The Effects of Social Networks on Employment and Inequality

By ANTONI CALVÓ-ARMENGOL AND MATTHEW O. JACKSON*

We develop a model where agents obtain information about job opportunities through an explicitly modeled network of social contacts. We show that employment is positively correlated across time and agents. Moreover, unemployment exhibits duration dependence: the probability of obtaining a job decreases in the length of time that an agent has been unemployed. Finally, we examine inequality between two groups. If staying in the labor market is costly and one group starts with a worse employment status, then that group's drop-out rate will be higher and their employment prospects will be persistently below that of the other group. (*JEL* A14, J64, J31, J70)

The importance of social networks in labor markets is pervasive and well documented. Mark Granovetter (1973, 1995) found in a survey of residents of a Massachusetts town that over 50 percent of jobs were obtained through social contacts. Earlier work by Albert Rees (1966) found numbers of over 60 percent in a similar study. Exploration in a large number of studies documents similar figures for a variety of occupations, skill levels, and socioeconomic backgrounds.[1]

In this paper, we take the role of social networks as a manner of obtaining information about job opportunities as a given and explore its implications for the dynamics of employment. In particular, we examine a simple model of the transmission of job information through a network of social contacts. Each agent is connected to others through a network. Information about jobs arrives randomly to agents. Agents who are unemployed and directly hear of a job use the information to obtain a job. Agents who are already employed, depending on whether the job is more attractive than their current job, might keep the job or else might pass along information to one (or more) of their direct connections in the network. Also, in each period some of the agents who are employed randomly lose their jobs. After documenting some of the basic characteristics and dynamics of this model, we extend it to analyze the decision of agents to drop out of the labor force based on the status of their network. This permits us to compare the dynamics of drop-out rates and employment status across groups.

The fact that participation in the labor force is different across groups such as whites and blacks is well documented. For instance, David Card and Alan B. Krueger (1992) see a drop-out rate 2.5 to 3 times higher for blacks compared to whites. Amitabh Chandra (2000) provides a breakdown of differences in participation rates by education level and other characteristics, and finds ratios of a similar magnitude. Moreover, the analysis of James Heckman et al. (2000) suggests that differences in drop-out rates are an important part of the inequality in wages across races and that accounting for dropouts actually increases the black-white wage gap.[2,3]

* Calvó-Armengol: Universitat Autònoma de Barcelona, Department of Economics, Edifici B, 08193 Bellaterra (Barcelona), Spain (e-mail: antoni.calvo@uab.es; http://selene.uab.es/acalvo); Jackson: Division of the Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena, CA 91125 (e-mail: jacksonm@hss.caltech.edu; http://www.hss.caltech.edu/~jacksonm/Jackson.html). We thank Valentina Bali, Kim Border, Antonio Cabrales, Janet Currie, François Fontaine, Isa Hafalir, Eddie Lazear, Massimo Morelli, David Pérez-Castrillo, and Giorgio Topa for helpful conversations and discussions. We also thank seminar participants and three anonymous referees for their comments and suggestions. We gratefully acknowledge the financial support of Spain's Ministry of Science and Technology under Grant Nos. SEC2001-0973 and BEC2002-02130, and the Lee Center for Advanced Networking at Caltech.

[1] See James Montgomery (1991) for further discussion and references.

[2] Ignoring dropouts biases estimated wages upwards. Given much higher drop-out rates for blacks, this can bias the wage differential.

[3] The persistent inequality in wages between whites and blacks is one of the most extensively studied areas in labor

One is then left to explain why drop-out rates should differ across races. An analysis of social networks provides a basis for observing both higher drop-out rates in one race versus another and sustained inequality in wages and employment rates, even among those remaining in the labor force, as follows.[4] The starting point of our model is that the better the employment status of a given agent's connections (e.g., relatives, friends, acquaintances), the more likely it is that those connections will pass information concerning a job opening to the agent. This might be for any number of reasons. One is that as the employment status of a connection improves it is less likely that the connection will want to keep a new job for him or herself. Another is that as the employment status of other agents in the network improves, the more likely that a given agent's connections will be employed and a source of information, and the more likely that they will choose to pass news to the given agent rather than to some other agent who already has a job. This sort of information passing leads to positive correlation between the employment status of agents who are directly or indirectly connected in the network, within a period and across time, as we establish below.

Such correlation of employment is observed in the data, both in the indirect form of the inequality mentioned previously, and also in terms of direct measured correlation patterns. Giorgio Topa (2001) demonstrates geographic correlation in unemployment across neighborhoods in Chicago, and also finds a significantly positive amount of social interactions across such neighborhoods. Timothy Conley and Topa (2001) find that correlation also exists under metrics of travel time, occupation, and ethnicity; and that ethnicity and race are dominant factors in explaining correlation patterns.

The positive correlation that we establish between the employment of agents in a network then provides a basis for understanding the sustained difference in drop-out rates and resulting inequality in employment rates. Consider two identical networks, except that one starts with each of its agents having a better employment status than their counterparts in the other network. Now consider the decision of an agent to either remain in the labor market or to drop out. Remaining in the labor market involves some costs, which could include things like costs of skills maintenance, education, and opportunity costs. Agents in the network with worse initial starting conditions have a lower expected discounted stream of future income from remaining in the network than agents in the network with better initial starting conditions. This minor difference might cause some agents to drop out in the worse network but remain in the better network. The important observation is that dropping out has a contagion effect. When some of an agent's connections drop out, that agent's future prospects worsen since the agent's network is no longer as useful a source of job information. Thus, some agents connected to dropouts also drop out due to this indirect effect. This can escalate, so a slight change in initial conditions can lead to a substantial difference in drop-out decisions. As a larger drop-out rate in a network leads to worse employment status for those agents who remain in the network, we find that slight differences in initial conditions can lead to large differences in drop-out rates and sustained differences in employment rates.[5]

Before proceeding to the model, let us also mention a fourth feature that is also exhibited by the model. Unemployment exhibits duration dependence and persistence. That is, when conditioning on a history of unemployment, the expected probability of obtaining a job decreases in the length of time that an agent has been unemployed. Such duration dependence is well documented in the empirical literature, for example, in studies by Stuart O. Schweitzer and

economics. James P. Smith and Finis R. Welch (1989), using statistics from census data, break the gap down across a variety of dimensions and time. The gap is roughly on the order of 25 percent to 40 percent, and can be partly explained by differences in skill levels, quality of education, and other factors [e.g., see Card and Krueger (1992); Chandra (2000)]. See Reynolds Farley (1990) for a comparison of labor market outcomes for 50 racial-ethnic groups in the United States.

[4] In this paper we focus on differences in drop-out rates and employment, and we refer the reader to Calvó-Armengol and Jackson (2003) for an analysis of wage inequality.

[5] Similarly, small differences in network architectures may also lead to sharp differences in drop-out patterns across two groups with identical starting conditions.

Ralph E. Smith (1974), Heckman and George Borjas (1980), Christopher Flinn and Heckman (1982), and Lisa M. Lynch (1989). To get a feeling for the magnitude, Lynch (1989) finds average probabilities of finding employment on the order of 0.30 after one week of unemployment, 0.08 after eight weeks of unemployment, and 0.02 after a year of unemployment.

The reason that we see duration dependence in a networked model of labor markets is a simple one. A longer history of unemployment is more likely to come when the direct and indirect connections of an agent are unemployed. Thus, seeing a long spell of unemployment for some agent leads to a high conditional expectation that the agent's contacts are unemployed. This in turn leads to a lower probability of obtaining information about jobs through the social network. This network explanation is orthogonal to standard explanations such as unobserved heterogeneity.

At this point, let us preview a policy prediction that emerges from a networked model. Due to the network effects, improving the status of a given agent also improves the outlook for that agent's connections. This is the contagion effect discussed above in reverse. As a result, in a networked model there are local increasing returns to subsidizing education, and other policies like affirmative action.[6] One implication is that it can be more efficient to concentrate subsidies or programs so that a cluster of agents who are interconnected in a network are targeted, rather than spreading resources more broadly.

Before presenting the model, we note that we are certainly not the first researchers to recognize the importance of social networks in labor markets. Just a few of the studies of labor markets that have taken network transmission of job information seriously are Scott A. Boorman (1975), Montgomery (1991, 1992, 1994), Kenneth Arrow and Ron Borzekowski (2001), Topa (2001), and Calvó-Armengol (2004)—not to mention the vast literature in sociology.[7] The contribution here is that we are the first to study an explicit network model and prove some of the resulting implications for the patterns and dynamics of employment, as well as the inequality across races.

I. A Simple Network Model

The model we consider here is one where all jobs are identical. We refer the reader to a companion paper, Calvó-Armengol and Jackson (2003), for a more general model that nests this model, and also looks at wage dynamics, and allows for heterogeneity in jobs, decisions as to whether to switch jobs, repeated and selective passing of information, competing offers for employment, and other extensions of the model presented here. In short, the results presented here extend to wage inequality as well, and are quite robust to the formulation.

There are n agents. Time evolves in discrete periods indexed by t. The vector $\mathbf{s}_t$ describes the employment status of the agents at time t. If agent i is employed at the end of period t, then $s_{it} = 1$ and if i is unemployed then $s_{it} = 0$.

A period t begins with some agents being employed and others not, as described by the status $\mathbf{s}_{t-1}$ from the last period. Next, information about job openings arrives. In particular, any given agent hears about a job opening with a probability a that is between 0 and 1. This job arrival process is independent across agents. If the agent is unemployed, then he or she will take the job. However, if the agent is already employed then he or she will pass the information along to a friend, relative, or acquaintance who is unemployed. This is where the network pattern of relationships is important, as it describes who passes information to whom, which is ultimately crucial in determining a person's long-term employment prospects. We now describe these network relationships and the process of information exchange.

Any two people either know each other or do not, and in this model information only flows between agents who know each other. A graph g summarizes the links of all agents, where $g_{ij} = 1$ indicates that i and j know each other,

[6] In our model, improving the status of one agent has positive external effects on other agents' expected future employment. There are, of course, other factors that might counterbalance this sort of welfare improvement: for instance, the difficulty that an agent might have adapting to new circumstances under affirmative action as discussed by George A. Akerlof (1997).

[7] Some related references can be found in Montgomery (1991) and Granovetter (1995).

and $g_{ij} = 0$ indicates that they do not know each other. It is assumed that $g_{ij} = g_{ji}$, meaning that the acquaintance relationship is a reciprocal one.

If an agent hears about a job and is already employed, then this agent randomly picks an unemployed acquaintance to pass the job information to. If all of an agent's acquaintance are already employed, then the job information is simply lost. The probability of the joint event that agent i learns about a job and this job ends up in agent j's hands, is described by $p_{ij}(\mathbf{s})$, where

$$p_{ij}(\mathbf{s}) = \begin{cases} a & \text{if } s_i = 0 \text{ and } i = j, \\ \dfrac{a}{\sum_{k:s_k=0} g_{ik}} & \text{if } s_i = 1,\ s_j = 0, \text{ and } g_{ij} = 1; \text{ and} \\ 0 & \text{otherwise,} \end{cases}$$

and where the vector $\mathbf{s}$ describes the employment status of all the agents at the beginning of the period.

In what follows, we will also keep track of some indirect relationships, as friends of a friend will play a couple of roles. First, they are competitors for job information in the short run. Second, they help keep an agent's friends employed, which is a benefit in the longer run. We say that i and j are *path-connected* under the network g if there exists a sequence of links that form a path between i and j.

Finally, the last thing that happens in a period is that some agents lose their jobs. This happens randomly according to an exogenous breakup rate, b, between 0 and 1. This is the probability that any given employed agent will lose his or her job at the end of a given period, and this is also independent across agents.

II. The Dynamics and Patterns of Employment

As time unfolds, employment evolves as a function of both past employment status and the network of connections which, together, randomly determine the new employment. Employment thus follows a finite state Markov process, where the state is the vector of agents' employment status at the end of a period and transition probabilities are dependent on the network of relationships. We wish to characterize the behavior of this stochastic process.

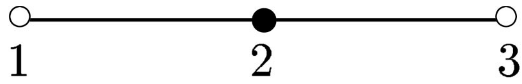


FIGURE 1. NEGATIVE CORRELATION IN CONDITIONAL EMPLOYMENT

The relationship between the one-period-ahead employment status of an agent and his pattern of connections, as described by the $p_{ij}(\mathbf{s})$'s above, is clear. Having links to employed agents improves i's prospects for hearing about a job if i is unemployed. In addition, decreasing the competition for information from two-link-away connections is helpful. That is, if friends of my friends are employed rather than unemployed, then I have a higher chance of being the one that my friends will pass information to. Further indirect relationships (more than two-links away) do not enter the calculation for the one-period-ahead employment status of an agent. However, once we take a longer time perspective, the evolution of employment across time depends on the larger network and status of other agents. This, of course, is because the larger network and status of other agents affect the employment status of i's connections.

We first present an example which makes it clear why a full analysis of the dynamics of employment is subtle.

Example 1 (Negative Conditional Correlation): Consider Figure 1, a network with three agents, and suppose the employment from the end of the last period is $\mathbf{s}_{t-1} = (0, 1, 0)$. In the picture, a darkened node represents an employed agent (agent 2), while unemployed agents (1 and 3) are represented by empty nodes. A line between two nodes indicates that those two agents are linked.

Conditional on this state $\mathbf{s}_{t-1}$, the employment states s_{1t} and s_{3t} are negatively correlated. This is due to the fact that agents 1 and 3 are "competitors" for any job news that is first heard by agent 2.

Despite this negative (conditional) correlation in the shorter run, agent 1 can benefit from 3's presence in the longer run. Indeed, agent 3's presence helps improve agent 2's employment

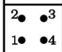
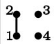
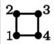
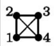
g	Prob$(s_1 = 0)$	Corr(s_1, s_2)	Corr(s_1, s_3)
	0.132	–	–
	0.083	0.041	–
	0.063	0.025	0.019
	0.050	0.025	0.025

FIGURE 2. CORRELATION AND NETWORK STRUCTURE I

status. Also, when agent 3 is employed, agent 1 is more likely to hear about any job that agent 2 hears about. These aspects of the problem counter the local (conditional) negative correlation, and help induce a positive correlation between the employment status of agents 1 and 3.

