

Social Networks in Determining Employment and Wages: Patterns, Dynamics, and Inequality *

Antoni Calvó-Armengol[†] and Matthew O. Jackson[‡]

October 2001

Revision: October 21, 2002

Abstract

We develop a model where agents obtain information about job opportunities through an explicitly modeled network of social contacts. We show that an improvement in the employment status of either an agent's direct or indirect contacts leads to an increase in the agent's employment probability and expected wages, in the sense of first order stochastic dominance. A similar effect results from an increase in the network contacts of an agent. In terms of dynamics and patterns, we show that employment is positively correlated across time and agents, and the same is true for wages. Moreover, unemployment exhibits persistence in the sense of 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 (in opportunity costs, education costs, or skills maintenance) and one group starts with a worse employment status or a smaller network, then that group's drop-out rate will be higher and their employment prospects and wages will be persistently below that of the other group.

Keywords: Networks, Labor Markets, Employment, Unemployment, Wages, Wage Inequality, Drop-Out Rates, Duration Dependence.

JEL Classification Numbers: A14, J64, J31, J70.

*We thank Valentina Bali, Kim Border, Antonio Cabrales, Janet Currie, Isa Hafalir, Eddie Lazear, Massimo Morelli, and David Pérez-Castrillo, for helpful conversations and discussions. We also thank seminar participants as well as Preston McAfee and three anonymous referees for their comments and suggestions. We gratefully acknowledge the financial support of Spain's Ministry of Education under grant SEC2001-0973 and the Lee Center for Advanced Networking at Caltech.

[†]Universitat Autònoma de Barcelona, Department of Economics, Edifici B, 08193 Bellaterra (Barcelona), Spain. Email: antoni.calvo@uab.es, <http://selene.uab.es/acalvo>

[‡]Division of Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena, California 91125, USA. Email: jacksonm@hss.caltech.edu, <http://www.hss.caltech.edu/~jacksonm/Jackson.html>

1 Introduction

The importance of social networks in labor markets is pervasive and well-documented. Granovetter (1973, 1995) found that over 50% of jobs in a survey of residents of a Massachusetts town obtained jobs through social contacts. Earlier work by Rees (1966) found numbers of over 60% in a similar study. Exploration in a large number of studies documents similar figures for a variety of occupations, skill levels, and socio-economic 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 and wages. 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, employment status, and wages across groups.

There are several issues that we are interested in analyzing in the context of this model. These issues are all interrelated and each is important in its own right. Let us motivate these from one particular perspective, but they do not need to be viewed only from this vantage.

The persistent inequality in wages between whites and blacks is one of the most extensively studied areas in labor economics. Smith and 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% to 40%, 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)).² The analysis of Heckman, Lyons, and Todd (2000) suggests that differences in drop-out rates are an important part of the inequality and that accounting for drop-outs actually increases the gap.³ The fact that participation in the labor force is different across groups such as whites and blacks is well-documented. For instance, Card and Krueger (1992) quote a difference in drop-out rates of 2.5 to 3 times for blacks compared to whites. Chandra (2000) provides a breakdown of differences in participation rates by education level and other characteristics, and finds ratios of a similar magnitude.

Even if one believes the inequality to be entirely explainable by differences in factors such as education, skills, and drop-out rates; one is then left to explain why those should differ across races.⁴ An analysis of social networks provides a basis for observing both higher drop-out rates

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

²See Farley (1990) for a comparison of labor market outcomes for 50 racial-ethnic groups in the U.S..

³Ignoring drop-outs biases estimated wages upwards. Given much higher drop-out rates for blacks, this can bias the wage differential.

⁴The extent to which inequality is explainable by such factors is still a point of some debate. See for instance,

in one race versus another and sustained inequality in wages and employment rates, even among those remaining in the labor force.

In order to understand why a model based on network transmission of job information exhibits these features, let us discuss the patterns and dynamics of wages and employment that a network model exhibits. Consider a given agent in a network. In the model we consider, the better the employment and wage status of the 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 given agent. This might be for any number of reasons. One is that as the employment and wage status of a connection improves it is less likely that the connection will want to keep the job for him or herself. Another reason is that the improved employment and wage status of a connection might improve their access to information about openings. There is also an indirect effect. As the employment status of other agents in the network improves, the more likely indirect information might be passed along, and also the more likely that an agent's connection might choose to pass it to that agent rather than an agent who already has a (good) job. The result of this sort of information passing is positive correlation between the employment and wage status of agents who are directly or indirectly connected in the network, within a period and across time. Establishing this turns out to be much trickier than the above explanation would suggest, for reasons that we will detail below.

Let us mention that correlation of employment and wages is observed in the data. It can be seen on a basic level in the inequality mentioned previously, as this indirectly suggests correlation within race. One can also look directly for correlation patterns. The correlation across likely social contacts is documented in recent work by Topa (2001) who demonstrates geographic correlation in unemployment across neighborhoods in Chicago, and finds a significantly positive amount of social interactions across such neighborhoods. 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 wage and employment statuses of different agents in a network then provides a basis for understanding sustained difference in drop-out rates, and resulting inequality in employment and wages. The difference and resulting inequality can arise for (at least) two reasons. One has to do with differences in initial conditions in a network, and the other has to do with differences in network structure. Let us discuss these in turn.

