On the Formation of Interaction Networks in Social Coordination Games[1]

Matthew O. Jackson and Alison Watts

This Draft: October 28, 1999

Abstract: There are many situations where two interacting individuals can benefit from coordinating their actions. We examine the endogenous choice of partners in such social coordination games and the implications for resulting play. We model the interaction pattern as a network where individuals periodically have the discretion to add or sever links to other players. A player chooses whether to add or sever a link based on the (prospective) partner's past behavior. With such endogenous interaction patterns we see multiple stochastically stable states of play, including some that involve play of equilibria in the coordination game that are neither efficient nor risk dominant.

Journal of Economic Literature Classification Numbers: A14, D20, J00.

Humanities and Social Sciences 228-77,
Caltech, Pasadena, CA 91125, USA
(jacksonm@hss.caltech.edu)

Department of Economics,
Box 1819, Station B, Vanderbilt University, Nashville, TN 37235, USA
(alison.watts@vanderbilt.edu)

---

1. Introduction

There are many situations where two interacting individuals can benefit from coordinating their actions. Examples where coordination is important include research and development partnerships, joint production, political and trade alliances, as well as the choice of compatible technologies or conventions such as the choice of a software or language.[2] For instance, a person may select a long distance service or internet provider keeping in mind which is best at facilitating communication with friends and family. Often there is an advantage to having individuals coordinate on selecting the same or compatible systems, with examples including long distance telephone service with MCI's friends and family plan and internet service with AOL's buddy list system. In many such situations, individuals select a strategy (e.g., technology) that they then use in interactions with many other individuals. Such situations are often characterized by multiple equilibria corresponding to coordination on different technologies or strategies, where the equilibria may be Pareto ranked. The welfare implications of such a ranking provide an obvious importance to understanding what behavior might be predicted and which factors determine whether efficient coordination is attained.

In the context of symmetric 2x2 coordination games, Kandori, Mailath and Rob (1993) and Young (1993) have shown that populations of individuals, who follow a large class of strategy adjustment procedures (including myopic best response) and are subjected to small random perturbations in their strategy choices, tend in the long run to coordinate on risk-dominant strategies as defined by Harsanyi and Selten (1988). Thus the risk-dominant equilibrium is selected from among the set of strict Nash equilibria, even if the risk dominant equilibrium is the inefficient equilibrium. This result has a natural and simple intuition: the basin of attraction of the risk-dominant equilibrium is larger than that of the non risk-dominant equilibrium. In particular, more than half of the population must be playing the non-risk dominant strategy (which may be the efficient strategy) for that strategy to be a best response. So, if everyone initially plays the risk dominant strategy, then more than half the population must randomly change to the efficient strategy for the dynamics to move towards the efficient equilibrium, while if everyone initially plays the efficient strategy, then less than half the population needs to randomly change to the risk dominant strategy for the dynamics to move to the risk dominant solution. In the long run this leads to a higher probability that in any given period players will be playing the risk dominant equilibrium, and the risk dominant solution is the stochastically stable convention, in the sense coined by Foster and Young (1990).

In the Kandori, Mailath, and Rob (1993) and Young (1993) models each player plays against every other player in the population (or equivalently faces each other player with an equal probability). While this is plausible in some contexts, many situations where social coordination is an issue involve more specific interaction patterns. These results are robust to some alternative interaction structures, for instance where individuals interact according to fixed neighborhood

---

[2] For overviews and more detailed discussion of other relevant examples and applications, see Young (1998) and the *Journal of Economic Perspectives* symposium on network externalities (Katz and Shapiro (1994), Besen and Farrell (1994) and Liebowitz and Margolis (1994)).

structures as shown by Ellison (1993) (see also Young (1998)). Thus we are left with the somewhat pessimistic result that a society can be expected in the long run[3] to coordinate on the risk-dominant equilibrium, even in cases where it is inefficient and so in society's common interest to do otherwise.   However, we show in an example here that for some fixed network structures it is possible to have multiple stochastically stable states, with some involving play of equilibria that are not risk dominant. This example points out the important role of the particular network structure in determining stochastic stability, and underlines the importance of endogenizing the network structure.

Ely (1998) (see also Mailath, Samuelson, and Shaked (1997)) considers models where the interaction structure is endogenized by locational choices. Individuals select a location at which to reside and then interact according to a pattern governed by that location and the location of other individuals. Conditions are given under which the efficient equilibrium is the one that is reached by a society even when it is not risk dominant. In the model of Ely (1998) if some individual randomly moves to an unoccupied location and plays the efficient strategy, then other individuals would like to move to that location and play the efficient strategy rather than staying at a location where they play the inefficient strategy. This leads to efficient play. While this result is encouraging in showing how endogenizing the interaction pattern can lead to efficiency, this result depends on the locational aspect of the interaction patterns.[4]   In particular, in changing locations agents are essentially allowed to sever all old ties, form new ties, and switch technologies simultaneously. While there are situations where location is the major factor in determining with whom an individual interacts, in many applications individuals choose with whom they interact in a more discretionary manner, not having to completely uproot to form new relationships.