The benefits from having other agents in the network outweigh the local negative correlation effects, if we take a long-run perspective. The following examples illustrate the long-run behavior of the Markov process regulating employment as shaped by the underlying network of contacts between agents.

Example 2 (Correlation and Network Structure): Consider an example with $n = 4$ agents and let $a = 0.100$ and $b = 0.015$. If we think about these numbers from the perspective of a time period being a week, then an agent loses a job roughly on average once in every 67 weeks, and hears (directly) about a job on average once in every ten weeks. Figure 2 shows unemployment probabilities and correlations between agents' employment statuses under the long-run steady-state distribution.[8]

If there is no network relationship at all, then we see an average unemployment rate of 13.2 percent. Even moving to just a single link ($g_{12} = g_{21} = 1$) substantially decreases the probability (for the linked agents) of being unemployed, as it drops by more than a third, to 8.3 percent. The resulting unemployment rate aggregated over the four agents is 10.75 percent. As we see from Figure 2, adding more links further decreases the unemployment rate, but with a decreasing marginal impact. This makes sense, as the value to having an additional link comes only in providing job information when all of the existing avenues of information fail to provide any. The probability of this is decreasing in the number of connections.

The correlation between two agents' employment is (weakly[9]) decreasing in the number of links that each an agent has, and the correlation between agents' employment is higher for direct compared to indirect connections. The decrease as a function of the number of links is due to the decreased importance of any single link if an agent has many links. The difference between direct and indirect connections in terms of correlation is due to the fact that direct connections provide information, while indirect connections only help by indirect provision of information that keeps friends, friends of friends, etc., employed.

Also, note that the correlation between agents 1 and 3 in the third row of Figure 2 is positive (1.9 percent). Thus, even though agents 1 and 3 are in competition for information from both agents 2 and 4 in the shorter run, their employment is positively correlated in the long run. This will be true more generally, as stated in the propositions below.

Next, Figure 3 examines some eight-person networks, with the same information arrival and job breakup rates, $a = 0.100$ and $b = 0.015$.[10]

Here, again, the probability of unemployment falls with the number of links, and the correlation between two employed agents decreases with the distance of the shortest path of links (geodesic) between them.

Also, we can see some comparisons to the four-person networks: an agent has a lower unemployment rate in a complete four-person network than in an eight-person circle. In this example, the direct connection is worth more than a number of indirect ones. More generally,

[8] The numbers for more than one agent are obtained from simulations in Maple®. We simulate the economy over a large number of periods (hundreds of thousands) and calculate observed unemployment averages and correlations. The programs are available upon request from the authors. The correlation numbers are only moderately accurate, even after several hundred thousand periods.

[9] In some cases, the correlations are indistinguishable to the accuracy of our simulations.

[10] We know from Propositions 1 and 2 that Corr(s_1, s_5) is positive for the top network in Figure 3. However, it is too small to accurately report its numerical value based on our simulations.

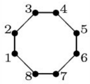
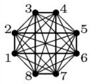
g	Prob$(s_1 = 0)$	Corr(s_1, s_2)	Corr(s_1, s_3)	Corr(s_1, s_4)	Corr(s_1, s_5)
	0.060	0.023	0.003	0.001	–
	0.030	0.014	0.014	0.014	0.014

FIGURE 3. CORRELATION AND NETWORK STRUCTURE II

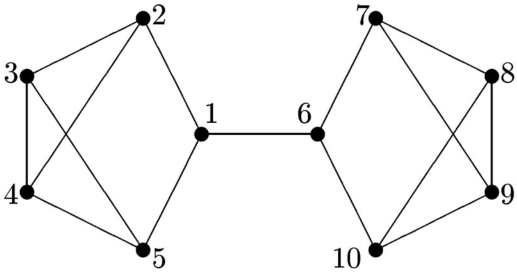


FIGURE 4. A NETWORK WITH A BRIDGE

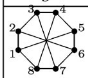
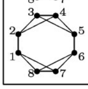
g	average path length	average unemployment
	1.571	0.048
	1.786	0.049

FIGURE 5. PATH LENGTH MATTERS

the trade-off between direct connections and indirect ones will depend on the network architecture and the arrival and breakup rates. In this example, agents rarely lose jobs, and hear about them relatively frequently, and so direct connections are employed with a high probability regardless of the number of their neighbors, and so indirect connections are less important than direct ones. In situations with higher breakup rates and lower arrival rates, indirect connections can become more important.

The model also provides a tool for analyzing asymmetries in the network.

Example 3 (Bridges and Asymmetries): Consider the network in Figure 4. Again we calculate employment from simulations using the same arrival and breakup rates as in the previous examples.

In this network the steady-state unemployment probabilities are 4.7 percent for agents 1 and 6, 4.8 percent for agents, 2, 5, 7, and 10, and 5.0 percent for the rest. While these are fairly close, simple differences of an agent's position in the network affects his or her unemployment rate, even though all agents all have the same number of connections. Here agents 1 and 6 have lower unemployment rates than the others, and 3, 4, 8, and 9 are the worst off. If we compare agent 3 to agent 1, we note the following: the average geodesic (minimum path) distance between any two agents who are directly connected to agent 3 is 4/3. In a sense, the agents that agent 3 is connected to are not "well diversified." In contrast, the average geodesic (minimum path) distance between any two agents who are directly connected to agent 1 is 2. Moreover, agents 5 and 6 are only path-connected through agent 1. In fact, 1 and 6 form what is referred to as a "bridge" in the social networks literature.[11]

Example 4 (Structure Matters: Densely Versus Closely Knit Networks): The model can also show how other details of the network structure matter. Compare the long-run average unemployment rates on two eight-person networks with 12 links each. In both networks, all agents have exactly three links. But, the average length of the paths connecting agents is different across networks. Again, we run simulations with $a = 0.100$ and $b = 0.015$; see Figure 5.

The average path length is lower for the circle with diameters than for the circle with local

[11] The lower unemployment (higher employment) rate of these agents is then consistent with ideas such as Ronald S. Burt's (1992) structural holes, although for different reasons than the power reasoning behind Burt's theory.

four-agents clusters, meaning that the latter is more closely knit than the former. Indeed, the average path length decreases when the span of network contacts spreads; that is, when the relationships get less introverted or less closely knit.[12] The average unemployment increases with closed-knittedness, reflecting the fact that the wider the breadth of current social ties, the more diversified are the sources of information.

The fact that the long-run employment status of path-connected agents is positively correlated in the above examples is something that holds generally. In particular, as we divide a and b both by some larger and larger factor—so that we are looking at arbitrarily short time periods—then we can begin to sort out the short- and longer-run effects. We call this the "subdivision" of periods. In the limit we approximate a continuous time (Poisson) process, which is the natural situation where such temporary competition for jobs is short lived and inconsequential, while the overall status of indirect connections tells one a great deal about the possible status of direct connections, and the longer-run effects come to dominate.

PROPOSITION 1: *Under fine enough subdivisions of periods, the unique steady-state long-run distribution on employment is such that the employment statuses of any path-connected agents are positively correlated.*[13]

The proposition shows that despite the short-run conditional negative correlation between the employment of competitors for jobs and information, in the longer run any interconnected agents' employment is positively correlated. This implies that there is a clustering of agents by employment status, and employed workers tend to be connected with employed workers, and vice versa. This is consistent with the sort of clustering observed by Topa (2001). The intuition is clear: conditional on knowing that some set of agents are employed, it is more likely that their neighbors will end up receiving information about jobs, and so on.

Moreover, the positive correlation holds not only under the steady-state distribution, but also across any arbitrary time periods. That is, agent i's employment status at time t is positively correlated with agent j's status at time t' for general values of t and t'.

PROPOSITION 2: *Under fine enough subdivisions of periods, starting under the steady-state distribution, the employment statuses of any two path-connected agents are positively correlated across arbitrary periods.*[14]

III. Duration Dependence and Persistence in Unemployment

As mentioned in the introduction, there are some other patterns of unemployment that have been observed in the data and are exhibited by a networked model. To see this, let us examine some of the serial patterns of employment that emerge.

Again, consider job arrival and breakup rates

[12] Here, close-knittedness is measured by the average path length, which is just the average across all pairs of agents of the length of the shortest path between them. An alternative measure of the tightness of a network is, for instance, the clustering coefficient, which reflects the level of intraconnectedness among agents with a common friend. Both measures are roughly equivalent, though in some cases, when network links are randomly rewired, the average path length drops sharply, in contrast with a lower corresponding decrease of the clustering coefficient, a phenomenon often termed a "small world" effect (see, e.g., Duncan J. Watts and Steven H. Strogatz, 1998).

[13] More formally, any network g on the population of n agents, arrival probability $a \in (0, 1)$ and breakdown probability $b \in (0, 1)$ define a finite-state irreducible and aperiodic Markov process $\mathcal{M} = (g, a, b)$ on employment statuses. Denote by μ the (long-run) steady-state distribution associated to this process, which is uniquely defined. By dividing a and b by some common T, we obtain an associated Markov process $\mathcal{M}^T = (g, a/T, b/T)$, that we name the $T -$ *period subdivision* of $\mathcal{M}$, with steady-state distribution μ^T. We show that there exists some T' such that, for all $T \geq T'$, the employment statuses of any path-connected agents are positively correlated under μ^T.

[14] More formally, we show that there exists some T' such that $\mathrm{Cov}(S_{it}, S_{jt'}) > 0$ for any path-connected i and j and periods t and t', under the Markov process $\mathcal{M}^T$ for any subdivision of $T \geq T'$. "Starting under the steady-state distribution" means that the starting state is randomly drawn according to the steady-state distribution, and then all expectations account for the dependence of the process on this initial randomness. This is necessary, as Example 1 shows that starting from some particular states, one cannot escape short-run negative correlation. Starting from the steady-state distribution is "as if" our expectations are taken where we let the process run for a long time and average over all dates that are $t - t'$ apart from each other.

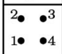
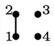
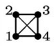
g	1 period	2 periods	10 periods	limit
	0.099	0.099	0.099	0.099
	0.176	0.175	0.170	0.099
	0.305	0.300	0.278	0.099

FIGURE 6. DURATION DEPENDENCE

of $a = 0.100$ and $b = 0.015$.[15] Ask the following question: suppose that a person has been unemployed for at least each of the last X periods. What is the probability that he or she will be employed at the end of this period? We examine the answer to this question in Figure 6 as we vary the number of periods of observed past unemployment and the network. The entries represent the probability that agent 1 will be employed next period, conditional upon the network structure and having been unemployed for at least the given number of consecutive previous periods.

In the case where there is no network, the probability of becoming employed is independent of the observed history—it is simply the probability that an agent hears about a job (and then does not lose it in this period). Once there is a network in place, the length of the unemployment history does give us information: it provides insight into the likelihood that the agents' connections are unemployed.

The patterns observed here are not particular to the example but hold more generally.

PROPOSITION 3: *Under fine enough subdivisions of periods and starting under the steady-state distribution, the conditional probability that an individual will become employed in a given period is decreasing with the length of their observed* (*individual*) *unemployment spell.*

Indeed, longer past unemployment histories lead to worse inferences regarding the state of one's connections and the overall state of the network. This leads to worse inferences regarding the probability that an agent will hear indirect news about a job. That is, the longer i has been unemployed, the higher the expectation that i's connections and path connections are themselves also unemployed. This makes it more likely that i's connections will take any information they hear of directly, and less likely that they will pass it on to i. In other words, a longer individual unemployment spell makes it more likely that the state of one's social environment is poor, which in turn leads to low forecasts of future employment prospects.

As we mentioned in the introduction, this explanation for duration dependence is complementary to many of the previous explanations. For instance, one (among a number of) explanations that has been offered for duration dependence is unobserved heterogeneity.[16] A simple variant of unobserved heterogeneity is that agents have idiosyncratic features that are relevant to their attractiveness as an employee and are unobservable to the econometrician but observed by employers. With such idiosyncratic features some agents will be quickly reemployed while others will have longer spells of unemployment, and so the duration dependence is due to the unobserved feature of the worker. While the network model also predicts duration dependence, we find that over the lifetime of a single worker, the worker may have different likelihoods (which are serially correlated) of reemployment depending on the current state of their surrounding network. So, it also predicts that controlling for the state of the network should help explain the duration dependence. In particular, it offers an explanation for why workers of a particular type in a particular location (assuming networks correlate with location) might experience different employment characteristics than the same types of workers in another location, all other variables held constant. So for example, variables such as location that capture network effects should interact with

[15] These calculations are also from simulations, where here we can directly calculate these conditional probabilities, by looking at conditional frequencies on observed strings of unemployment of X periods long. The limit numbers are obtained analytically, and are simply the same as having no network.

[16] Theoretical models predicting duration dependence, though, are a bit scarcer. In Olivier J. Blanchard and Peter Diamond (1994), long unemployment spells reduce the reemployment probability through a stigma effect that induces firms to hire applicants with lower unemployment durations (see also Tara Vishwanath, 1989, for a model with a stigma effect). In Christopher A. Pissarides (1992), duration dependence arises as a consequence of a decline in worker skills during the unemployment spell.

TABLE 1—PROBABILITY OF FINDING EMPLOYMENT FOR AGENTS IN THE BRIDGE NETWORK

Number of employed	0	1	2	3	4	5	6	7	8	9
$a = 0.100$; $b = 0.015$	10.0	10.4	12.0	14.5	17.9	20.7	25.4	25.7	28.7	34.4
$a = 0.050$; $b = 0.050$	5.0	5.9	6.2	6.9	8.6	9.3	11.3	12.2	15.0	18.5

other worker characteristic variables which would not be predicted by other models.[17]

A. *Comments on Stickiness in the Dynamics of Employment*

The duration dependence for individuals is reflective of a more general persistence in employment dynamics. This persistence can be understood by first noting a simple feature of our model. When aggregate employment is relatively high, unemployed agents have relatively more of their connections employed and face relatively less competition for job information, and are more likely to hear about jobs. Conversely, when aggregate employment is relatively low, unemployed agents are relatively less likely to hear about jobs.