Consider two identical networks, except that one starts with each of its agents having a better employment and wage 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 comes from our results on the dynamics and

Darity and Mason (1998) and Heckman (1998). Independent of whether there is a significant residual gap, one still needs to explain why any differences should exist and why things like drop-out rates should differ.

correlation patterns of employment and wages. This might even be a very minor difference at first. This minor difference might cause some agents to drop-out in the worse network but remain in the better network. However, dropping-out has a contagion effect. As some of an agent's connections drop-out, that agent's future prospects worsen since the agent's connections are no longer as useful a source of job information. Thus, other agents will be more likely to drop out, and this can escalate. This means that even slight differences in initial conditions can lead to substantial differences in drop-out rates, and then worse employment and wage status for those agents who remain in the network with more drop-outs, not to mention a substantial difference in overall employment and wage status.

The above discussion shows how differences in initial conditions between two networks can lead to sustained differences in drop-out rates, employment and wages over time. Let us also discuss how differences in network structure might matter. Consider two networks where each agent has the same number of connections, but one of which is a smaller and thus "tighter-knit" network; with the smaller network representing the minority group.⁵ Even with the same arrival rates of per-capita job information, the expected future employment and wage status of agents in the smaller network will be worse than for the agents in the larger network even with similar starting conditions, as essentially a smaller network has a more introverted path structure. Even very small differences can then be magnified through a sort of contagion in drop-out rates similar to that discussed above. This again leads to inequality in drop-out rates, employment and wages.

Up to this point, we have discussed three features that we show emerge from a networked labor market:

- Employment and wages are positively correlated across agents both within and across periods.
- A poor status of social connections strengthens the incentives to withdraw from the labor force, and can lead to substantial differences in drop-out rates across groups. Moreover, small differences in starting conditions or network structure can lead to large differences in drop-out rates due to contagion effects.
- Higher drop-out rates are consistent with persistent employment and wage inequality. Not only do the drop-outs have low employment and wage status, but also the short-run as well as the long-run steady state distributions of employment and wages will be worse (in the sense of first order stochastic dominance) for the group with the higher drop-out rate. Thus, inequality in wages and employment will persist.

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

⁵A large number of studies in the economic sociology of immigration document the tightness and closeness of immigrant networks. See for instance Portes and Sensenbrenner (1993) and references therein.

a history of unemployment, the expected probability of obtaining a job and expected future wages decrease in the length of time that an agent has been unemployed.

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 direct and indirect contacts are unemployed (or have a lower wage status). 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.

Such duration dependence is well-documented in the empirical literature (e.g., see Schweitzer and Smith (1974), Heckman and Borjas (1980), Flinn and Heckman (1982), and Lynch (1989)). For instance, Lynch (1989) finds average probabilities of finding employment on the order of .30 after one week of unemployment, .08 after eight weeks of unemployment and .02 after a year of unemployment.

While there are other explanations for why one might observe duration dependence, it is still useful to note that the network model is consistent with it. Also, this explanation is quite orthogonal and hence complementary to the standard ones such as unobserved heterogeneity. We discuss this in more detail when we present the result.

At this point, let us preview a difference in policy prediction that emerges from a networked model compared to other labor market models. For instance, in the case of inequality in employment and wages, there is a predicted synergy across the network. Improving the status of a given agent also improves the outlook for that agent's connections. This is the contagion effect mentioned 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 so that only a small fraction of agents in any part of a network are affected. The model also provides suggestions to change the network structure itself.

Before presenting the model let us point out 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 Boorman (1975), Montgomery (1991, 1992, 1994), Calvó-Armengol (2000), Arrow and Borzekowski (2001), Topa (2001); not to mention the vast literature in sociology.⁷ The contribution here is that this is the first to study an explicit network model and prove some of the resulting implications for the patterns and dynamics of employment and wages, as well as the inequality across races.

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

⁶In our model, improving the status of one agent has positive external effects on other agents' expected future employment and wage status. 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 Akerlof (1997).

⁷Some related references can be found in Granovetter (1995), Montgomery (1991), and Dutta and Jackson (2002).

transmission. An example is whether or not individuals take advantage of certain available welfare programs. Recent studies by Bertrand, Luttmer, and Mullanaithan (2000) and Aizer and Currie (2002) point to the importance of social networks in such contexts.

2 A Special Case of the Network Model

In Section 3 we will present the general network model. Before presenting the general model, however, we first present a very special case where many of the intuitions are easily seen and the material can be presented informally. Answers to any questions regarding full details of the model can be found in Section 3.

The special case of the model that we start with is one that we refer to as homogeneous job networks.