Our model is not a location model, but rather one where players choose their interaction patterns on an individual-by-individual basis.   We model the interaction pattern as a network where individuals periodically have the discretion to add or sever links to other players.   Players choose whether to add or sever links based on their (prospective) partner's past behavior. With such endogenous interaction patterns there exist multiple stochastically stable states of play, including some that involve play of equilibria in the coordination game that are *neither* efficient nor risk dominant. Thus, it is possible to have inefficient play be stochastically stable *even when* the efficient strategy is risk dominant, in a model where individuals completely control with whom they interact.

2. Network Interactions and Stochastic Stability

Networks

---

[3] Ellison (1993) and Young (1998) also present interesting results on how the time associated with the ``long run'' varies depending on the network structure.
[4] Mailath, Samuelson and Shaked discuss the importance of having locations where individuals can isolate themselves from others in order to obtain efficiency. In their model, individuals cannot necessarily escape being matched with undesired opponents, and the ability to isolate is important in determining whether efficiency is attained.

The set N={1,...,n} is a finite set of individual players. These may be people, firms, computers, countries, or other relevant participants. We assume that n>2, as networks among two players are not very interesting.

The network relations among these individuals are represented by a graph whose nodes or vertices are the individuals, and whose links or edges are connections between the individuals. The complete network or graph, denoted g^N, is the set of all subsets of N of size 2. The set of all possible networks on N is $\{g \mid g \subset g^N\}$.[5] Let ij denote the subset of N containing i and j and is referred to as the *link ij*. The interpretation of ij∈g is that in the network g, individuals i and j are linked.

Note that i is linked to j if and only if j is linked to i, so that we study non-directed networks where interaction requires mutual consent.

Before endogenizing the network, let us discuss how interaction occurs given a fixed network.

Coordination Games played on a Fixed Network

Consider the following situation as described by Young (1998). A population of n players plays a game repeatedly. Players are located on a fixed network g. In each period t, a player i chooses an action $a^t_i \in \{A,B\}$, and then receives a payoff which is

$$u_i(g;a^t_1,...,a^t_n) = \sum_{i \neq j} \pi_{ij}(g)v_i(a^t_i,a^t_j) \qquad (1)$$

where $\pi_{ij}(g)$ is a factor that weighs the relative importance of the interaction between i and j. For our purposes, set $\pi_{ij}(g)=1$ if ij∈g and $\pi_{ij}(g)=0$ if ij ∉ g. Thus, each player interacts only with the players that he is directly connected to.

The following matrix describes the payoff function v_i. The matrix lists only the payoff to player i, with the payoff to player j being symmetrically determined.

		Player j	
		A	B
	A	a	c
Player i			
	B	d	b

Let a>d and b>c, so that the game is a coordination game, with two pure strategy equilibria, A,A and B,B. Let (a-d)>(b-c), so that A,A is the risk-dominant equilibrium in the sense of Harsanyi and Selten (1988). This equilibrium has the property that each player is choosing a strategy that

---

[5]There are many ways in which graphs may be encoded (with a standard one being in a matrix). This notation proves to be convenient for the study of individual incentives in network formation.

is also a best response to the other player mixing 50/50. Thus, A is the strategy that is a best response to the largest set of beliefs over possible plays of the opponent, and so the risk dominant equilibrium is the pure strategy equilibrium with a larger basin of attraction than the other pure strategy equilibrium. Specifically, playing A is a player's best response if the fraction of his opponents who play A is greater than or equal to $(b-c)/(a-d+b-c) < 1/2$. Thus if $a=3$, $b=1$, and $c=d=0$, then playing A is a best response if ¼ or more of a player's opponents play A.

In the case where n is even, assume that A is the best response to a mixture of $(n-2)/[2(n-1)]$ on A and $n/[2(n-1)]$ on B. So, for instance, if $n=10$ then A is a best response for a player who is playing against 4 other players who play A and 5 who play B. This assumption ensures an asymmetry among strategies, so that our results of multiple stochastically stable states are not artificial.

The dynamic process is described as follows. Each period one player is chosen at random (say with equal probability, although that is not important) to update his strategy. A player updates his strategy myopically, best responding to what the other players with whom he interacts did in the *previous* period. However, there is also a probability $1>\varepsilon>0$ that a player trembles, and chooses a strategy that he did not intend to. Thus, with probability $1-\varepsilon$ the strategy chosen is $a_i^t = \arg\max_{a_i} u_i(a_i, a_{-i}^{t-1})$ and with probability ε the strategy is $a_i^t \neq \arg\max_{a_i} u_i(a_i, a_{-i}^{t-1})$.[6] The probabilities of trembles are identical and independent across players, strategies, and periods.[7] These trembles can be thought of as mistakes made by players or exogenous factors that influence players' choices. Once initial strategies are specified, the above process leads to a well-defined Markov chain where the state is the vector of actions, a^t, that are played in period t. The Markov chain has a unique stationary distribution, denoted $\mu^\varepsilon(a)$. Thus, for any given initial strategies, $\mu^\varepsilon(a)$ describes the probability that *a* will be the state in some period (arbitrarily) far in the future. Let $\mu = \lim_\varepsilon \mu^\varepsilon$. Following the terminology of Foster and Young (1990), a given state *a* is *stochastically stable* if it is in the support of μ. Thus, a state is stochastically stable if there is a probability bounded away from zero that the system will be in that state according to the steady state distribution, for arbitrarily small probabilities of trembles.

