows. If the edge that led away from g in the original g'-tree had a positive resistance, then the same argument as above works. If not, then it must be that the edge leading away from g in the original g'-tree pointed to some network in the closed cycle containing g. In that case, there must have been some g'' in the closed cycle containing g that had an edge exiting the closed cycle. Construct the g'' tree from the g' tree as described above. By the definition of closed cycle it must be that $g' \in \operatorname{im}(g'')$ and $g'' \notin \operatorname{im}(g')$. Thus the above argument holds, establishing that g'' is stochastically stable. So, it follows that g must also be stochastically stable as it is part of the same closed cycle. ∎

Proof of Theorem 2. Consider the following results about the core stable networks (see Roth and Sotomayor [34]):

THEOREM A (Gale and Shapley [17], Knuth [28]). *There exists a man-optimal core stable network that all men weakly prefer to any other core stable network. Similarly, there exists a woman-optimal core stable network. If men all weakly prefer one core stable network g to another g', then the women all weakly prefer g' to g.*

THEOREM B (2.22 in Roth and Sotomayor [34]). *The set of players with no links is the same in all core stable networks.*

THEOREM C (Roth and Vande Vate [35]). *From any g that is not core stable, there exists a simultaneous improving path that leads to a core stable network g'.*

We use these in proving the following series of claims.

CLAIM 1. *There exist no closed cycles.*

This follows from Theorem C.

CLAIM 2. *If there is a singleton core stable network, then it is the unique S-stochastically stable network.*

This follows from Claim 1.

CLAIM 3. *Any core stable network, g, that is not man-optimal is one link away from an improving path that leads to a core stable network, g', where the number of men married to their man-optimal mates is larger in g' than in g.*

Find a man in g who is not linked to his man-optimal mate and sever that link. (By Theorem B, that man must be linked under g.) Call this man m_i and call this man's former mate w_i. By Theorem B, w_i must have a man-optimal mate who is linked. Let this man, m_j, link with w_i severing his existing link. (By Theorem A and strict preferences, m_j must prefer w_i to his current mate.) Since w_i is unlinked she will benefit from forming a link with m_j (as she must prefer him to being alone given that she is linked to him in the man-optimal core stable network). Now m_j's former mate is unmarried, let her man-optimal mate sever his link and link to her. This process can be continued as long as there is still a woman (and thus man) who is disconnected and would be linked in a core stable network. (Recall from Theorem B that the set of players involved in severing old links and adding new ones is only the set of players who are connected in core stable networks.) When this process ends, it must be that m_i proposes to his man-optimal mate. Let g'' be the network reached. Note that this process

follows a simultaneous improving path and that any man who has a different mate under g and g'' must have his man-optimal mate under g''. If g'' is core stable then stop. Then by construction the number of men married to their man-optimal mates is larger in g'' than in g. If not, then there exists a couple, say w_k and m_k, who would both prefer to link and sever their existing links under g''. Let them do so. (Note that w_k was originally matched to her man-optimal mate or, by Theorem A, to someone she prefers to her man-optimal mate. Thus by strict preferences w_k must strictly prefer m_k to her man-optimal mate. So, m_k must strictly prefer his man-optimal mate to w_k; otherwise the man-optimal network is not core stable since w_k and m_k would both prefer to sever their ties and marry. Thus if m_k is linked to w_k and m_k has the chance to propose to his man-optimal mate he will.) Now m_k's former mate is unmarried. Let her man-optimal mate propose. Continue as described above, which continues on a simultaneous improving path. Since there are no cycles (Claim 1), this process ends at a core stable network g'. By construction the number of men married to their man-optimal mates is larger under g' than under g.

CLAIM 4. *If there are $k > 1$ core stable networks, then the resistance of the man optimal network is $k - 1$.*

This follows from Claims 1 and 3 by constructing a restricted man-optimal tree by directing each core-stable g that is not man-optimal to a core stable g' as defined in Claim 3. Note that there must be some network that connects to the man optimal network.