To illustrate this point, consider the bridge network in Figure 4 that connects ten agents. We calculate the (average) individual probability that an unemployed agent finds a job within the current period, conditional on the total number of employed agents in the network. We provide calculations for two different pairs of parameter values, $a = 0.100$, $b = 0.015$, and $a = b = 0.050$. The probabilities are expressed in percentage terms.

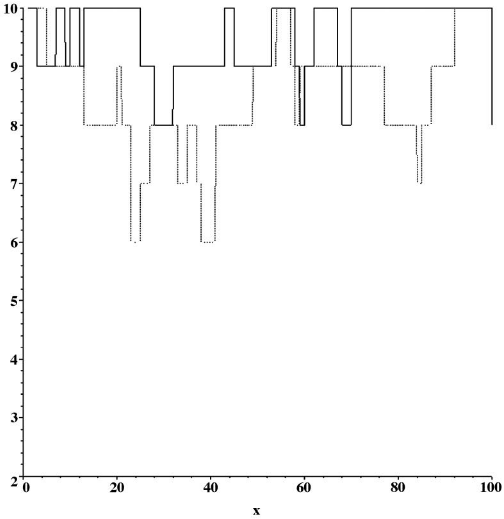
When there is no network connecting agents, the probability that an unemployed agent finds a job is simply the arrival rate a. In contrast, when agents are connected through a network (here, the bridge network of Figure 4), the probability of finding a job varies with the employment state. This conditional probability is a when everybody is unemployed, but then increases with the number of employed agents in the network, as shown in Table 1.

This state dependence of the probability of hearing about a job, then implies a persistence in aggregate employment dynamics. As a network gets closer to full employment, unemployed agents become even more likely to become employed. Symmetrically, the lower the employment rate, the lower the probability that a given unemployed agent hears about a job.[18] Although the process oscillates between full employment and unemployment, it exhibits a stickiness and attraction so that the closer it gets to one extreme (high employment or high unemployment) the greater the pull is from that extreme. This leads to a sort of boom and bust effect, as illustrated in Figure 7.

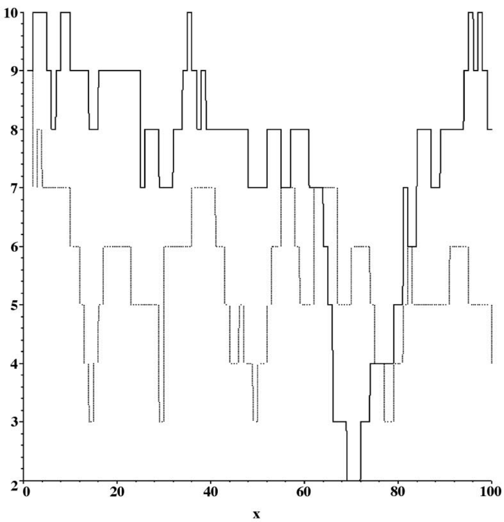
Starting from full employment, Figure 7 plots the dynamics of a simulation of aggregate employment over 100 periods for an empty network (the dotted line) and for the bridge network of Figure 4 (the plain line). We ran the dynamics for two different parameter pairs. First, when $a = 0.100$ and $b = 0.015$, the economy oscillates between full and high employment in the bridge network while it oscillates more widely between high and low employment in the empty network. An important feature is that the spells of unemployment are shorter in the bridge network; this is reflective of the fact that unemployed agents hear about jobs more quickly when the economy is near full employment. The fact that the arrival rate is relatively high compared to the breakup rate means that the bridge network stays very close to full employment most of the time. In the second simulation, $a = 0.050$ and $b = 0.050$, and so the arrival rate and breakup rates are equal, and the economy oscillates more widely between high and low employment in both networks. Still the empty network experiences lower average employment as we should expect; but more importantly, the bridge network snaps back to full employment more

[17] We thank Eddie Lazear for pointing this out to us.

[18] We have not explicitly modeled equilibrium wages and the job arrival process. Incorporating these effects might mitigate some of the effects our model exhibits. However, taking the arrival process as exogenous helps us show how the network pushes the process to have certain characteristics. See Randall Wright (1986) for a search model that generates fluctuating dynamics in a proper market setting.



Aggregate employment over time for a=0.100 and b=0.015

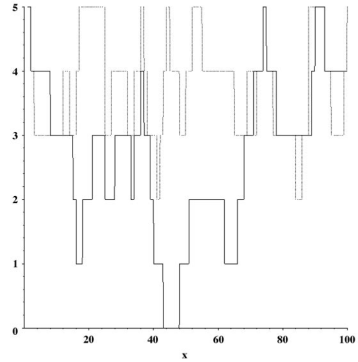


Aggregate employment over time for a=0.050 and b=0.050

FIGURE 7. TIME SERIES OF EMPLOYMENT FOR NETWORKED VERSUS DISCONNECTED AGENTS

quickly when pushed away from it. In line with the stickiness of the process, note that the only situation where more than two agents are unemployed for more than five periods appears between periods 60 and 80, and there we hit the lowest employment overall—which is illustrative of the relationship between level of unemployment and duration.

We also point out that employment need not be evenly spread on the network, especially in a network such as the bridge network from Figure 4. As a result temporal patterns may be asynchronous across different parts of the network, with some parts experiencing booms and other parts experiencing recessions at the same time. This asynchronous behavior is illustrated in Figure 8, which plots separately over 100 periods the aggregate employment of agents 1 to 5 (the dotted line) and that of agents 6 to 10 (the plain line) in the bridge network from Figure 4 from a simulation with $a = 0.050$ and $b = 0.050$.



Aggregate employment of agents 1 to 5 versus agents 6 to 10 over time for a=0.050 and b=0.050

FIGURE 8. ASYNCHRONOUS PATTERNS OF EMPLOYMENT ACROSS NETWORK SECTIONS

IV. Dropping Out and Inequality in Employment

We now turn to showing how the network model has important implications for inequality across agents, and how that inequality can persist.

Our results so far show that an agent's employment status will depend in important ways on the status of those agents who are path-connected with the agent. This leads to some heterogeneity across agents, as their networks and the local conditions in their networks will vary. Note, however, that in the absence of some structural heterogeneity across agents, their long-run prospects will look similar. That is, if the horizon is long enough, then the importance of the starting state will disappear.

However, expanding the model slightly can introduce substantial and sustained inequality among agents, even if their network structures are identical. The expansion in the model comes in the form of endogenizing the network by allowing agents to "drop out" of the network. This decision can be sensitive to starting conditions, and have lasting and far-reaching effects on the network dynamics. Let us take a closer look.

Suppose that agents have to decide whether to be in the labor market network or to drop out. We model this as a once-and-for-all decision, with no reentry allowed for agents choosing to drop out. Staying in the labor market results in an expected present value of costs $c_i \geq 0$. These include costs of education, skills maintenance, opportunity costs, etc. We normalize the outside option to have a value of 0, so that an agent chooses to stay in the labor force when the discounted expected future wages exceed the costs.

There are two simplifications made here. One is that we model this as if all agents make their decisions simultaneously. In reality although the bulk of a given agent's connections may be at similar stages of their lives, the agent may also be connected to other agents at different stages in their lives. The second simplification is that the drop-out decision is made just once. Implicit in this is the idea that the bulk of the costs (education and opportunity) appear at an early stage of an agent's career, and once those costs are sunk there is little incentive to drop out. These are both clearly crude approximations, but reasonable starting points.

There are some obvious comparative statics. Drop-out percentages will be decreasing in wages and increasing in costs. Drop-out decisions also depend on how well an agent is connected. With better connections (for instance, larger numbers of links holding all else constant), there is a larger chance of hearing about jobs and so the future prospects of employment are higher, leading to a higher threshold of costs where the agent would choose to drop out.

The part that is less transparent, but still quite intuitive, is the interaction between the decisions of different agents. Positive correlation of employment for path-connected agents both within and across periods implies that having more agents participate is better news for a given agent as it effectively improves the agent's network connections and prospects for future employment. Therefore, the decisions to stay in the labor force are strategic complements, implying that the drop-out game is supermodular. That is, as more of the other players decide to stay in, a given player's decision to stay in is increasingly favored relative to dropping out. The theory of supermodular games then guarantees the existence of a *maximal* Nash equilibrium in pure strategies. A maximal equilibrium is such that the set of agents staying in the market is maximal, so that the set of agents staying in at any other equilibrium of the game is a subset of those staying in at the maximal equilibrium. We restrict attention to such maximal equilibria.[19,20]

This supermodular aspect of the drop-out decisions is where we see the emergence of the contagion effects discussed in the introduction. The fact that an agent drops out leads to worse future employment prospects for that agent's connections. This in turn increases the chance that those agents drop out, and so forth. Thus, drop-out decisions are not independently and identically distributed (i.i.d.), even when the costs of staying in the labor force are i.i.d. across agents. This effect, as well as how the initial condition of the network affects drop-out rates, are illustrated in the following example.

Example 5 (Initial Conditions, Dropouts, and Contagion): To measure the contagion effect, we first ask how many people would drop out without any equilibrium effect, that is, if they each did the calculation supposing that everyone else was going to stay in. Then we can calculate how many people drop out in equilibrium, and any extra people dropping out are due to somebody else dropping out, which is what we attribute to the contagion effect.

For these calculations, we take the cost of staying in the network, c_i, to be uniformly distributed on [0.8, 1] and fix the per period wage to be $w = 1$. We do the calculations with complete networks, where each participating agent is directly linked to every other agent. We

[19] Formally, let $d_i \in \{0, 1\}$ denote i's decision of whether to stay in the labor market, where $d_i = 1$ stands for staying in. A vector of decisions, $\mathbf{d^{**}}$, is a maximal equilibrium if it is an equilibrium and, for every other equilibrium $\mathbf{d^*}$, we have $\mathbf{d^{**}} \geq \mathbf{d^*}$, where $\geq$ is the componentwise ordering on $\{0, 1\}^n$.

[20] There is a coordination game going on (players may all wish to drop out if all others do, etc.), and here looking at the maximal equilibrium eliminates coordination errors, and allows us to focus on the network effects of the model.

TABLE 2—DROPOUTS AND CONTAGION—STARTING EMPLOYED

$\mathbf{s}_0 = (1, \ldots, 1)$	$n = 1$	$n = 2$	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n \to \infty$
Drop-out percentage	58.3	44.5	26.2	14.7	9.7	7.8	6.8
Percentage due to contagion	0	8.8	5.0	1.4	0.4	0.2	0

TABLE 3—DROPOUTS AND CONTAGION—STARTING UNEMPLOYED

$\mathbf{s}_0 = (0, \ldots, 0)$	$n = 1$	$n = 2$	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n \to \infty$
Drop-out percentage	100	97.8	92.9	82.2	68.0	60.6	56.8
Percentage due to contagion	0	12.1	21.7	18.9	8.7	3.0	0

compute the percentage of dropouts for different values of n, and we also calculate the percentage of dropouts due to contagion. We do the calculation for two initial states: everybody employed, $\mathbf{s}_0 = (1, \ldots, 1)$, and everybody unemployed, $\mathbf{s}_0 = (0, \ldots, 0)$.

For Tables 2 and 3, the calculations are done for a discount rate of 0.9, where we simplify things by assuming that an agent starts in the initial state, and then jumps to the steady state in the next "period." This just gives us a rough calculation, but enough to see the effects. So, an agent who stays in gets a net payoff of $0.1s_i + 0.9p_i - c_i$, where p_i is the agent's steady-state employment probability in the maximal equilibrium. We again set $a = 0.100$ and $b = 0.015$.[21]

So, for instance, in Table 3, when $n = 16$ and everybody is initially unemployed, we have 68 percent of the people dropping out on average. This means that we expect about 11 people to drop out on average and about 5 people to stay in. The 8.7 percent due to contagion means that about 1.5 ($=0.087 \times 16$) of the people dropping out are doing so because others drop out, and they would be willing to stay in if all the others were willing to. Thus about 9.5 of the 11 people would drop out even if all stayed in, and 1.5 of the 11 drop out because of the fact that some others have dropped out.

Note that the contagion effect is larger for the worse starting state and is also larger for smaller networks (although not entirely monotone). This is true because the impact of someone dropping out is more pronounced in worse starting states and smaller networks. In the limit, the impact of having people drop out is negligible and so the contagion effect disappears when agents have very large numbers of connections (holding all else fixed). For $n = 1$, there cannot be a contagion effect, so the number is 0 there as well.

The nonmonotonicity of the contagion effect in n is a bit subtle. The possibility of contagion is initially nonexistent. It then increases as the number of connections increases, since there are more possible combinations of neighbor dropouts that can take place with three connections (when $n = 4$) than one connection (when $n = 2$), and any single dropout can then trigger another. Eventually, with large numbers of connections, the marginal impact of an additional connection to a given agent is already very low, and in fact becomes second order in the number of agents. The fact that it shrinks so much means that eventually the contagion effect disappears as even having some fraction of one's connections drop out is no longer a problem if there are still a large number of connections staying in.

The previous example shows that different social groups with identical network relationships but differing by their starting employment

[21] In these calculations we estimate p_i by an approximation formula, to save on calculations, as we need to iterate on both the c_i's and the number of agents. The approximation formula leads to some slight biases. In the simulations, for any given n, we first randomly draw the c_i's. We then calculate the p_i's if all agents stay in. From this we can find out which agents would drop out. We can then recalculate the p_i's for the new network, and see which agents drop out with the new p_i's. We continue this process until no additional agents drop out. This provides the maximal equilibrium for this draw of c_i's. We then run iterations on this algorithm for new draws of c_i's and calculate sample averages over the iterations.

state, have different drop-out rates. Because dropping out hurts the prospects for the group further, this can have strong implications for inequality patterns. We now show that this holds more generally.

In analyzing the drop-out game for the proposition below we do not alter the network structure when an agent drops out; instead we simply set the dropout's employment status to be 0 forever. Thus a dropout's social contacts do not change and his or her contacts still pass on news about jobs. It is as if the agent's connections still pass the agent information thinking that the agent is simply unemployed rather than a dropout. This might introduce a bias to the extent that information is no longer passed on to someone who has dropped out. Accounting for such effects would complicate the analysis as one would need to keep track of any modifications of the network structure as agents drop out.