Let us consider a specific example that illustrates the importance of the network configuration in determining the stochastically stable states. Let $n=4$ and $a=3$, $b=1$, and $c=d=0$. This is a coordination game where A,A is both the efficient and risk dominant equilibrium. In this situation, for either the complete network or circle networks the unique stochastically stable state is all players playing A; but, for the star network there are two stochastically stable states: all players playing A and all players playing B.

Example - The Complete Network

First, consider the case where players are located on the complete graph, g^N. Recall that $\pi_{ij}=1$ for each $ij \in g$. Thus for the graph g^N, each player plays once against every other player in a period (or

___

[6]Assume that $\arg\max_{a_i} u_i(a_i, a_{-i}^{t-1})$ is single valued, which is true generically in choices of the payoff matrix.

[7] The structure of trembles is not innocuous. As detailed by Bergin and Lipman (1996), varying the perturbation structure can have a profound influence on the dynamic process. We return to discuss this shortly.

equivalently faces each other player with an equal probability). The stochastically stable state in this case is for all players to play A, as shown in Kandori, Mailath, and Rob (1993) and Young (1993).

Let us briefly go over the intuition behind this example and the Kandori, Mailath, and Rob (1993) and Young (1993), as this will be useful later on. If at least one of the other players played A last period, then a player who is updating her strategy will choose to play A. A player who is updating her strategy will choose to play B only if all of the other players played B last period. So, consider a situation where all players are playing A. If due to an ε-error some player switches to B, the other players when called on to update will not choose to switch and will continue to play A. This situation will update back to the situation where all players play A. It takes three trembles on three different players in order to have the remaining player choose to switch to B. So, the process takes at least three trembles (without any intermediate updating by those players) to switch the process from all playing A to all playing B. If all players are playing B, then if one player switches to playing A the other players when called on to update will choose to switch to A. Thus, with only one tremble the process can lead from all playing B to all playing A. While this is not a complete description of the stochastic process it outlines why it is relatively easier (by orders of magnitude in the number of trembles needed) to move from social coordination on B to coordination on A than the other way around.

Example - Circle Networks

Next consider the case where players are located on a circle, as for instance in the network g={12,23,34,41}. The unique stochastically stable state will be all players playing A as shown by Ellison (1993), and as he shows this result also holds for some more general neighborhood structures on circles.

The intuition behind this example is similar to that of complete networks. Here, each player only cares about what his or her two neighbors are playing. Again, as long as one neighbor is playing A, an updating player will choose to play A. Thus it takes both neighbors switching to B in order to have a player want to switch from playing A to playing B, while it only takes one neighbor switching to A to get a player to want to switch from playing B to playing A.

Example - Star Networks

Lastly, consider the situation where players are located on a star such as g={12,13,14}. The dynamics associated with this network differ in important ways from those described around the previous two examples. With a star network there are two stochastically stable states: one where all players play A and the other where all players play B.

To see the intuition behind this example note that now players 2, 3 and 4 care only about what player 1 is playing, and they will update to play whatever 1 played last period when called on to update. Player 1, in contrast, cares about what all the players are doing. Thus one tremble by player 1 can lead from a network where all play A to one where all play B. Alternatively, any tremble of any player changing from B to A can lead from a situation where all play B to one

where all play A. Thus starting from either equilibrium of all play A or all play B, we need only one tremble to have updating lead naturally to the other equilibrium. As the relative number of trembles is the important factor in determining the set of stochastically stable states, both of these states are stochastically stable.

Note that the relative probability in the limiting distribution of the state where all play A is higher than that of the state where all play B. This follows from the fact that any single tremble can lead to a transition from the state where all play B to the state where all play A, while a specific tremble is needed to lead to a transition from the state where all play B to that where all play A. Nevertheless, the state where all play the inefficient and non-risk dominant equilibrium still receives positive probability in the limit distribution and is stochastically stable.

The star example shows that it is possible for a network of individuals to have multiple stochastically stable states in the coordination game, including some where the inefficient and non risk-dominant strategy is selected. This result is in contrast with the previous literature (e.g., Kandori, Mailath and Rob (1993), Young (1993), and Ellison (1993 and 1995)) where the risk dominant equilibrium is always selected.[8] This result has seemingly negative implications, as there may be inefficient play. On the other hand, the result can also have positive implications in the case where the risk dominant equilibrium A, A is inefficient. For instance, if a-c>b-d but b>a, then the analysis of the above examples still holds, and under the complete or circle networks only play of the inefficient equilibrium A,A is stochastically stable, while on the star network play of the efficient equilibrium B,B is also stochastically stable.