CLAIM 5. *If there are $k > 1$ core stable networks, then the resistance of any core-stable network is $k - 1$.*

By Claim 4, we know this for the man-optimal (and correspondingly the woman optimal) network. Consider some other core stable g. We know that this is man-optimal under a change of preferences where each man's preferences are changed so that any woman that he preferred to his mate under g is now unacceptable. Then, by the logic of Claim 4, there exists a restricted g tree where the resistance of each edge is 1 among the core stable networks of the modified preferences. (The resistance of each of these edges is the same under the original preferences since no one is ever linked to an unacceptable mate). Next we show that this restricted g-tree is identical to a restricted g-tree on the set of core stable networks that all men find weakly less preferred to g under the original preferences. This statement is true if the set of core stable networks that all men find weakly less preferred to g under the original preferences (call this set G_B) equals the set of core stable networks of the modified preferences (call this set G_A). Next we show that $G_A = G_B$. It is clear that $G_B \subset G_A$ since if a network in

G_B is not blocked under the original preferences it is not blocked under the modified preferences. It is also obvious that if $g' \in G_A$ and g' was core stable under the original preferences then g' must be weakly less preferred to g by all men under the original preferences. All that is left is to show that if g' is core stable under the modified preferences then g' is core stable under the original preferences. First of all, since the set of acceptable mates under the original preferences includes all those that are acceptable under the modified preferences, no individual finds his or her g' mate unacceptable under the original preferences. All that is left is to consider blocking by a pair of a man and a woman (as we know from Gale and Shapley [17] that we need only examine blocking coalitions of two or fewer individuals). As only the man's preferences have been altered, if g' is blocked under the original preferences by such a pair (m_k, w_k) but not under the modified preferences, then it must be that m_k strictly prefers w_k to his mate in g', under his original preferences but not under the modified ones, while w_k is indifferent between her mate in g' and m_k. We show that this cannot be the case by showing that in such a case w_k must strictly prefer her mate in g' to m_k. Consider two cases. First, suppose m_k weakly prefers his mate in g to w_k under the original preferences. Then both w_k and $m_k(g')$ are no better under the original preferences than $m_k(g)$, and hence m_k's ranking of these are unchanged in the modified preferences, which contradicts the above. Second, suppose m_k strictly prefers w_k to $m_k(g)$. Since g was core stable under the original preferences, it must be that w_k strictly prefers $w_k(g)$ to m_k. The modified preferences are constructed so that w_k is the most preferred mate of $w_k(g)$. By the core stability of g' under the modified preferences, it must be that w_k weakly prefers $w_k(g')$ to $w_k(g)$. Therefore, w_k weakly prefers $w_k(g')$ to $w_k(g)$ and so w_k strictly prefers $w_k(g')$ to m_k.

Thus, we can find a restricted g-tree on the set of core stable networks that all men find weakly less preferred to g under the original preferences, where the resistance of each edge is 1. Similarly we can find a restricted g-tree on the set of core stable networks that all men find weakly preferred to g (using Theorem A and working to the woman-optimal network) where the resistance of each edge is 1. The remaining core stable networks that we have not yet described edges for must not be uniformly ranked by men relative to g. For those networks keep the same directed edges as in the tree in Claim 4. Since from the proof of claim 3 those edges always point to networks that all men weakly prefer and some strictly prefer to the starting network, these directed edges cannot cycle and must eventually lead to a core stable network g' that is weakly preferred by all men to g. Since there is a directed path from g' to g, this defines a restricted g-tree and it has resistance $k-1$.

Claims 1, 2, and 5 establish the theorem. ∎

REFERENCES

1. R. Aumann and R. Myerson, Endogenous formation of links between players and coalitions: An application of the Shapley value, *in* "The Shapley Value" (A. Roth, Ed.), pp. 175–191, Cambridge Univ. Press, Cambridge, UK, 1988.

2. V. Bala and S. Goyal, Self-organization in communication networks, *Econometrica* **68** (2000), 1181–1229.

3. K. Basu and J. Foster, On measuring literacy, Working Paper 96-W02, Vanderbilt University, 1996.

4. A. Bell, Bilateral trading on a network: Convergence and optimality results, mimeo, Vanderbilt University, 1996.

5. J. Bergin and B. Lipman, Evolution with state-dependent mutations, *Econometrica* **64** (1996), 943–956.

6. S. Boorman, A combinatorial optimization model for transmission of job information through contact networks, *Bell J. Econ.* **6** (1975), 216–249.

7. E. Droste, R. Gilles, and C. Johnson, Evolution of conventions in endogenous social networks, mimeo, Virginia Tech, 2000.

8. B. Dutta and M. O. Jackson, The stability and efficiency of directed communication networks, *Rev. Econ. Design* **5** (2000), 251–272.

9. B. Dutta and M. O. Jackson, On the formation of networks and groups, introductory chapter, *in* "Models of the Formation of Networks and Groups" (B. Dutta and M. O. Jackson, Eds.), Springer-Verlag, Heidelberg, 2001, in press.

10. B. Dutta and S. Mutuswami, Stable networks, *J. Econ. Theory* **76** (1997), 322–344.

11. G. Ellison, Learning, local interaction, and coordination, *Econometrica* **61** (1993), 1047–1071.

12. G. Ellison and D. Fudenberg, Word-of-mouth communication and social learning, *Quart. J. Econ.* **110** (1995), 93–126.