PROPOSITION 4: *Consider two social groups with identical network structures. If the starting state person-by-person is higher for one group than the other, then the set of agents who drop out of the first group in the maximal equilibrium is a subset of their counterparts in the second group. These differences in drop-out rates generate persistent inequality in probabilities of employment in the steady-state distributions, with the first group having weakly better employment probabilities than their counterparts. There is a strict difference in employment probabilities for all agents in any component of the network for which the equilibrium drop-out decisions differ across the two groups.*

So we have established that a networked model can generate persistent differences among two social groups with identical economic characteristics except that they differ in their starting state. As mentioned in the introduction, this is consistent with documented differences in drop-out rates among blacks and whites, as well as studies that show that accounting for voluntary dropouts from the labor force negatively affect the standard measures of black economic progress (e.g., Chandra, 2000, Heckman et al., 2000). While this comparison is stylized, the fact that we consider two completely identical networks except for their starting states emphasizes how important starting conditions can be.[22]

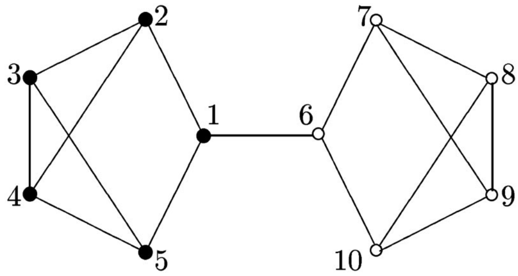


FIGURE 9. THE BRIDGE NETWORK WITH ASYMMETRIC STARTING STATES

Just to show that the inequality we are seeing is not due to the isolation of the groups of agents with different starting conditions, let us examine drop-out decisions when there is a bridge between two groups.

Example 6 (Connected Social Groups and Dropouts): Consider the network structure from Example 3; see Figure 9.

Agents 1 to 5 start employed and agents 6 to 10 start unemployed.

We do drop-out calculations as in Example 5. We take the c_i to be uniformly distributed on [0.8, 1], fix $w = 1$, use a discount rate of 0.9, and have agents who stay in get a net payoff of $0.1s_i + 0.9p_i - c_i$, where p_i is the agent's steady-state employment probability in the maximal equilibrium of the drop-out game, and s_i is their starting employment state.

The drop-out probabilities for the different agents are illustrated in Table 4.

Note that the drop-out rates are significantly higher for the agents who start in the unemployed state, even though the network is connected. The agent who forms the bridge from the unemployed side is less likely to drop out than the other unemployed agents: while the counterpart agent who forms the bridge from

[22] Differences in network structure, rather than differences in starting conditions, can also lead to different drop-out decisions and sustained inequality. For instance, as we saw in Example 2, under exactly the same *a* and *b* and with each agent having two links, the expected long-run unemployment of an agent in a network of four agents is 6.3 percent while it is 6 percent for an agent in a network of eight agents. While the difference in this example is small (on the order of a 5-percent change in unemployment), it can easily become amplified through the contagion effect.

TABLE 4—DROP-OUTS RATES IN THE BRIDGE NETWORK WITH ASYMMETRIC STARTING STATES

Agent	1	2	3	4	5	6	7	8	9	10
Drop-out rate	0.47	0.42	0.42	0.42	0.42	0.91	0.93	0.93	0.93	0.93

the employed side is more likely to drop out than other employed agents.[23] While this example is clearly highly stylized, it does provide some intuitive predictions, and shows that starting conditions can lead to sustained differences across groups even in a connected network.

Let us discuss the relation of our analysis of drop-out rates to theories of discrimination. Classical theories of discrimination, such as that of Gary Becker (1957) or Thomas C. Schelling (1971), postulate that individuals have an intrinsic preference for individuals in their own societal group.[24] Because of such preferences and externalities, individuals end up segregated in the workplace, and the resulting sorting patterns by group affiliation can breed wage inequality.[25] Our model offers an alternative and novel explanation for inequality in wages and employment.[26,27] Two otherwise identical individuals embedded in two societal groups with different collective employment histories (or with different networks as discussed below) typically will experience different employment outcomes. In other words, social networks influence economic success of individuals at least in part due to the different composition and history of individuals' networks. When coupled with drop-out decisions, sustained inequality can be the result of differences in history. We discuss some policy implications of this network viewpoint below.

V. A Look at Policy Implications

Let us mention some lessons that can be learned from our model about policy in the presence of network effects, and some of which we will illustrate with some examples below. One obvious lesson is that the dynamics of the model show that policies that affect current employment will have both delayed and long-lasting effects.

Another lesson is that there is a positive externality between the status of connected individuals. So, for instance, if we consider improving the status of some number of individuals who are scattered around the network, or some group that are more tightly clustered, there will be two sorts of advantages to concentrating the improvements in tighter clusters. The first is that this will improve the transition probabilities of those directly involved, but the second is that this will improve the transition probabilities of those connected with

[23] There may also be differences between the drop-out rates of agents 2 and 3, or 8 and 9, as things are not symmetric; but these differences are beyond the accuracy of our simulations.

[24] There is also an important literature on "statistical" discrimination that follows Arrow (1972), John J. McCall (1972), Edmond S. Phelps (1972), and others. Our work is quite complementary to that work as well.

[25] We use the word "can" because it may be that some employers discriminate while the market wages do not end up unequal. As Becker (1957) points out, the ultimate outcome in the market will depend on such factors as the number of nondiscriminating employers and elasticities of labor supply and demand.

[26] While we have not included "firms" in our model, note that to the extent to which the job information comes initially from an employee's own firm, there would also be correlation patterns among which firms connected agents work for. That is, if an agent's acquaintance is more likely to be getting information about job openings in the acquaintance's own firm, then that agent has a more than uniformly random likelihood of ending up employed in the acquaintance's firm. This would produce some segregation patterns beyond what one would expect in a standard labor market model.

[27] Two other important explanations for inequality can be found in Glenn C. Loury (1981) and Steven Durlauf (1996). As in our model, both papers relate social background to individual earning prospects. In Loury's paper, the key aspect of social background is captured by family income which then determines investment decisions in education. In Durlauf's work, education is modeled as a local public good, and community income, rather than family incomes, affects human capital formation. In both cases, because the social background imposes constraints on human capital investment, income disparities are passed on across generations. In our paper, we focus instead on the larger societal group within which one is embedded, its network structure, collective employment history, and access to information about jobs. This offers a complementary, rather than competing, explanation for sustained inequality.

these individuals. Moreover, concentrated improvements lead to a greater improvement of the status of connections than dispersed improvements. This will then propagate through the network.

To get a better picture of this, consider the drop-out game. Suppose that we are in a situation where all agents decide to drop out. Consider two different subsidies: in the first, we pick agents distributed around the network to subsidize; while in the second we subsidize a group of agents who are clustered together. In the first case, other agents might now just have one (if any) connection to an agent who is subsidized. This might not be enough to induce them to stay in, and so nobody other than the subsidized agents stay in the market. Here the main impact on the drop-out rate is directly through the subsidy. In contrast, in the clustered case, a number of agents now have several connections who are in the market. This may induce these other agents to stay in. This can then have a contagion effect, carrying over to agents connected with them and so on. Beyond the direct impact of the subsidy, this has an indirect contagion effect that decreases the drop-out rate, and then improves the future status of all of the agents involved even further through the improved network effect.

Exactly how one wants to distribute subsidies to maximize their impact is a subtle matter. We look at an example to highlight the subtleties.

A. *Concentration of Subsidies*

Let us again consider a society of eight individuals, again where $a = 0.100$ and $b = 0.015$. Suppose the costs of staying in the network, c_i, are drawn at random from a uniform distribution with support [0.8, 1]. Initially, everybody is unemployed, so $\mathbf{s}_0 = (0, \ldots, 0)$. We work with drop-out decisions when the discount rate is 0.9, as in the previous examples.

The experiment we perform here is the following. In each case we subsidize two agents to stay in the market—simply by lowering their cost c_i to 0.[28] The question is which two agents we subsidize. In the network, each agent has four connections. The network structure is as follows. Each agent has three links—two immediate neighbors and one that is slightly further away. This is pictured in Figure 10.

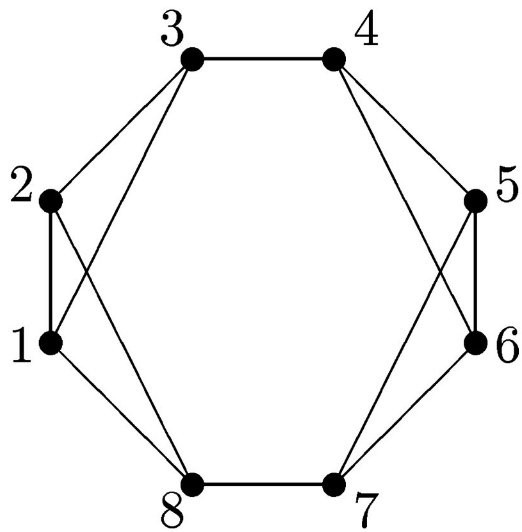


FIGURE 10. THE STARTING NETWORK STRUCTURE

Table 5 provides the percentage of agents who stay in the network as a function of who is subsidized (two agents in each case) and what the range of costs (randomly drawn) are.[29]

There are some interesting things to note.

In the highest cost range, even having one neighbor stay in is not enough to induce an agent to stay, and so the only agents staying in are the subsidized ones. Here it is irrelevant which agents are subsidized as they are the only ones staying in.

In the lowest two cost ranges, having one neighbor stay in has a big impact, and so spreading the subsidies out has the maximal impact. Agents 1 and 5 are on opposite ends of the circle and have no direct contact in common. Subsidizing agents 1 and 5 thus amounts for spreading subsidies out, and it is indeed the best policy in terms of maximizing

[28] This might in fact overestimate the necessary subsidy costs, as one might not need to lower c_i by so much. Moreover, we would end up with even lower subsidy costs by concentrating subsidies which increases the future prospects of the subsidized agents. Thus, lower costs of subsidization would be a further reason for concentrating subsidies.

[29] Note that the different cases of who are subsidized cover all possible configurations, up to a relabeling of the agents.

TABLE 5—SUBSIDIZATION STRUCTURE AND PERCENTAGE OF AGENTS WHO STAY IN

	Cost range			
Agents subsidized	0.80 to 1	0.82 to 1	0.84 to 1	0.86 to 1
1 and 2	52.9	39.4	27.8	25.0
1 and 3	53.6	39.4	27.1	25.0
1 and 4	57.2	43.4	27.9	25.0
1 and 5	**57.9**	43.8	27.0	25.0
1 and 6	57.9	**44.0**	27.0	25.0
1 and 7	57.1	43.4	27.8	25.0
1 and 8	53.5	39.4	27.1	25.0
3 and 4	54.5	39.3	26.1	25.0
3 and 7	57.7	43.6	27.4	25.0
3 and 8	56.2	42.9	**29.1**	25.0

the number of agents who stay in the market when the cost is at its lowest level.[30] When the cost is in the 0.82 to 1 range, we begin to see a slight change, where now subsidizing agent 1 and 6 is better, and these agents are slightly closer together.

The places where things favor a different sort of policy is in the middle range of costs. Here costs are high enough so that it is quite likely that an agent will drop out if she has only one neighbor who stays in, but is less likely if she has two neighbors who stay in, and so contagion effects are relatively high. Spreading the subsidies out to agents 1 and 5, or 3 and 7, etc., does worse than having them close together (1 and 2, 1 and 3, 1 and 4) and the best possible subsidy here is of the form 3 and 8. What matters is the number of contacts subsidized agents have in common. Agents 3 and 8 are well-placed since both 1 and 2 are connected to both of them. Thus, this concentrates subsidies in a way that provides a high probability that 1 and 2 will stay in. Without such a concentration of subsidies, we get a higher drop-out rate.

What this suggests is that, in designing subsidy or affirmative action programs, attention to network effects is important. Concentrating efforts more locally can end up having a higher or lower impact, depending on the network configuration.

[30] It is almost a tie with 1 and 6, but slightly ahead in the next decimal.

VI. Possible Empirical Tests

While as we have discussed, the model generates patterns of employment and dropouts that are consistent with the stylized facts from a number of studies, one might want to look at some additional and more direct tests of the model's predictions.

Note that drop-out rates and contagion effects depend both on the costs ranges and on the values for the arrival rate and breakup rate. Some comparative statics are quite obvious: (1) as the expected cost increases (relative to wages), the drop-out rate increases; (2) as the breakup rate increases, the drop-out rate increases; and (3) as the arrival rate increases, the drop-out rate decreases. However, there are also some more subtle comparisons that can be made.

For instance, let us examine what happens as job turnover increases. Here, as the arrival and breakup rates are both scaled up by the same factor, we can see the effects on the drop-out rates. Note that such a change leaves the base employment rate (that of an isolated agent) unchanged, and so the differences are attributable entirely to the network effects. Table 6 pulls out various rescalings of the arrival and breakup rates for the two cost ranges when $n = 4$ and agents are related through a complete network. As before, the first figure is the drop-out rate and the second is the amount attributable to contagion effects.

As we can see, higher turnover rates (higher rescalings of a and b) lead to higher drop-out rates. The intuition behind this is as follows.

TABLE 6—DEPENDENCE OF DROPOUTS AND CONTAGION ON ARRIVAL AND BREAKUP RATES

Scaled by a and b	1	3	5	7	9
	0.05, 0.015	0.15, 0.045	0.25, 0.075	0.35, 0.105	0.45, 0.135
$c_i \sim [0.8, 1]$	69:27	76:27	83:26	88:24	96:20
$c_i \sim [0.6, 1]$	24:3	28:3	34:5	37:5	42:5

With higher turnover rates, when an agent becomes unemployed it is more likely that some of his neighbors are unemployed. This is actually quite subtle, as the base unemployment rate has not changed. However, higher turnover makes it more likely that several agents lose their jobs at the same time, and end up competing for information. This effect then lowers the employment rate, which in turn feeds back and results in less valuable connections.