This example seems to contradict Theorem 6.1 (and Corollary 6.1) in Young (1998). Young's result states that for *any* fixed network structure the unique stochastically stable state in a symmetric coordination game is for all players to coordinate on the risk dominant strategy. However, there is no contradiction between the results as the perturbation structure is different in the two analyses. We assume that the probability of a player trembling (i.e., a perturbation or mistake) from A to B when A is the best response is the same as the probability of a player trembling from B to A when B is the best response. In contrast, Young (1998) assumes that updating takes place according to a distribution that is proportional to a factor that is log-linear in payoff. In particular, as ε goes to 0 in the Young (1998) framework, the probability of any player trembling from A to B when others are playing A becomes *infinitely* less likely than the probability of trembling from B to A when the others are playing B. Such an error structure is very tractable and powerful in its predictions (e.g., see Blume (1993) and Young (1998)), and reflects the belief that errors should be less relatively likely when they are more costly. However, such an error structure is extreme in that very small differences in relative payoff comparisons across strategies, lead to infinite differences in limiting perturbation probability. If an error reflects exogenous factors or limits in a player's calculation ability, then there is no reason to assume that such an error should become infinitely more likely for some strategies versus others. While we work with the other extreme assumption of equal probabilities of trembles, the results

---

[8] Bergin and Lipman (1996) have shown for any specific equilibrium, there is a set of perturbations varying in their order of magnitude across strategies that will select the desired equilibrium as the stochastically stable state. What we are showing here is different, in that there are identical trembles resulting in multiple stochastically stable states, one of which may be neither risk dominant nor efficient, and instead it is the network structure that matters.

we present here would still hold if the probabilities of single errors vary across strategy choices as long as the ratio of these error probabilities does not become infinite.

Given that the set of stochastically stable states varies with the network structure, we now turn to the question of which networks arise endogenously if the network structure is at the players' discretion. If players choose both with whom they interact as well as what strategies they play, then which networks should we expect to see? This is the question we address next.

3. Endogenous Networks and Stochastic Stability

The following outlines an approach for endogenizing the network. It is a simple variation on the process discussed above.

A Dynamic Process

Let g^{t-1} denote the network at the end of period t-1 and a^{t-1} denote the action profile at the end of period t-1. In an arbitrary period t three things occur.

1. First, one link ij is chosen at random according to the fixed probability distribution $\{p_{ij}\}$ where $p_{ij} > 0$ for each ij. This is the only link that can be formed or severed at time t. Players decide whether to add or sever the link, and make this choice based on the assumption that players (including themselves) will play the same strategy as in the previous period. If the link is not in the network then it is added if at least one player's utility increases and the other player's does not decrease. If the link is already in the network then it is severed if either player would benefit from its removal. After the choice is made, with probability $1>\gamma>0$ the choice is reversed by a tremble. This process determines a network g^t according to well-defined probabilities.[9]

2. Second, one player i is randomly selected to adjust their strategy according to the fixed probability distribution $\{q_i\}$ where $q_i > 0$ for each i. This player chooses the strategy that is a best response to the current network g^t and to the previous periods' play configuration a^{t-1}.[10] After the choice is made, with probability $1>\varepsilon>0$ it is reversed by a tremble. All trembles and random selections are independent. This determines a strategy profile a^t according to well defined probabilities.[11]

---

[9] So, conditional on ij being selected the process is as follows: If $ij \in g^t$ and $u_i(g^{t-1}-ij, a^{t-1}) > u_i(g^{t-1}, a^{t-1})$ or $u_j(g^{t-1}-ij, a^{t-1}) > u_j(g^{t-1}, a^{t-1})$ then $g^t = g^{t-1}-ij$ with probability $(1-\gamma)$ and g^{t-1} with probability γ. If $ij \in g^t$ and $u_i(g^{t-1}-ij, a^{t-1}) \le u_i(g^{t-1}, a^{t-1})$ and $u_j(g^{t-1}-ij, a^{t-1}) \le u_j(g^{t-1}, a^{t-1})$ then $g^t = g^{t-1}$ with probability $(1-\gamma)$ and $g^{t-1}-ij$ with probability γ. If $ij \notin g^t$ and $u_i(g^{t-1}-ij, a^{t-1}) \ge u_i(g^{t-1}, a^{t-1})$ and $u_j(g^{t-1}-ij, a^{t-1}) \ge u_j(g^{t-1}, a^{t-1})$ with one inequality strict, then $g^t = g^{t-1}+ij$ with probability $(1-\gamma)$ and g^{t-1} with probability γ. Otherwise, $g^t = g^{t-1}$ with probability $(1-\gamma)$ and $g^t = g^{t-1}+ij$ with probability γ. Taking probabilities across ij according to p_{ij} leads to a distribution over g^t as a function of g^{t-1}.