13. J. Ely, Local conventions, *Adv. Theoret. Econ.* **2** (2002).

14. M. Fafchamps and S. Lund, Risk-sharing networks in rural Philippines, mimeo, Stanford University, 1997.

15. D. Foster and P. Young, Cooperation in the short and in the long run, *Games Econ. Behavior* **3** (1991), 145–156.

16. M. Freidlin and A. Wentzell, *in* "Random Perturbations of Dynamical Systems," pp. 161–211, Springer-Verlag, New York, 1984.

17. D. Gale and L. Shapley, College admissions and the stability of marriage, *Amer. Math. Monthly* **69** (1962), 9–15.

18. E. Glaeser, B. Sacerdote, and J. Scheinkman, Crime and social interactions, *Quart. J. Econ.* **111** (1996), 507–548.

19. S. Goyal and F. Vega-Redondo, Learning, network formation and coordination, mimeo, Erasmus University, 1999.

20. N. Hummon, A utility analysis of dynamic social networks, mimeo, University of Pittsburgh, 1998.

21. D. Iacobucci and N. Hopkins, Modeling dyadic interactions and networks in marketing, *J. Market. Res.* **29** (1992), 5–17.

22. M. O. Jackson, The stability and efficiency of economic and social networks, *in* "Models of the Formation of Networks and Groups" (B. Dutta and M. O. Jackson, Eds.), Springer-Verlag, Heidelberg, 2001, in press.

23. M. O. Jackson and A. Watts, The evolution of social and economic networks, working paper version, California Institute of Technology WP1044, 1998.

24. M. O. Jackson and A. Watts, On the formation of interaction networks in social coordination games, *Games Econ. Behav.*, in press.

25. M. O. Jackson and A. Wolinsky, A strategic model of social and economic networks, *J. Econ. Theory* **71** (1996), 44–74.
26. M. Kandori, G. Mailath, and R. Rob, Learning, mutation, and long run equilibria in games, *Econometrica* **61** (1993), 29–56.
27. A. Kirman, C. Oddou, and S. Weber, Stochastic communication and coalition formation, *Econometrica* **54** (1986), 129–138.
28. D. Knuth, "Mariages Stables," Les Presses de l'Université de Montreal, Montreal, 1976.
29. R. Maxfield, General equilibrium and the theory of directed graphs, *J. Math. Econ.* **27** (1997), 23–51.
30. D. Monderer, D. Samet, and L. Shapley, Weighted values and the core, *Internat. J. Game Theory* **21** (1992), 27–39.
31. J. Montgomery, Social networks and labor market outcomes, *Amer. Econ. Rev.* **81** (1991), 1408–1418.
32. R. Myerson, Graphs and cooperation in games, *Math. Oper. Res.* **2** (1977), 225–229.
33. A. Ritzberger and J. Weibull, Evolutionary selection in normal-form games, *Econometrica* **63** (1995), 1371–1399.
34. A. Roth and M. Sotomayor, "Two Sided Matching," Econometric Society Monographs 18, Cambridge Univ. Press, Cambridge, UK, 1989.
35. A. Roth and J. Van de Vate, Random paths to stability in two-sided matching, *Econometrica* **58** (1990), 1475–1480.
36. M. Slikker and A. van den Nouweland, A one-stage model of link formation and payoff division, *Games Econ. Behav.* **34** (2001), 153–175.
37. G. Topa, Social interactions, local spillovers and unemployment, *Rev. Econ. Stud.* **68** (2001), 261–295.
38. L. Tesfatsion, A trade network game with endogenous partner selection, *in* "Computational Approaches to Economic Problems" (H. Amman, *et al.*, Eds.), pp. 249–269, Kluwer Academic, Dordrecht, 1997.
39. L. Tesfatsion, Gale–Shapley matching in an evolutionary trade network game, Iowa State Unversity Economic Report 43, 1998.
40. A. Watts, A dynamic model of network formation, *Games Econ. Behav.* **34** (2001), 331–341.
41. P. Wang and Q. Wen, Network bargaining, mimeo, Penn State University, 1998.
42. G. Weisbuch, A. Kirman, and D. Herreiner, Market organization, *Econ. J.* **110** (2000), 411–436.
43. H. P. Young, The evolution of conventions, *Econometrica* **61** (1993), 57–84.
44. H. P. Young, "Individual Strategy and Social Structure," Princeton Univ. Press, Princeton, NJ, 1998.