This effect provides for a testable implication: industries with higher turnover rates, all else held equal, should have higher drop-out rates. Operationalizing this requires some care, however, as we do not model the career choices for workers or an equilibrium in wages. Nonetheless, it is clear that the prediction is that the wage to education cost ratio must be relatively higher in order to induce workers to enter careers where the turnover rate is high compared to those where it is low, even after correcting for any risk aversion or income-smoothing motives.

Let us briefly mention some other possible empirical tests of the model. To the extent that direct data on network relationships is available, one can directly test the model. In fact, such information in the form of survey data (the General Social Survey) has been used extensively in the sociology literature and also in conjunction with wage data (e.g., Troy Tassier, 2001).

There are also other tests that are possible. For instance, there is data concerning how the reliance on networks for finding jobs varies across professions, age, and race groups, etc. (see the table in Montgomery, 1991, for instance, to see some differences across professions). Our model then predicts that the intensity of clustering, duration dependence, and drop-out rates should also vary across these socioeconomic groups. Moreover, even within a specific socioeconomic group, our model predicts differences across separate components of the network as the local status of the connections changes.

VII. Concluding Discussion

As we have mentioned several times, we treat the network structure as given, except that we consider drop-out decisions. Of course, people have more specific control over whom they socialize with both in direct choice of their friendships, as well as through more indirect means such as education and career choices that affect whom they meet and fraternize with on a regular basis. Examining the network formation and evolution process in more detail could provide a fuller picture of how the labor market and the social structure co-evolve by mutually influencing each other: network connections shape the labor market outcomes and, in turn, are shaped by them.[31]

In addition to further endogenizing the network, we can also look deeper behind the information exchange procedure.[32] There are a wide variety of explanations (especially in the sociology literature, for instance see Granovetter, 1995) for why networks are important in job markets. The explanations range from assortive matching (employers can find workers with

[31] There is a growing literature on the formation of networks that now provides a ready set of tools for analyzing this problem. See Jackson (2004) for a survey of models applying to networks of the form analyzed here, as well as Sanjeev Goyal (2004), Frank Page (2004), and Anne van den Nouweland (2004) for issues relating to learning, farsightedness, and cooperation structures.

[32] Recall that the results stated so far extend to a framework more general than the simple communication protocol where only direct contacts can communicate with each other and unemployed direct contacts are treated on an equal footing. In particular, the general framework accommodates a priori ranking among contacts, indirect passing of job information, heterogeneous jobs with different wages, idiosyncratic arrival and breakup rates, information passing, and job turnover dependent on the overall wage distribution, etc. See Calvó-Armengol and Jackson (2003) for more details.

similar characteristics by searching through them), to information asymmetries (in hiring models with adverse selection), and simple insurance motives (to help cope with the uncertainty due to the labor market turnover). In each different circumstance or setting, there may be a different impetus behind the network. This may in turn lead to different characteristics of how the network is structured and how it operates. Developing a deeper understanding along these lines might further explain differences in the importance of networks across different occupations.

Another aspect of changes in the network over time is that network relationships can change as workers are unemployed and lose contact with former connections. Long unemployment spells can generate a desocialization process leading to a progressive removal from labor market opportunities and to the formation of unemployment traps. This is worth further investigation.

Another important avenue for extension of the model is to endogenize the labor market equilibrium so that the probability of hearing about a job depends on the current overall employment and wages are equilibrium ones. This would begin to give insights into how network structure influences equilibrium structure.

Finally, we point out that although our focus in this paper is on labor markets, this model can easily be adapted to other sorts of behaviors where social networks play a key role in information transmission. An example is whether or not individuals take advantage of certain available welfare programs. Recent studies by Marianne Bertrand et al. (2000) and Anna Aizer and Janet Currie (2002) point to the importance of social networks in such contexts.

Appendix

We first provide some definitions that are necessary for the proofs that follow. Here we specialize the definitions to random vectors $\mathbf{S} = (S_1, \ldots, S_n)$ whose components take on values of 0 or 1. We follow the convention of representing random variables by capital letters and realizations by small letters.

Association

While first-order stochastic dominance is well suited for capturing distributions over a single agent's status, we need a richer tool for discussing interrelationships between a number of agents at once, and this is needed in the proofs that follow. The following definition is first due to James D. Esary et al. (1967).

Let μ be a joint probability distribution describing $\mathbf{S}$.

μ is associated if $\text{Cov}_\mu(f, g) \geq 0$ for all pairs of nondecreasing functions $f : \{0, 1\}^n \to \mathbb{R}$ and $g : \{0, 1\}^n \to \mathbb{R}$, where $\text{Cov}(f, g)$ is the covariance $E_\mu[f(\mathbf{S})g(\mathbf{S})] - E_\mu[f(\mathbf{S})]E_\mu[g(\mathbf{S})]$.

If $S_1, \ldots, S_n$ are the random variables described by a measure μ that is associated, then we say that $S_1, \ldots, S_n$ are associated. Note that association of μ implies that S_i and S_j are nonnegatively correlated for any i and j. Essentially, association is a way of saying that all dimensions of $\mathbf{S}$ are nonnegatively interrelated.[33]

Strong Association

As we often want to establish strictly positive relationships, and not just nonnegative ones, we need to define a strong version of association. Since positive correlations can only hold between agents who are path-connected, we need to define a version of strong association that respects such a relationship.

Consider a partition Π of $\{1, \ldots, n\}$ that captures which random variables might be positively related; which here will be determined by the components of the graph g.

A probability distribution μ governing $\mathbf{S}$ is *strongly* associated relative to the partition Π if it is associated, and for any $\pi \in \Pi$ and nondecreasing functions f and g

[33] This is still a weaker concept than affiliation, which requires association for all conditionals. It is imperative that we work with the weaker notion as affiliation will not hold in our setting.

$$\mathrm{Cov}_\mu(f, g) > 0$$

whenever there exist i and j such that f is increasing in s_i for all s_{-i}, g is increasing in s_j for all s_{-j}, and i and j are in π.

One implication of strong association is that S_i and S_j are positively correlated for any i and j in π.

Domination

Consider two probability distributions μ and ν.

μ *dominates* ν if

$$E_\mu[f] \geq E_\nu[f]$$

for every nondecreasing function $f: \{0, 1\}^n \to \mathbb{R}$. The domination is *strict* if strict inequality holds for some nondecreasing f.

Domination captures the idea that "higher" realizations of the state are more likely under μ than under ν. In the case where $n = 1$, domination reduces to first-order stochastic dominance.

LEMMA 5: *Consider two probability distributions μ and ν on $\{0, 1\}^n$. μ dominates ν if and only if there exists a Markov transition function $\phi : \{0, 1\}^n \to \mathcal{P}(\{0, 1\}^n)$ [where $\mathcal{P}(\{0, 1\}^n)$ is set of all the probability distributions on $\{0, 1\}^n$] such that*

$$\mu(\mathbf{s}') = \sum_{\mathbf{s}} \phi_{\mathbf{ss}'} \nu(\mathbf{s}),$$

where ϕ is a dilation (that is $\phi_{\mathbf{ss}'} > 0$ implies that $\mathbf{s}' \geq \mathbf{s}$). Strict domination holds if $\phi_{\mathbf{ss}'} > 0$ for some $\mathbf{s}' \neq \mathbf{s}$.

Thus, μ is derived from ν by a shifting of mass "upwards" (under the partial order $\geq$) over states in $\mathbf{S}$. This lemma follows from Theorem 18.40 in Charalambos Aliprantis and Kim C. Border (2000).

Let

$$\mathcal{E} = \{E \subset \{0, 1\}^n | \mathbf{s} \in E, \mathbf{s}' \geq \mathbf{s} \Rightarrow \mathbf{s}' \in E\}.$$

$\mathcal{E}$ is the set of subsets of states such that if one state is in the event then all states with at least as high employment status (person by person) are also in. Variations of the following useful lemma appear in the statistics literature (e.g., see Section 3.3 in Esary et al., 1967). A proof of this version can be found in Calvó-Armengol and Jackson (2003).

LEMMA 6: *Consider two probability distributions μ and ν on $\{0, 1\}^n$.*

$$\mu(E) \geq \nu(E)$$

for every $E \in \mathcal{E}$, if and only if μ dominates ν. Strict domination holds if and only if the first inequality is strict for at least one $E \in \mathcal{E}$. The probability measure μ is associated if and only if

$$\mu(EE') \geq \mu(E)\mu(E')$$

for every E and $E' \in \mathcal{E}$. The association is strong (relative to Π) if the inequality is strict whenever E and E' are both sensitive to some $\pi \in \Pi$.[34]

[34] E is sensitive to π if its indicator function is. A nondecreasing function $f: \{0, 1\}^n \to \mathbb{R}$ is sensitive to $\pi \in \Pi$ (relative to μ) if there exist $\mathbf{s}$ and $\tilde{\mathbf{s}}_\pi$ such that $f(\mathbf{s}) \neq f(\mathbf{s}_{-\pi}, \tilde{\mathbf{s}}_\pi)$ and $\mathbf{s}$ and $\mathbf{s}_{-\pi}, \tilde{\mathbf{s}}_\pi$ are in the support of μ.

The proof of the following lemma is straightforward and omitted.

LEMMA 7: *Let μ be associated and have full support on values of* $\mathbf{S}$. *If f is nondecreasing and is increasing in S_i for some i, and g is a nondecreasing function which is increasing in S_j for some j, and* $\mathrm{Cov}_{\mu}(S_i, S_j) > 0$, *then* $\mathrm{Cov}_{\mu}(f, g) > 0$.

Fix $\mathcal{M} = (g, a, b)$. Let $\mathbf{P}^T$ denote the matrix of transitions between different $\mathbf{s}$'s under the T-period subdivision $\mathcal{M}^T = (g, a/T, b/T)$. So $\mathbf{P}^T_{\mathbf{ss}'}$ is the probability that $\mathbf{S}_t = \mathbf{s}'$ conditional on $\mathbf{S}_{t-1} = \mathbf{s}$. Let

$$\mathbf{P}^T_{\mathbf{s}E} = \sum_{\mathbf{s}' \in E} \mathbf{P}^T_{\mathbf{ss}'}.$$

LEMMA 8: *Consider an economy* $\mathcal{M} = (g, a, b)$. *Consider* $\mathbf{s}' \in \mathbf{S}$ *and* $\mathbf{s} \in \mathbf{S}$ *such that* $\mathbf{s}' \geq \mathbf{s}$, *and any* $t \geq 1$. *Then for all T and* $E \in \mathcal{E}$

$$\mathbf{P}^T_{\mathbf{s}'E} \geq \mathbf{P}^T_{\mathbf{s}E}.$$

Moreover, if $\mathbf{s}' \neq \mathbf{s}$ *then the inequality is strict for at least one E.*

PROOF OF LEMMA 8:

Let us say that two states $\mathbf{s}'$ and $\mathbf{s}$ are adjacent if there exists ℓ such that $\mathbf{s}'_{-\ell} = \mathbf{s}_{-\ell}$ and $s'_\ell > s_\ell$ (that is, $s'_\ell = 1$ and $s_\ell = 0$). We show that $\mathbf{P}^T_{\mathbf{s}'E} \geq \mathbf{P}^T_{\mathbf{s}E}$ for adjacent $\mathbf{s}$ and $\mathbf{s}'$, as the statement then follows from a chain of comparisons across such $\mathbf{s}'$ and $\mathbf{s}$.

Let ℓ be such that $s'_\ell > s_\ell$. By adjacency, $s'_i = s_i$, for all $i \neq \ell$.

Since $s'_k = s_k$ for all $k \neq \ell$, it follows by our definition of $p_{ij}(\mathbf{s})$ that $p_{ij}(\mathbf{s}') \geq p_{ij}(\mathbf{s})$ for all $j \neq \ell$ and for all i. These inequalities imply that $\mathrm{Prob}^T_{\mathbf{s}'}(\mathbf{S}_{-\ell,t})$ dominates $\mathrm{Prob}^T_{\mathbf{s}}(\mathbf{S}_{-\ell,t})$, where $\mathrm{Prob}^T_{\mathbf{s}}$ is the probability distribution conditional on $\mathbf{S}_{t-1} = \mathbf{s}$. Given that $1 = s'_\ell > s_\ell = 0$, we then have that $\mathrm{Prob}^T_{\mathbf{s}'}(\mathbf{S}_t)$ dominates $\mathrm{Prob}^T_{\mathbf{s}}(\mathbf{S}_t)$. The conclusion that $\mathbf{P}^T_{\mathbf{s}'E} \geq \mathbf{P}^T_{\mathbf{s}E}$ follows from Lemma 6.

To see the strict domination, consider $E = \{\tilde{\mathbf{s}} | \tilde{s}_\ell \geq s'_\ell\}$. Since there is a positive probability that ℓ hears 0 offers under $\mathbf{s}$, the inequality is strict.

Given a measure ξ on $\{0, 1\}^n$, let $\xi\mathbf{P}^T$ denote the measure induced by multiplying the $(1 \times n)$ vector ξ by the $(n \times n)$ transition matrix $\mathbf{P}^T$. This is the distribution over states induced by a starting distribution ξ multiplied by the transition probabilities $\mathbf{P}^T$.

LEMMA 9: *Consider an economy* $\mathcal{M} = (g, a, b)$ *and two measures* μ *and* ν *on* $\mathbf{S}$. *For all T, if* μ *dominates* ν, *then* $\mu\mathbf{P}^T$ *dominates* $\nu\mathbf{s}^T$. *Moreover, if* μ *strictly dominates* ν, *then* $\mu\mathbf{P}^T$ *strictly dominates* $\nu\mathbf{P}^T$.