[10] Again, assume that the a_i that maximizes $u_i(g^t, a_i, a^{t-1}_{-i})$ is unique for every g^t, a^{t-1}_{-i}.

[11] Conditional on i being selected, $a^t_{-i} = a^{t-1}_{-i}$ and $a^t_i = \text{argmax } u_i(g^t, a_i, a^{t-1}_{-i})$ with probability $1-\varepsilon$ and $a^t_i \ne \text{argmax } u_i(g^t, a_i, a^{t-1}_{-i})$ with probability ε. Taking probabilities across i according to q_i leads to a distribution over a^t as a function of a^{t-1}.

3. Lastly, players play the coordination game with the other players that they are directly connected to in the network and receive the payoff $u_i(g^t, a^t)$, as defined in equation (1) where $\pi_{ij}(g^t)=1$ if $ij \in g^t$, and $\pi_{ij}(g^t)=0$ if $ij \notin g^t$.

Assume that disconnected players, when identified in step 2, choose a best response to being uniformly randomly matched with any other player and to the previous period's play configuration a^{t-1}. As before, the process determines a finite state, irreducible, aperiodic Markov chain, and thus has a unique stationary distribution $\mu^{\gamma, \varepsilon}$ over states, where states are now network/strategy configurations.

A network/strategy configuration g, a is *stochastically stable* if it is in the support of $\mu = \lim_{\gamma=k\varepsilon \to 0} \mu^{\gamma, \varepsilon}$, where we take γ and ε to zero at the same rate; so $\gamma = k\varepsilon$ for some $k>0$. [12]

In the above process, players adjust their links and strategies independently. Players do not consider the possibility of changing their strategy when adding or severing a link and the possible implications that this might have for the future evolution of play. This sort of consideration may be important when there are relatively small numbers of forward-looking players who are well-informed about the network, strategies played, and the motivation of others. However, in larger networks and situations where players' information might be local and limited, or in situations where players significantly discount the future, myopic behavior is a more natural assumption.

Proposition 1: *If c>0 and d>0, then the unique stochastically stable state is the fully connected network (g^N) with all players playing A.*

The proof of Proposition 1 is in the appendix.

Proposition 1 states that if the payoff to mis-coordination is positive, then the unique stochastically stable state involves formation of the full network and all players coordinating on the risk dominant equilibrium regardless of whether or not it is efficient. This result suggests that endogenizing the network leads to outcomes similar to those suggested in the exogenous network literature of Kandori, Mailath and Rob (1993), Young (1993, 1998), and Ellison (1993, 1995).

Proposition 1 also has implications regarding speed of convergence to the stochastically stable state. As discussed in Ellison (1993, 1995) and Young (1998), network structures with features that are similar to that of a circle, i.e., where players have tight local interactions and are more loosely linked across neighborhoods, have faster speeds of convergence to the stochastically stable state. Young (1998) has an elegant characterization of such ``close knit'' networks. However, Proposition 1 says that when the network is endogenized, we will not end up with the close knit networks that have the nice speed-of-convergence properties, but instead we end up with fully connected networks. This occurs because links are essentially costless. Any two

---

[12]We do not have to worry about cycles or mixed strategies given this "one at a time" process and the structure of the payoffs.

9

disconnected players will benefit from adding a link since all links play in each period and all links guarantee positive payoffs. With this in mind, the intuition behind Proposition 1 is transparent as the usual evolutionary reasoning applies and so the full network with all playing A is the only stochastically stable state.

Proposition 1 should be interpreted cautiously, however, as it depends critically on there being no significant costs to having a link. This is captured in c>0 and d>0, where players still receive a positive payoff when they fail to coordinate. Things change if we add a significant cost to maintaining links. If there is a cost k>0 of maintaining a link, which is just subtracted from each entry of the payoff matrix, then for non-trivial k the payoff from mis-coordination can be negative. As the structure of the payoff matrix is maintained, this can be modeled simply by considering variations in the payoff parameters a,b,c, and d.

Proposition 2: *If failing to coordinate in the game is costly so that c<0 and d<0, (but coordinating links are still beneficial so that a>0 and b>0), then there are two stochastically stable states: a fully connected network with all players playing A, and a fully connected network with all players playing B.*

The proof of Proposition 2 appears in the appendix.

If a<0 (or b<0), then the state where all play A (or all play B) is no longer stable. If both are less than 0, then the disconnected network becomes the only stochastically stable state.

Proposition 2 shows that endogenizing the network has implications beyond predictions of the network structure; it also has implications for the strategies that are chosen in the game. Proposition 2 shows that there are stochastically stable states where actions are played that are neither risk dominant nor efficient. That is, all playing B can be part of a stochastically stable state even when a>b. Comparing this result to our exogenous network results, we find that while the star network does not arise endogenously, stochastically stable play is similar to play observed on the exogenous star network.

The intuition behind Proposition 2 builds off of the endogeneity of the network. Suppose that we are initially in a network where all players are playing A and are fully connected. Suppose that two trembles occur and two players start playing B. Then, players playing A would like to sever links to these players, as it is costly to maintain a link with B players. These B players would like to join together and form a component playing B. This component can then continue to grow as additional trembles occur. This process turns out to be symmetric in the way it moves between A and B, thus allowing both states to be stable. The other candidates for stochastically stable states are situations where there are two separate, fully-connected components to the network, one playing all A and the other playing all B. However, it can be shown that these are not stochastically stable.

Let us make some further remarks concerning Proposition 2.

First, although Proposition 2 has all players fully connected in the stochastically stable states, the

speed of transition from any state to one of the stochastically stable states may be faster than with an exogenous fully connected network. That is, the network adjusts as players' strategies change. So, even though the stable states have fully connected networks, the transitions involve changes in those connections that allow trembles to build up one-at-a-time.[13] In the exogenous fully connected network case, to go from all playing B to all playing A requires m trembles (where $m = n(b-c)/(a-d+b-c) < n/2$), which must occur simultaneously. In Proposition 2, to go from all playing B to all playing A requires $(n-1)$ trembles, but these trembles can occur one-at-a-time. Thus Proposition 2 will have a faster speed of convergence than the fully connected network if m is large, but will have a slower speed of convergence if m is small.

Second, Proposition 2 depends on the relative change rates of links and strategies. For instance, if link patterns are much more rigid than strategies (by an order of $1/\varepsilon$), then some of the reasoning above may not apply as players strategies may readjust before the network adjusts. The relative ease of change, of course, will depend on the application.

Third, the analysis depends on myopic choices on the part of players. Players are not forecasting the responses of other players, in terms of the strategies played or the network, when they decide on their own actions and links. This myopia is clearly important in the reasoning behind the results, and thus the analysis is best suited for large settings where history is the best benchmark for predicting behavior. Although this is a caution that applies fairly broadly to the literature on stochastic stability, it has to be kept in mind when interpreting the results.

Finally, the analysis above assumes that in each period, players get a separate payoff for each player with whom they are connected. If instead, there is a limit so that each player actually plays on only a subset of links, but still bears some cost for all links that he or she maintains, then the results can change and in particular we can find mixed stochastically stable states where there is play of different strategies by different groups. To get a feeling for this, consider an example where a player's payoff is the average (rather than the sum) of the payoffs from each of her links and that the player pays a positive cost for each link. Then the stochastically stable states are networks where players are paired and each pair of players plays the same strategy as her partner. However, it is straightforward to check that different pairs may play different strategies.

3. Concluding Remarks

From the previous literature one might take away two predictions: (i) With fixed networks of interaction society will coordinate on the risk-dominant equilibrium, and, (ii) with endogenous interaction patterns as determined by location (with homogeneous players[14]) society will coordinate on the efficient equilibrium. Here we have shown via an example that (i) may fail and one can see multiple stochastically stable states on a fixed network; and similarly that (ii) may fail as with endogenous networks, multiple stochastically stable states can include coordination on

---

[13] Speed of convergence and its relationship to one-at-a-time trembles is discussed formally in Ellison (1995).
[14] Mailath, Samuelson and Shaked (1997) give an example of a locational model where there exists a stable state with heterogeneous play among the population. Their example builds off of heterogeneity of players in the population, where players' types affect their matching probabilities and their ability to choose with whom to interact. We are referring to situations where players are initially homogeneous.

equilibria that are neither efficient nor risk dominant.

In our model, whether or not efficient play or risk dominant play (or neither) is reached depends on a number of assumptions, such as whether or not failure to coordinate is costly, whether players are limited in the number of interactions that they can undertake, and whether players are forward looking or myopic. These assumptions are in addition to more basic differences between an endogenous network approach and a locational approach, such as caring about the sum of payoffs across interactions rather than an average[15] and being able to refuse to connect with an individual.

This sensitivity to a wide variety of assumptions suggests that the conclusions in this literature be interpreted cautiously.[16] Given this sensitivity to a number of aspects of behavior and interaction technology, there may not be broad-sweeping predictions that one can make concerning a society's ability to reach efficient coordination, but instead such predictions are likely to be dependent on the details of the application. Rather than trying to catalogue results based on a matrix of modeling choices, we suggest tailoring the models to match specific applications as they arise.

4. References

Bergin, J. and B. Lipman (1996) "Evolution with State-Dependent Mutations," **Econometrica**, 64, 943-956.

Besen, S. and J. Farrell (1994) "Choosing How to Compete: Strategies and Tactics in Standardization," **Journal of Economic Perspectives**, 8, 117-131.

Blume, L. (1993) "The Statistical Mechanics of Strategic Interaction," **Games and Economic Behavior**, 4, 387-424.

Ellison, G. (1993) ``Learning, Local Interaction, and Coordination,'' **Econometrica**, 61, 1047-1071.