PROOF OF LEMMA 9:

$$[\mu\mathbf{P}^T](E) - [\nu\mathbf{P}^T](E) = \sum_{\mathbf{s}} \mathbf{P}^T_{\mathbf{s}E}(\mu_{\mathbf{s}} - \nu_{\mathbf{s}}).$$

By Lemma 5 we rewrite this as

$$[\mu\mathbf{P}^T](E) - [\nu\mathbf{P}^T](E) = \sum_{\mathbf{s}} \mathbf{P}^T_{\mathbf{s}E}\left(\sum_{\mathbf{s}'} \nu_{\mathbf{s}'}\phi_{\mathbf{s}'\mathbf{s}} - \nu_{\mathbf{s}}\right).$$

We rewrite this as

$$[\mu\mathbf{P}^T](E) - [\nu\mathbf{P}^T](E) = \sum_{\mathbf{s}}\sum_{\mathbf{s}'} \nu_{\mathbf{s}'}\phi_{\mathbf{s}'\mathbf{s}}\mathbf{P}^T_{\mathbf{s}E} - \sum_{\mathbf{s}} \nu_{\mathbf{s}}\mathbf{P}^T_{\mathbf{s}E}.$$

As the second term depends only on s, we rewrite that sum on s' so we obtain

$$[\mu\mathbf{P}^T](E) - [\nu\mathbf{P}^T](E) = \sum_{\mathbf{s}'}\left(\sum_{\mathbf{s}} \nu_{\mathbf{s}'}\phi_{\mathbf{s}'\mathbf{s}}\mathbf{P}^T_{\mathbf{s}E} - \nu_{\mathbf{s}'}\mathbf{P}^T_{\mathbf{s}'E}\right).$$

Since ϕ is a dilation, $\phi_{\mathbf{s}'\mathbf{s}} > 0$ only if $\mathbf{s} \geq \mathbf{s}'$. So, we can sum over $\mathbf{s} \geq \mathbf{s}'$:

$$[\mu\mathbf{P}^T](E) - [\nu\mathbf{P}^T](E) = \sum_{\mathbf{s}'} \nu_{\mathbf{s}'}\left(\sum_{\mathbf{s}\geq\mathbf{s}'} \phi_{\mathbf{s}'\mathbf{s}}\mathbf{P}^T_{\mathbf{s}E} - \mathbf{P}^T_{\mathbf{s}'E}\right).$$

Lemma 8 implies that $\mathbf{P}^T_{\mathbf{s}E} \geq \mathbf{P}^T_{\mathbf{s}'E}$ whenever $\mathbf{s} \geq \mathbf{s}'$. Thus since $\phi_{s's} \geq 0$ and $\sum_{\mathbf{s}\geq\mathbf{s}'} \phi_{\mathbf{s}'\mathbf{s}} = 1$, the result follows.

Suppose that μ strictly dominates ν. It follows from Lemma 5 that there exists some $\mathbf{s} \neq \mathbf{s}'$ such that $\phi_{\mathbf{s}'\mathbf{s}} > 0$. By Lemma 8, there exists some $E \in \mathcal{E}$ such that $\mathbf{P}^T_{\mathbf{s}E} > \mathbf{P}^T_{\mathbf{s}'E}$. Then $[\mu\mathbf{P}^T](E) > [\nu\mathbf{P}^T](E)$ for such E, implying (by Lemma 6) that $\mu\mathbf{P}^T$ strictly dominates $\nu\mathbf{P}^T$.

We prove Proposition 1 and Proposition 2 as follows.

PROOF OF PROPOSITION 1:

Since $\mathbf{P}^T$ represents an irreducible and aperiodic Markov chain, it has a unique steady-state distribution that we denote by μ^T. The steady-state distributions μ^T converge to a unique limit distribution (see H. Peyton Young, 1993), which we denote μ^*.

Let $\bar{\mathbf{P}}^T$ be the transition matrix where the process is modified as follows. Let $p_i(\mathbf{s}) = \sum_{j\in N} p_{ji}(\mathbf{s})$. Starting in state $\mathbf{s}$, in the hiring phase each agent i hears about a new job (and at most one) with probability $p_i(\mathbf{s})/T$ and this is *independent* of what happens to other agents, while the breakup phase is as before with independent probabilities b/T of losing jobs. Let $\bar{\mu}^T$ be the associated (again unique) steady-state distribution, and $\bar{\mu}^* = \lim_T \bar{\mu}^T$ (which is well defined as shown in the proof of Claim 1 below).

The following claims establish the proposition.

Claim 1: $\bar{\mu}^* = \mu^*$.

Claim 2: $\bar{\mu}^*$ is strongly associated.

The following lemma is useful in the proof of Claim 1.

Let $\mathbf{P}$ be a transition matrix for an aperiodic, irreducible Markov chain on a finite-state space Z. For any $\mathbf{z} \in Z$, let a $\mathbf{z}$-tree be a directed graph on the set of vertices Z, with a unique directed path leading from each state $\mathbf{z}' \neq \mathbf{z}$ to $\mathbf{z}$. Denote the set of all $\mathbf{z}$-trees by $\mathcal{T}_z$. Let

(A1) $$p_{\mathbf{z}} = \sum_{\tau\in\mathcal{T}_{\mathbf{z}}} [\times_{\mathbf{z}',\mathbf{z}''\in\tau}\mathbf{P}_{\mathbf{z}'\mathbf{z}''}].$$

LEMMA 10 Mark Freidlin and Alexander D. Wentzell (1984)[35]: *If* $\mathbf{P}$ *is a transition matrix for an aperiodic, irreducible Markov chain on a finite-state space* Z*, then its unique steady-state distribution* μ *is described by*

[35] See Chapter 6, Lemma 3.1; also see Young (1993) for the adaptation to discrete processes.

$$\mu(\mathbf{z}) = \frac{p_{\mathbf{z}}}{\sum_{\mathbf{z}' \in \mathbf{Z}} p_{\mathbf{z}'}},$$

where $p_{\mathbf{z}}$ is as in (A1) *above.*

PROOF OF CLAIM 1:

Given $\mathbf{s} \in \{0, 1\}^n$, we consider a special subset of the set of $\mathcal{T}_{\mathbf{s}}$, which we denote $\mathcal{T}^*_{\mathbf{s}}$. This is the set of s-trees such that if $\mathbf{s}'$ is directed to $\mathbf{s}''$ under the tree τ, then $\mathbf{s}'$ and $\mathbf{s}''$ are adjacent. As $\mathbf{P}^T_{\mathbf{s}'\mathbf{s}''}$ goes to 0 at the rate $1/T$ when $\mathbf{s}'$ and $\mathbf{s}''$ are adjacent,[36] and other transition probabilities go to 0 at a rate of at least $1/T^2$, it follows from Lemma 10 that for large enough $T\mu^T(\mathbf{s})$ may be approximated by

$$\frac{\sum_{\tau \in \mathcal{T}^*_{\mathbf{s}}} [\times_{\mathbf{s}',\mathbf{s}'' \in \tau} \mathbf{P}^T_{\mathbf{s}'\mathbf{s}''}]}{\sum_{\hat{\mathbf{s}}} \sum_{\tau \in \mathcal{T}^*_{\hat{\mathbf{s}}}} [\times_{\mathbf{s}',\mathbf{s}'' \in \tau} \mathbf{P}^T_{\mathbf{s}'\mathbf{s}''}]}.$$

Moreover, note that for large T and adjacent $\mathbf{s}'$ and $\mathbf{s}''$, $\mathbf{P}^T_{\mathbf{s}'\mathbf{s}''}$ is either $b/T + o(1/T^2)$ (when $s'_i > s''_i$ for some i) or $p_i(\mathbf{s}')/T + o(1/T^2)$ (when $s'_i < s''_i$ for some i), where $o(1/T^2)$ indicates a term that goes to zero at the rate of $1/T^2$. For adjacent $\mathbf{s}'$ and $\mathbf{s}''$, let $\tilde{\mathbf{P}}^T_{\mathbf{s}'\mathbf{s}''} = b/T$ when $s'_i > s''_i$ for some i, and $p_i(\mathbf{s}')/T$ when $s'_i < s''_i$ for some i.[37] It then follows that

$$\text{(A2)} \qquad \mu^*(\mathbf{s}) = \lim_{T \to \infty} \frac{\sum_{\tau \in \mathcal{T}^*_{\mathbf{s}}} [\times_{\mathbf{s}',\mathbf{s}'' \in \tau} \mathbf{P}^T_{\mathbf{s}'\mathbf{s}''}]}{\sum_{\hat{\mathbf{s}}} \sum_{\tau \in \mathcal{T}^*_{\hat{\mathbf{s}}}} [\times_{\mathbf{s}',\mathbf{s}'' \in \tau} \tilde{\mathbf{P}}^T_{\mathbf{s}'\mathbf{s}''}]}.$$

By a parallel argument, this is the same as $\bar{\mu}^*(\mathbf{s})$.

PROOF OF CLAIM 2:

Equation (A2) and Claim 1 imply that

$$\bar{\mu}^*(\mathbf{s}) = \lim_{T \to \infty} \frac{\sum_{\tau \in \mathcal{T}^*_{\mathbf{s}}} [\times_{\mathbf{s}',\mathbf{s}'' \in \tau} \tilde{\mathbf{P}}^T_{\mathbf{s}'\mathbf{s}''}]}{\sum_{\hat{\mathbf{s}}} \sum_{\tau \in \mathcal{T}^*_{\hat{\mathbf{s}}}} [\times_{\mathbf{s}',\mathbf{s}'' \in \tau} \tilde{\mathbf{P}}^T_{\mathbf{s}'\mathbf{s}''}]}.$$

Multiplying top and bottom of the fraction on the right-hand side by T, we find that

$$\text{(A3)} \qquad \bar{\mu}^*(\mathbf{s}) = \frac{\sum_{\tau \in \mathcal{T}^*_{\mathbf{s}}} [\times_{\mathbf{s}',\mathbf{s}'' \in \tau} \hat{\mathbf{P}}_{\mathbf{s}'\mathbf{s}''}]}{\sum_{\hat{\mathbf{s}}} \sum_{\tau \in \mathcal{T}^*_{\hat{\mathbf{s}}}} [\times_{\mathbf{s}',\mathbf{s}'' \in \tau} \hat{\mathbf{P}}_{\mathbf{s}'\mathbf{s}''}]},$$

where $\hat{\mathbf{P}}$ is set as follows. For adjacent $\mathbf{s}'$ and $\mathbf{s}''$ (letting i be the agent for whom $s'_i \neq s''_i$), $\hat{\mathbf{P}}_{\mathbf{s}'\mathbf{s}''} = b$ when $s'_i > s''_i$, and $\hat{\mathbf{P}}_{\mathbf{s}'\mathbf{s}''} = p_i(\mathbf{s}')$ when $s'_i < s''_i$,[38] and $\hat{\mathbf{P}}_{\mathbf{s}'\mathbf{s}''} = 0$ for nonadjacent $\mathbf{s}'$ and $\mathbf{s}''$.

The proof of the claim is then established via the following steps.

[36] Note that, since $\mathbf{s}'$ and $\mathbf{s}''$ are adjacent, then $\mathbf{P}^T_{\mathbf{s}'\mathbf{s}''} \neq 0$.

[37] We take T high enough such that all coefficients of the transition matrix $\tilde{\mathbf{P}}$ are between 0 and 1.

[38] If $p_i(\mathbf{s}') > 1$ for some i and $\mathbf{s}'$, we can divide top and bottom through by some fixed constant to adjust, without changing the steady-state distribution.

Step 1: $\bar{\mu}^*$ is associated.

Step 2: $\bar{\mu}^*$ is strongly associated.

PROOF OF STEP 1:

We show that for any T and any associated μ, $\mu\bar{\mathbf{P}}^T$ is associated. From this, it follows that if we start from an associated μ_0 at time 0 (say an independent distribution), then $\mu_0(\bar{\mathbf{P}}^T)^k$ is associated for any k. Since $\bar{\mu}^T = \lim_k \mu_0(\bar{\mathbf{P}}^T)^k$ for any μ_0 (as $\bar{\mu}^T$ is the steady-state distribution), and association is preserved under (weak) convergence,[39] this implies that $\bar{\mu}^T$ is associated for all T. Then again, since association is preserved under (weak) convergence, this implies that $\lim_T \bar{\mu}^T = \bar{\mu}^*$ is associated.

So, let us now show that for any T and any associated μ, $\nu = \mu\bar{\mathbf{P}}^T$ is associated. By Lemma 6, we need to show that

$$\nu(EE') - \nu(E)\nu(E') \geq 0 \tag{A4}$$

for any E and E' in $\mathcal{E}$. Write

$$\nu(EE') - \nu(E)\nu(E') = \sum_{\mathbf{s}} \mu(\mathbf{s})(\bar{\mathbf{P}}^T_{\mathbf{s}EE'} - \bar{\mathbf{P}}^T_{\mathbf{s}E}\nu(E')).$$

Since S_t is independent conditional on $\mathbf{S}_{t-1} = \mathbf{s}$, it is associated.[40] Hence,

$$\bar{\mathbf{P}}^T_{\mathbf{s}EE'} \geq \bar{\mathbf{P}}^T_{\mathbf{s}E}\bar{\mathbf{P}}^T_{\mathbf{s}E'}.$$

Substituting into the previous expression we find that

$$\nu(EE') - \nu(E)\nu(E') \geq \sum_{\mathbf{s}} \mu(\mathbf{s})(\bar{\mathbf{P}}^T_{\mathbf{s}E}\bar{\mathbf{P}}^T_{\mathbf{s}E'} - \bar{\mathbf{P}}^T_{\mathbf{s}E}\nu(E'))$$

or

$$\nu(EE') - \nu(E)\nu(E') \geq \sum_{\mathbf{s}} \mu(\mathbf{s})\bar{\mathbf{P}}^T_{\mathbf{s}E}(\bar{\mathbf{P}}^T_{\mathbf{s}E'} - \nu(E')). \tag{A5}$$

Both $\bar{\mathbf{P}}^T_{\mathbf{s}E}$ and $(\bar{\mathbf{P}}^T_{\mathbf{s}E'} - \nu(E'))$ are nondecreasing functions of $\mathbf{s}$. Thus, since μ is associated, it follows from (A5) that

$$\nu(EE') - \nu(E)\nu(E') \geq \left[\sum_{\mathbf{s}} \mu(\mathbf{s})\bar{\mathbf{P}}^T_{\mathbf{s}E}\right]\left[\sum_{\mathbf{s}} \mu(\mathbf{s})(\bar{\mathbf{P}}^T_{\mathbf{s}E'} - \nu(E'))\right].$$

Then since $\sum_{\mathbf{s}} \mu(\mathbf{s})(\bar{\mathbf{P}}^T_{\mathbf{s}E'} - \nu(E')) = 0$ (by the definition of ν), the above inequality implies (A4).

PROOF OF STEP 2:

We have already established association. Thus, we need to establish that for any f and g that are increasing in some s_i and s_j respectively, where i and j are path-connected,

$$\text{Cov}_{\bar{\mu}^*}(f, g) > 0.$$

[39] See, for instance, P5 in Section 3.1 of Ryszard Szekli (1995).
[40] See, for instance, P2 in Section 3.1 of Szekli (1995).

By Lemma 7 it suffices to verify that

$$\mathrm{Cov}_{\bar{\mu}^*}(S_i, S_j) > 0.$$

For any transition matrix $\mathbf{P}$, let $\mathbf{P}_{\mathbf{s}ij} = \sum_{\mathbf{s}'} \mathbf{P}_{\mathbf{ss}'} s_i' s_j'$, and similarly $\mathbf{P}_{\mathbf{s}i} = \sum_{\mathbf{s}'} \mathbf{P}_{\mathbf{ss}'} s_i'$. Thus, these are the expected values of the product $S_i S_j$ and S_i conditional on starting at $\mathbf{s}$ in the previous period, respectively. Let

$$\mathrm{Cov}_{ij}^T = \sum_{\mathbf{s}} \bar{\mu}^T(\mathbf{s}) \bar{\mathbf{P}}_{\mathbf{s}ij}^T - \sum_{\mathbf{s}} \bar{\mu}^T(\mathbf{s}) \bar{\mathbf{P}}_{\mathbf{s}i}^T \sum_{\mathbf{s}'} \bar{\mu}^T(\mathbf{s}') \bar{\mathbf{P}}_{\mathbf{s}'j}^T.$$

It suffices to show that for each i, j for all large enough T

$$\mathrm{Cov}_{ij}^T > 0.$$

The matrix $\bar{\mathbf{P}}^T$ has diagonal entries $\bar{\mathbf{P}}_{\mathbf{ss}}^T$ which tend to 1 as $T \to \infty$ while other entries tend to 0. Thus, we use a closely associated matrix, which has the same steady-state distribution, but for which some other entries do not tend to 0.

Let

$$\underline{\mathbf{P}}_{\mathbf{ss}'}^T = \begin{cases} T\bar{\mathbf{P}}_{\mathbf{ss}'}^T & \text{if } \mathbf{s} \neq \mathbf{s}' \\ 1 - \sum_{\mathbf{s}'' \neq \mathbf{s}} T\bar{\mathbf{P}}_{\mathbf{ss}''}^T & \text{if } \mathbf{s}' = \mathbf{s}. \end{cases}$$

One can directly check that the unique steady-state distribution of $\underline{\mathbf{P}}^T$ is the same as that of $\bar{\mathbf{P}}^T$, and thus also that

$$\mathrm{Cov}_{ij}^T = \sum_{\mathbf{s}} \bar{\mu}^T(\mathbf{s}) \underline{\mathbf{P}}_{\mathbf{s}ij}^T - \sum_{\mathbf{s}} \bar{\mu}^T(\mathbf{s}) \underline{\mathbf{P}}_{\mathbf{s}i}^T \sum_{\mathbf{s}'} \bar{\mu}^T(\mathbf{s}') \underline{\mathbf{P}}_{\mathbf{s}'j}^T.$$

Note also that transitions are still independent under $\underline{\mathbf{P}}^T$. This implies that starting from any $\mathbf{s}$, the distribution $\underline{\mathbf{P}}_{\mathbf{s}}^T$ is associated and so

$$\underline{\mathbf{P}}_{\mathbf{s}ij}^T \geq \underline{\mathbf{P}}_{\mathbf{s}i}^T \underline{\mathbf{P}}_{\mathbf{s}j}^T.$$

Therefore,

$$\mathrm{Cov}_{ij}^T \geq \sum_{\mathbf{s}} \bar{\mu}^T(\mathbf{s}) \underline{\mathbf{P}}_{\mathbf{s}i}^T \underline{\mathbf{P}}_{\mathbf{s}j}^T - \sum_{\mathbf{s}} \bar{\mu}^T(\mathbf{s}) \underline{\mathbf{P}}_{\mathbf{s}i}^T \sum_{\mathbf{s}'} \bar{\mu}^T(\mathbf{s}') \underline{\mathbf{P}}_{\mathbf{s}'j}^T.$$

Note that $\underline{\mathbf{P}}_{\mathbf{s}i}^T$ converges to $\tilde{\mathbf{P}}_{\mathbf{s}i}$, where $\tilde{\mathbf{P}}_{\mathbf{s}i}$, is the rescaled version of $\hat{\mathbf{P}}$ (defined above),

$$\tilde{\mathbf{P}}_{\mathbf{ss}'} = \begin{cases} T\hat{\mathbf{P}}_{\mathbf{ss}'} & \text{if } \mathbf{s} \neq \mathbf{s}' \\ 1 - \sum_{\mathbf{s}'' \neq \mathbf{s}} T\hat{\mathbf{P}}_{\mathbf{ss}''} & \text{if } \mathbf{s}' = \mathbf{s}. \end{cases}$$

It follows that

$$\lim_{T \to \infty} \mathrm{Cov}_{ij}^T \geq \sum_{\mathbf{s}} \bar{\mu}^*(\mathbf{s}) \tilde{\mathbf{P}}_{\mathbf{s}i} \tilde{\mathbf{P}}_{\mathbf{s}j} - \sum_{\mathbf{s}} \bar{\mu}^*(\mathbf{s}) \tilde{\mathbf{P}}_{\mathbf{s}i} \sum_{\mathbf{s}'} \bar{\mu}^*(\mathbf{s}') \tilde{\mathbf{P}}_{\mathbf{s}'j}.$$

Thus, to complete the proof, it suffices to show that

(A6)
$$\sum_{\mathbf{s}} \bar{\mu}^*(\mathbf{s})\tilde{\mathbf{P}}_{\mathbf{s}i}\tilde{\mathbf{P}}_{\mathbf{s}j} > \sum_{\mathbf{s}} \bar{\mu}^*(\mathbf{s})\tilde{\mathbf{P}}_{\mathbf{s}i} \sum_{\mathbf{s}'} \bar{\mu}^*(\mathbf{s}')\tilde{\mathbf{P}}_{\mathbf{s}'j}.$$

Viewing $\tilde{\mathbf{P}}_{\mathbf{s}i}$ as a function of $\mathbf{s}$, this is equivalent to showing that $\text{Cov}(\tilde{\mathbf{P}}_{\mathbf{s}i}, \tilde{\mathbf{P}}_{\mathbf{s}j}) > 0$. From Step 1 we know that $\bar{\mu}^*$ is associated. We also know that $\tilde{\mathbf{P}}_{\mathbf{s}i}$ and $\tilde{\mathbf{P}}_{\mathbf{s}j}$ are both nondecreasing functions of $\mathbf{s}$.

First let us consider the case where $g_{ij} = 1$.[41] We know that $\tilde{\mathbf{P}}_{\mathbf{s}i}$ is increasing in s_i, and also that $\tilde{\mathbf{P}}_{\mathbf{s}i}$ is increasing in s_j for $g_{ij} = 1$. Similarly, $\tilde{\mathbf{P}}_{\mathbf{s}i}$ is increasing in s_j. Equation (A6) then follows from Lemma 7 (where we apply it to the case where $S_i = S_j$), as both $\tilde{\mathbf{P}}_{\mathbf{s}i}$ and $\tilde{\mathbf{P}}_{\mathbf{s}j}$ are increasing in s_j.

Next, consider any k such that $g_{jk} = 1$. Repeating the argument above, since $\tilde{\mathbf{P}}_{\mathbf{s}j}$ is increasing s_j we apply Lemma 7 again to find that S_i and S_k are positively correlated. Repeating this argument inductively leads to the conclusion that S_i and S_k are positively correlated for any i and k that are path-connected.

Proposition 1 now follows from Claim 2 since $\mu^T \rightarrow \bar{\mu}^*$.

PROOF OF PROPOSITION 2:

We know from Claim 2 that $\bar{\mu}^*$ is strongly associated. The result then follows by induction using Lemma 9, and then taking a large enough T so that μ^T is close enough to $\bar{\mu}^*$ for the desired strict inequalities to hold.

PROOF OF PROPOSITION 3:

For any $t > t' \geq 0$, let $h_{i0}^{t',t}$ be the event that $S_{it'} = S_{it'+1} \cdots = S_{it-1} = S_{it} = 0$. Let $h_{i1}^{t',t}$ be the event that $S_{it'} = 1$ and $S_{it'+1} \cdots = S_{it-1} = S_{it} = 0$. So, $h_{i0}^{t',t}$ and $h_{i1}^{t',t}$ differ only in i's status at date t'. We want to show that

(A7)
$$P(S_{i,t+1} = 1 | h_{i0}^{0t}) < P(S_{i,t+1} = 1 | h_{i0}^{1t}).$$

Since (paying close attention to the subscripts and superscripts in the definition of $h_{\cdot t}^{\cdot t}$) $P(S_{i,t+1} = 1 | h_{i0}^{1t})$ is a weighted average of $P(S_{i,t+1} = 1 | h_{i0}^{0t})$ and $P(S_{i,t+1} = 1 | h_{i1}^{0t})$, (A7) is equivalent to showing that

(A8)
$$P(S_{i,t+1} = 1 | h_{i0}^{0t}) < P(S_{i,t+1} = 1 | h_{i1}^{0t}).$$

By Bayes' rule,

$$P(S_{i,t+1} = 1 | h_{i0}^{0t}) = \frac{P(S_{i,t+1} = 1, h_{i0}^{0t})}{P(S_{i,t+1} = 1, h_{i0}^{0t}) + P(S_{i,t+1} = 0, h_{i0}^{0t})}$$

and

$$P(S_{i,t+1} = 1 | h_{i1}^{0t}) = \frac{P(S_{i,t+1} = 1, h_{i1}^{0t})}{P(S_{i,t+1} = 1, h_{i1}^{0t}) + P(S_{i,t+1} = 0, h_{i1}^{0t})}.$$

From the two above equations, we rewrite (A8) as

(A9)
$$\frac{P(S_{i,t+1} = 1, h_{i0}^{0t})}{P(S_{i,t+1} = 1, h_{i0}^{0t}) + P(S_{i,t+1} = 0, h_{i0}^{0t})} < \frac{P(S_{i,t+1} = 1, h_{i1}^{0t})}{P(S_{i,t+1} = 1, h_{i1}^{0t}) + P(S_{i,t+1} = 0, h_{i1}^{0t})}.$$

Rearranging terms, (A9) is equivalent to

[41] If i is such that $g_{ij} = 0$ for all $j \neq i$, then strong association is trivial. So we treat the case where at least two agents are path-connected.

$$P(S_{i,t+1} = 1, h_{i0}^{0t})P(S_{i,t+1} = 0, h_{i1}^{0t}) < P(S_{i,t+1} = 1, h_{i1}^{0t})P(S_{i,t+1} = 0, h_{i0}^{0t}).$$

For any τ, let E_{i0}^{τ} be the set of s_τ such that $s_{i\tau} = 0$ and E_{i1}^{τ} be the set of $\mathbf{s}_\tau$ such that $s_{i\tau} = 1$.

Letting μ^* be the limiting steady-state distribution, we divide each side of the above inequality by $\mu^*(E_{i0}^0)\mu^*(E_{i1}^0)$ to obtain

$$\frac{P(S_{i,t+1} = 1, h_{i0}^{0t})}{\mu^*(E_{i0}^0)} \frac{P(S_{i,t+1} = 0, h_{i1}^{0t})}{\mu^*(E_{i1}^0)} < \frac{P(S_{i,t+1} = 1, h_{i1}^{0t})}{\mu^*(E_{i1}^0)} \frac{P(S_{i,t+1} = 0, h_{i0}^{0t})}{\mu^*(E_{i0}^0)}.$$

Thus, to establish (A7) it is enough to show that

$$\text{(A10)} \qquad \frac{P(S_{i,t+1} = 1, h_{i0}^{0t})}{\mu^*(E_{i0}^0)} < \frac{P(S_{i,t+1} = 1, h_{i1}^{0t})}{\mu^*(E_{i1}^0)}$$

and

$$\text{(A11)} \qquad \frac{P(S_{i,t+1} = 0, h_{i1}^{0t})}{\mu^*(E_{i1}^0)} < \frac{P(S_{i,t+1} = 0, h_{i0}^{0t})}{\mu^*(E_{i0}^0)}.$$

Let us show (A10), as the argument for (A11) is analogous.

Then,

$$\frac{P(S_{i,t+1} = 1, h_{i0}^{0t})}{\mu^*(E_{i0}^0)} = \sum_{\mathbf{s}^0 \in E_{i0}^0} \sum_{\mathbf{s}^1 \in E_{i0}^1} \cdots \sum_{\mathbf{s}^{t+1} \in E_{i1}^{t+1}} \frac{\mu^*(\mathbf{s}^0)}{\mu^*(E_{i0}^0)} P_{\mathbf{s}^0\mathbf{s}^1} P_{\mathbf{s}^1\mathbf{s}^2} \cdots P_{\mathbf{s}^t\mathbf{s}^{t+1}},$$

which we rewrite as

$$\frac{P(S_{i,t+1} = 1, h_{i0}^{0t})}{\mu^*(E_{i0}^0)} = \sum_{\mathbf{s}^0} \sum_{\mathbf{s}^1 \in E_{i0}^1} \cdots \sum_{\mathbf{s}^{t+1} \in E_{i1}^{t+1}} \mu^*(\mathbf{s}^0 | E_{i0}^0) P_{\mathbf{s}^0\mathbf{s}^1} P_{\mathbf{s}^1\mathbf{s}^2} \cdots P_{\mathbf{s}^t\mathbf{s}^{t+1}}.$$

Similarly,

$$\frac{P(S_{i,t+1} = 1, h_{i1}^{0t})}{\mu^*(E_{i1}^0)} = \sum_{\mathbf{s}^0} \sum_{\mathbf{s}^1 \in E_{i0}^1} \cdots \sum_{\mathbf{s}^{t+1} \in E_{i1}^{t+1}} \mu^*(\mathbf{s}^0 | E_{i1}^0) P_{\mathbf{s}^0\mathbf{s}^1} P_{\mathbf{s}^1\mathbf{s}^2} \cdots P_{\mathbf{s}^t\mathbf{s}^{t+1}}.$$

Note that by Proposition 1, $\mu^*(\mathbf{s}^0 | E_{i1}^0)$ strictly dominates $\mu^*(\mathbf{s}^0 | E_{i0}^0)$ (with some strict inequalities since i is connected to at least one other agent). Then, by the above equations, and Lemma 9 applied iteratively,[42] we derive the desired conclusion that (A10) is satisfied.