Ellison, G. (1995) "Basins of Attraction, Long Run Equilibria, and the Speed of Step-by-Step Evolution," mimeo: M.I.T.

Ely, J. (1998) "Local Conventions," mimeo: Northwestern University.

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

Harsanyi, J. and R. Selten (1988) A General Theory of Equilibrium Selection in Games, Cambridge: MIT Press.

---

[15] In the locational models payoffs are expected over interactions, and so the size of the population (holding composition constant) at a given location does not affect the locational choice.

[16] This is complementary to a caution raised by Bergin and Lipman (1996) detailing the sensitivity of stochastic stability to the perturbation technology.

Jackson, M. and A. Watts (1998), "The Evolution of Social and Economic Networks," Caltech WP# 1044.

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

Katz, M. and C. Shapiro (1994) "Systems Competition and Networks Effects," **Journal of Economic Perspectives**, 8, 93-115.

Liebowitz, S. and S. Margolis (1994) "Network Externality: An Uncommon Tragedy," **Journal of Economic Perspectives**, 8, 133-150.

Mailath, G., L. Samuelson and A. Shaked (1997) "Endogenous Interactions," mimeo: University of Pennsylvania.

Young, H.P. (1993) ``The Evolution of Conventions, **Econometrica**, 61, 57-84.

Young, H.P. (1998) *Individual Strategy and Social Structure*, Princeton University Press: Princeton.

Appendix

Let $x = (g,a)$ represent a network and strategy combination.

Two network-strategy combinations x and x' are *adjacent* if they differ by at most one link and/or one player's strategy.

Define a *path* $(x_1,\ldots,x_k)$ as a sequence of network-strategy combinations such that x_i and x_{i+1} are adjacent but $x_i \neq x_{i+1}$ for each $i=1,\ldots,k-1$. Let $p(x_1,x_k)$ represent the set of all possible paths starting at x_1 and ending at x_k.

Given a two adjacent networks, (x,x'), let the *resistance* $r(x,x')$ be the number of trembles needed to go from x to x'. A tremble is needed if the players involved are not willing to make the change of the link and/or the strategy that is needed to move from x to x'. Thus to go from x to x' takes at most 2 trembles (to change both the link and the strategy) and at least 0 trembles.

To be precise, let $x =(g,a)$ and $x' = (g',a')$. No tremble is needed to move from g to g' (given a) if $g'=g$; $g'=g-ij$ and either $u_i(g, a)<u_i(g-ij, a)$ or $u_j(g, a)<u_j(g-ij, a)$; or if $g'=g+ij$ and $u_i(g+ij, a)\geq u_i(g, a)$ and $u_j(g+ij, a)\geq u_j(g, a)$ with one inequality holding strictly. Otherwise, one tremble is needed to move from g to g' (given a). No tremble is needed to move from a to a' (given g') if $a=a'$ or if $a_i \neq a'_i$ and $a'_i =\text{argmax } u_i (g',a_i, a'_{-i})$. Otherwise, one tremble is needed to move from a to a' (given g'). Then, $r(x,x')=2$ if one tremble is needed to go from g to g' (given a) and one tremble is needed to go from a to a' (given g'). $r(x,x')=1$ if one tremble is needed to go from g to g' (given a) and no trembles are needed to go from a to a' (given g'), or if no trembles are needed to go from g to g' (given a) and one trembles is needed to go from a to a' (given g'). $r(x,x')=0$ if

no trembles are needed to go from g to g' (given a) and no trembles are needed to go from a to a' (given g').

Define *the resistance of a path* $r(x_1,\ldots,x_k)$ as $\sum_{i=1}^{k-1} r(x_i,x_{i+1})$.

The path $(x_1,\ldots,x_k)$ is an *improving path* if $r(x_1,\ldots,x_k) = 0$. For further discussion of improving paths, see Jackson and Watts [1998].

Let $r(x_1,x_k) = \min_{(X1,\ldots,Xk)\,\in\, p(X1,Xk)} r(x_1,\ldots,x_k)$.

A theorem from Young [1993] is instrumental in the proofs of Propositions 1 and 2. Before stating Young's [1993] theorem, the following definitions from Young (1993) are needed.

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) $P(\varepsilon) \rightarrow P$, and (iii) $P(\varepsilon)_{xy} > 0$ implies that there exists $r \geq 0$ 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 n-th order transition probability associated with P of x to z is positive for some n.

The *recurrent communication classes* of P, denoted $X_1,\ldots,X_J$, 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 $x_1,\ldots x_k$ 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 $\Sigma_1^{K-1} r(x_k,x_{k+1})$, 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.

Given any state x, an *x-tree* is a directed graph with a vertex for each state and a unique directed

14

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 [1993]): 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.

Proof of Proposition 1: The stochastic process of Proposition 1 determines a finite state, irreducible, aperiodic Markov chain, and thus has a unique stationary distribution $\mu^{\gamma, \varepsilon}$ over states, where states are network/strategy combinations. Given that $\gamma = k\varepsilon$, we can write $\mu^{\gamma, \varepsilon}$ as μ^{ε}.

Note that from any x=(g,a) there is a path of zero resistance leading to either (g^N,A) or (g^N,B), where (g^N,A) and (g^N,B) represent the states where players are in the fully connected network g^N and all players play strategy A (or B, respectively). This follows since all payoffs are strictly positive and so players always prefer to add a link when given the opportunity, and the fact that all players must coordinate on strategy A or B given the symmetry of payoffs and the assumed uniqueness of the best response in the coordination game. Note also that any path from (g^N,A) or (g^N,B) to any adjacent network-strategy pair has positive resistance.

Thus, the recurrent communication classes as defined above must be $\{(g^N,A)\}$ and $\{(g^N,B)\}$. The set of stochastically stable states depends on which of these states has the smallest resistance. Given the theorem above, we need only find the minimum resistance over paths from $\{(g^N,A)\}$ to $\{(g^N,B)\}$ and compare that to the minimum resistance over paths from $\{(g^N,A)\}$ to $\{(g^N,B)\}$.

If agents are in the g^N network and all playing B, then given the payoff structure if we consider a path from (g^N,A) to (g^N,B) where only strategies are changed it takes less than n/2 trembles to get all agents to switch to all playing A.   So, $r((g^N,B), (g^N,A)) < n/2$. Next we bound $r((g^N,A), (g^N,B))$. If agents are in a g^N network and all agents play A, then given the payoff structure if no links are changed it will take at least n/2 trembles to get players to want to switch to play B. So, if the resistance is less than n/2 it must involve some trembles on links. Let us consider what it would take to get some single player to switch from playing A to B. For a player to switch from A to B more than half of the players that this player is linked to must play B. Thus, if k links are severed, then (n-k)/2 players must be playing B if n-k-1 is odd and (n-k+1)/2 if n-k-1 is even. So starting at (g^N,A) to get a single player to want to switch from A to B involves at least k+(n-k)/2 trembles and this is at least n/2. Thus $r((g^N,A), (g^N,B)) \geq n/2$. Since $r(g^N,A) < r(g^N,B), (g^N,A)$ is the unique stochastically stable state.

Proof of Proposition 2: If a>0>d and b>0>c then players will prefer to sever any link to a player who plays a strategy that is different than their own, and add a link to any player who plays a

strategy that is the same as their own. Thus, the set of recurrent communication classes are $\{(g^N,A)\}$, $\{(g^N,B)\}$, and each $\{(g^k,A/g^{n-k},B)\}$ for $2 \le k \le n-2$. Here $(g^k,A/g^{n-k},B)$ represents any state where there are two separate fully connected components, one of size k and the other of size (n-k) with all players in the k-size component playing A and all players in the (n-k) size component playing B. Note that for any k, there are many different states $(g^k,A/g^{n-k},B)$ since here are many ways to select k individuals. Each is a distinct recurrent communication class.

Given Young's theorem, we look for the set of recurrent communication classes with minimum resistance, found by constructing restricted (g,a)-trees. First we construct a restricted (g^N,A)-tree. Direct (g^N,B) to any of the $(g^2,A/g^{n-2},B)$ vertices; (g^N,B) has resistance 2 to an improving path leading to any $(g^2,A/g^{n-2},B)$ vertex. (Starting at (g^N,B) allow two trembles to change two player's strategies to A. Then there is an improving path leading to a $(g^2,A/g^{n-2},B)$ vertex. This improving path exists since all of the B players will receive a negative payoff from their link with the A players, and will sever it when given the opportunity. Note that (g^N,B) has a distance of more than 2 from an improving path to any other communication class, since more than 2 players strategies must be changed.) Any of the $(g^k,A/g^{n-k},B)$, $2 \le k \le n-3$, vertices has a distance of 1 from an improving path leading to one of the $(g^{k+1},A/g^{n-k-1},B)$ vertices; simply allow a tremble to change one of the B players to strategy A. Any of the $(g^{n-2},A/g^2,B)$ vertices has a distance of 1 from an improving path leading to (g^N,A). Simply allow a tremble to change one of the B players to strategy A, then the remaining B player will sever the tie to this player. The remaining B player will now be unlinked; since we assume that disconnected players, choose a best response to the current average play of all other players, this B player will change strategies to A and will then form links with all the other A players. Thus $r(g^N,A)$ is equal to the number of recurrent communication classes. Similar reasoning shows that $r(g^N,B)$ is the same.

Next we compute $r(g^k,A/g^{n-k},B)$ for $2 \le k \le n-2$. From the above, we know that g^N/A and g^N/B are both a distance of 2 or more from an improving path leading to any $(g^k,A/g^{n-k},B)$. Thus $r(g^k,A/g^{n-k},B)$ is strictly greater than the number of recurrent communication classes, and so larger than $r(g^N,A)=r(g^N,B)$. It follows from Young's theorem that the stochastically stable states are (g^N,A) and (g^N,B).