PROOF OF PROPOSITION 4:

Let $d_i \in \{0, 1\}$. A vector of decisions $\mathbf{d}$ is an equilibrium if for each $i \in \{1, \ldots, n\}$, $d_i = 1$ implies

[42] To be careful, at each stage we are applying the lemma to $\mathbf{P}$ where $\mathbf{P}_{\mathbf{s}\mathbf{s}'}$ only has positive probability on $\mathbf{s}'$ where $s_i' = 0$, except at time $t + 1$ when $s_i' = 1$. It is easy to see that Lemma 9 extends to this variation. Also, as seen in its proof, the lemma preserves some strict inequalities that correspond to the employment status of agents who are path-connected to i. For instance, for j connected to i, $\mu^*(E_{j1}^0 | E_{i1}^0) > \mu^*(E_{j1}^0 | E_{i0}^0)$. Through Lemma 9 this translates to a higher probability on E_{j1}^t (conditional on starting at E_{i1}^0 rather than E_{i0}^0) at each subsequent time through time t, which then leads to a strictly higher probability of i receiving a job offer at time $t + 1$.

$$E\left[\sum_t \delta_i^t S_{it} \middle| \mathbf{S}_0 = \mathbf{s}, \mathbf{d}_{-i}\right] \geq c_i,$$

and $d_i = 0$ implies the reverse inequality. The maximal equilibrium corresponding to a starting state $\mathbf{S}_0 = \mathbf{s}$ is denoted $\mathbf{d}^*(\mathbf{s})$. Let $\mathbf{s} \geq \mathbf{s}'$ and $\mathbf{d} \in \{0, 1\}^n$. We first show that for any T

$$E^T[f(\mathbf{S}_t)|\mathbf{S}_0 = \mathbf{s}', \mathbf{d}] \geq E^T[f(\mathbf{S}_t)|\mathbf{S}_0 = \mathbf{s}, \mathbf{d}].$$

Lemma 8 implies that any T-period subdivision and for every nondecreasing f,

$$E^T[f(\mathbf{S}_1)|\mathbf{S}_0 = \mathbf{s}', \mathbf{d}] \geq E^T[f(\mathbf{S}_1)|\mathbf{S}_0 = \mathbf{s}, \mathbf{d}].$$

Lemma 9 and a simple induction argument then establish the inequality for all $t \geq 1$. The inequality is strict whenever f is increasing and $\mathbf{s}' > \mathbf{s}$. Next, let $\mathbf{d} \geq \mathbf{d}'$. For a fine enough T-period subdivision and for every nondecreasing f, given that dropouts have value zero it follows that

$$E^T[f(\mathbf{S}_1)|\mathbf{S}_0 = \mathbf{s}, \mathbf{d}'] \geq E^T[f(\mathbf{S}_1)|\mathbf{S}_0 = \mathbf{s}, \mathbf{d}].$$

As before, the inequality extends to all $t \geq 1$ by induction. Again, f increasing and $\mathbf{d}' > \mathbf{d}$ imply a strict inequality. Combining these observations, we find that for any T when $\mathbf{s}' \geq \mathbf{s}$ and $\mathbf{d}' \geq \mathbf{d}$

$$E^T[f(\mathbf{S}_t)|\mathbf{S}_0 = \mathbf{s}', \mathbf{d}'] \geq E^T[f(\mathbf{S}_t)|\mathbf{S}_0 = \mathbf{s}, \mathbf{d}]. \tag{A12}$$

Consider the maximal equilibrium $\mathbf{d}^*(\mathbf{s})$. By (A12), for any T and all t

$$E^T[S_{it}|\mathbf{S}_0 = \mathbf{s}', \mathbf{d}^*(\mathbf{s})] \geq E^T[S_{it}|\mathbf{S}_0 = \mathbf{s}, \mathbf{d}^*(\mathbf{s})].$$

Thus,

$$\sum_t \delta_i^t E^T[S_{it}|\mathbf{S}_0 = \mathbf{s}', \mathbf{d}^*(\mathbf{s})] \geq \sum_t \delta_i^t E^T[S_{it}|\mathbf{S}_0 = \mathbf{s}, \mathbf{d}^*(\mathbf{s})].$$

If $\mathbf{d}^*(\mathbf{s})_i = 1$, then

$$\sum_t \delta_i^t E^T[S_{it}|\mathbf{S}_0 = \mathbf{s}', \mathbf{d}^*(\mathbf{s})] \geq \sum_t \delta_i^t E^T[S_{it}|\mathbf{S}_0 = \mathbf{s}, \mathbf{d}^*(\mathbf{s})] \geq c_i$$

and so also for all $\mathbf{d}' \geq \mathbf{d}^*(\mathbf{s})$, if i is such that $\mathbf{d}^*(\mathbf{s})_i = 1$, then

$$\sum_t \delta_i^t E^T[S_{it}|\mathbf{S}_0 = \mathbf{s}', \mathbf{d}'] \geq c_i. \tag{A13}$$

Set $d_i' = \mathbf{d}^*(\mathbf{s})_i$ for any i such that $\mathbf{d}^*(\mathbf{s})_i = 1$. Fixing $\mathbf{d}'$ for such i's, find a maximal equilibrium at $\mathbf{s}'$ for the remaining i's, and set $\mathbf{d}'$ accordingly. By (A13), it follows that $\mathbf{d}'$ is an equilibrium when considering all agents. It follows that $\mathbf{d}' \geq \mathbf{d}^*(\mathbf{s})$. Given the definition of maximal equilibrium, it then follows that $\mathbf{d}^*(\mathbf{s}') \geq \mathbf{d}' \geq \mathbf{d}^*(\mathbf{s})$.

REFERENCES

Aizer, Anna and Currie, Janet. "Networks, Neighborhoods, and the Utilization of Publicly-Funded Prenatal Care in California." Mimeo, UCLA, 2002.

Akerlof, George A. "Social Distance and Social Decisions." *Econometrica*, September 1997, *65*(5), pp. 1005–28.

Aliprantis, Charalambos D. and Border, Kim C. *Infinite dimensional analysis*. Heidelberg: Springer-Verlag, 2000.

Arrow, Kenneth. "Models of Job Discrimination," in A. Pascal, ed., *Racial discrimination in economic life*. Lexington, MA: Lexington Books, 1972.

Arrow, Kenneth and Borzekowski, Ron. "Limited Network Connections and the Distribution of Wages." Mimeo, Stanford University, 2001.

Becker, Gary. *The economics of discrimination*. Chicago: University of Chicago Press, 1957.

Bertrand, Marianne; Luttmer, Erzo F. P. and Mullainathan, Sendil. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics*, August 2000, *115*(3), pp. 1019–56.

Blanchard, Olivier J. and Diamond, Peter. "Ranking, Unemployment Duration and Wages." *Review of Economic Studies*, July 1994, *61*(3), pp. 417–34.

Boorman, Scott A. "A Combinatorial Optimization Model for Transmission of Job Information Through Contact Networks." *Bell Journal of Economics*, 1975, *6*(1), pp. 216–49.

Burt, Ronald S. *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press, 1992.

Calvó-Armengol, Antoni. "Job Contact Networks." *Journal of Economic Theory*, March 2004, *115*(1), pp 191–206.

Calvó-Armengol, Antoni and Jackson, Matthew O. "Networks in Labor Markets: Wage and Employment Dynamics and Inequality." Mimeo, Caltech and Universitat Autònoma de Barcelona, http://www.hss.caltech.edu/~jacksonm/dyngen.pdf, 2003.

Card, David and Krueger, Alan B. "School Quality and Black-White Relative Earnings: A Direct Assessment." *Quarterly Journal of Economics*, February 1992, *7*(1), pp. 151–200.

Chandra, Amitabh. "Labor-Market Dropouts and the Racial Wage Gap: 1940–1990." *American Economic Review*, May 2000 (*Papers and Proceedings*), *90*(2), pp. 333–38.

Conley, Timothy G. and Topa, Giorgio. "Socio-Economic Distance and Spatial Patterns in Unemployment." Mimeo, New York University, 2001.

Durlauf, Steven. "A Theory of Persistent Income Inequality." *Journal of Economic Growth*, 1996, (1), pp. 75–93.

Esary, James D.; Proschan, Frank and Walkup, David W. "Association of Random Variables, with Applications." *Annals of Mathematical Statistics*, October 1967, *38*(5), pp. 1466–74.

Farley, Reynolds. "Blacks, Hispanics, and White Ethnic Groups: Are Blacks Uniquely Disadvantaged?" *American Economic Review*, May 1990 (*Papers and Proceedings*), *80*(2), pp. 237–41.

Flinn, Christopher and Heckman, James. "New Methods for Analyzing Structural Models of Labor Force Dynamics." *Journal of Econometrics*, 1982, *18*(1), pp. 115–68.

Freidlin, Mark and Wentzell, Alexander D. *Random perturbations of dynamical systems*. New York: Springer-Verlag, 1984.

Goyal, Sanjeev. "Learning in Networks," in G. Demange and M. Wooders, eds., *Group formation in economics: Networks, clubs, and coalitions*. Cambridge: Cambridge University Press, 2004.

Granovetter, Mark. "The Strength of Weak Ties." *American Journal of Sociology*, May 1973, *78*(6), pp. 1360–80.

______. *Getting a job: A study of contacts and careers*, 2nd Ed. Chicago: University of Chicago Press, 1995.

Heckman, James J. and Borjas, George. "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence." *Economica*, August 1980, *47*(187), pp. 247–83.

Heckman, James; Lyons, Thomas M. and Todd, Petra E. "Understanding Black-White Wage Differentials, 1960–1990." *American Economic Review*, May 2000 (*Papers and Proceedings*), *90*(2), pp. 344–49.

Jackson, Matthew O. "A Survey of Models of Network Formation: Stability and Efficiency," in G. Demange and M. Wooders, eds., *Group formation in economics: Networks, clubs, and coalitions*. Cambridge: Cambridge University Press, 2004, http://www.hss.caltech.edu/~jacksonm/netsurv.pdf.

Loury, Glenn C. "Intergenerational Transfers and the Distribution of Earnings." *Econometrica*, July 1981, *49*(4), pp. 843–67.

Lynch, Lisa M. "The Youth Labor Market in the Eighties: Determinants of Reemployment Probabilities for Young Men and Women." *Review of Economics and Statistics*, February 1989, *71*(1), pp. 37–54.

McCall, John J. "The Simple Mathematics of Information, Job Search, and Prejudice," in A. Pascal, ed., *Racial discrimination in economic life*. Lexington, MA: Lexington Books, 1972.

Montgomery, James. "Social Networks and Labor Market Outcomes: Toward an Economic Analysis." *American Economic Review*, December 1991, *81*(5), pp. 1408–18.

______**.** "Job Search and Network Composition: Implications of the Strength-of-Weak-Ties Hypothesis." *American Sociological Review*, October 1992, *57*(5), pp. 586–96.

______**.** "Weak Ties, Employment, and Inequality: An Equilibrium Analysis." *American Journal of Sociology*, March 1994, *99*(5), pp. 1212–36.

Page, Frank. "Farsighted Stability in Network Formation," in G. Demange and M. Wooders, eds., *Group formation in economics: Networks, clubs, and coalitions*. Cambridge: Cambridge University Press, 2004.

Phelps, Edmond S. "The Statistical Theory of Racism and Sexism." *American Economic Review*, September 1972, *62*(4), pp. 659–61.

Pissarides, Christopher A. "Loss of Skill During Unemployment and the Persistence of Employment Shocks." *Quarterly Journal of Economics*, November 1992, *107*(4), pp. 1371–97.

Rees, Albert. "Information Networks in Labor Markets." *American Economic Review*, May 1966 (*Papers and Proceedings*), *56*(2), pp. 559–66.

Schelling, Thomas C. "Dynamics Models of Segregation." *Journal of Mathematical Sociology*, 1971, *1*(1), pp. 143–86.

Schweitzer, Stuart O. and Smith, Ralph E. "The Persistence of the Discouraged Worker Effect." *Industrial and Labor Relations Review*, 1974, *27*(2), pp. 249–60.

Smith, James P. and Welch, Finis R. "Black Economic Progress after Myrdal." *Journal of Economic Literature*, June 1989, *27*(2), pp. 519–64.

Szekli, Ryszard. *Stochastic ordering and dependence in applied probability*. New York: Springer-Verlag (Lecture Notes in Statistics # 97), 1995.

Tassier, Troy. "Labor Market Implications of the Breadth of Weak Ties." Mimeo, University of Iowa, 2001.

Topa, Giorgio. "Social Interactions, Local Spillovers and Unemployment." *Review of Economic Studies*, April 2001, *68*(2), pp. 261–95.

van den Nouweland, Anne. "Static Networks and Coalition Formation," in G. Demange and M. Wooders, eds., *Group formation in economics: Networks, clubs, and coalitions*. Cambridge: Cambridge University Press, 2004.

Vishwanath, Tara. "Job Search, Stigma Effect, and Escape Rate from Unemployment." *Journal of Labor Economics*, October 1989, *7*(4), pp. 487–502.

Watts, Duncan J. and Strogatz, Steven H. "Collective Dynamics of 'Small-World' Networks." *Nature*, June 1998, *393*(4), pp. 440–42.

Wright, Randall. "Job Search and Cyclical Unemployment." *Journal of Political Economy*, February 1986, *94*(1), pp. 38–55.

Young, H. Peyton. "The Evolution of Conventions." *Econometrica*, January 1993, *61*(1), pp. 57–84.