2.1 Homogeneous job networks

$N = \{1, \dots, n\}$ is a set of agents and time evolves in discrete periods, $t \in \{0, 1, 2, \dots\}$.

Homogeneous job networks are network economies where jobs are all identical (e.g., unskilled labor) and wages depend only on whether a worker is employed or not. Thus, for these networks all that we need to keep track of is whether or not an agent is employed.

Let s_{it} denote the employment status of agent i at time t . The vector $s_t \in \{0, 1\}^n$ represents a realization of the employment status at time t , with $s_{it} = 1$ indicating that agent i is employed at the end of period t .

So, a period begins with some agents being employed and others not. 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. If the agent is unemployed, then he or she will take the job. However, if the agent is unemployed 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 very important in determining a person's long-term employment prospects.

In the simple case we treat in this section, the passing is done in a very straightforward way. People either know each other or don't. If a person hears about a job and is already employed, then the person randomly picks an unemployed acquaintance to pass the job information to. If all of a person's acquaintances are already employed, then the job information is simply lost. The network of connections among agents is described by a graph g , which is an $n \times n$ matrix. For now, suppose that $g_{ij} \in \{0, 1\}$ and that g is a symmetric matrix. If $g_{ij} = 1$ then i is *linked* to j and $g_{ij} = 0$ if i is not linked to j . The interpretation is that if $g_{ij} > 0$, then when i hears about a job opening, i may tell j about the job. The symmetry of the network means that the acquaintance relationship is a reciprocal one: if i knows j , then j knows i .

In what follows, we will also be interested in indirect relationships, as friends of a friend will play a couple of roles: first, they are competition for job information in the short run, and second,

they help keep an agent's friends employed which is a benefit in the longer run. We say that two agents 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 (and again, we defer formal definitions to the next section).

The passing of information from i to j is thus described by the following probability $p_{ij}(s)$ as a function of the state of employment s at the beginning of the period

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

Finally, there is an exogenous breakup rate b between 0 and 1 which describes the (i.i.d.) probability that any given employed agent will lose their job in a given period. For convenience, we let this be the last thing that happens in a period.

2.2 The dynamics and patterns of employment

It is clear from the description above, that in this model employment follows a finite state Markov process, where the state is the vector of employment status at the end of the period. The relationship between the one-period-ahead employment status of an agent and his pattern of connections is clear, and described by the $p_{ij}(s)$'s above. Having more employed links improves i 's prospects, as does decreasing the competition for information from two-links away connections. The other indirect relationships in the network (more than two-links away) and status of other agents does not enter the calculation for one period ahead. 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 requires a close scrutiny.

Example 1 *Negative Conditional Correlations*

Consider a homogeneous job network with three agents, where agents 1 and 2 are linked, as are agents 2 and 3. Suppose the current employment state is $s_{t-1} = (0, 1, 0)$.

Conditional on this state, the employment states s_{1t} and s_{3t} are negatively correlated. In a sense, agents 1 and 3 are "competitors" for job information or a job offer from news first heard through agent 2.





Despite the fact that 1 and 3 are competitors for news from agent 2 and hence have a negative correlation in the shorter run, in the longer run agent 1 can benefit from 3's presence. Agent 3's presence can ultimately help improve 2's employment status. Also, when agent 3 is employed then 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 ultimately 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 a simple homogeneous network setting with $n = 4$ agents. Let $a = .100$ and $b = .015$. For instance, if we think about these numbers from the perspective of a time period being a week, then an agent loses a job approximately every 67 weeks, and hears (directly) about a job every ten weeks. Through the network, these lead to the following results:⁸

g	$\text{Prob}(s_1 = 0)$	$\text{Corr}(s_1, s_2)$	$\text{Corr}(s_1, s_3)$
	.132	—	—
	.083	.041	—
	.063	.025	.019
	.050	.025	.025

If there is no network relationship at all, then $\text{Prob}(s_i = 0) = .132$ under the steady state distribution. This represents an unemployment rate of 13.2%. Even moving to just a single link ($g_{12} = g_{21} = 1$) affects the individual probability for the linked agents of being employed substantially as it drops by more than a third to 8.3%. The resulting aggregate unemployment rate is 10.75%. As we see from the table, 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.

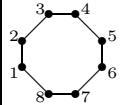
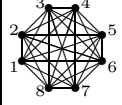
We also see that employment of both directly linked and indirectly linked agents is correlated. The correlation is decreasing in the number of links that an agent has, and is higher for direct compared to indirect connections. The decrease as a function of the number of links is due to the decreased 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 or friends or friends, etc., employed.

Also, note that the correlation between agents 1 and 3 in the “circle” ($g_{12} = g_{23} = g_{34} = g_{41} = 1$) is positive (1.9%). Thus, even though they are in competition for information from both agents

⁸The numbers for more than one agent are obtained from simulations in Maple. The programs are available upon request from the authors. The correlation numbers are only moderately accurate, even after several hundred thousand periods.

2 and 4 in the shorter run, their employment is positively correlated in the long run steady state distribution. We will see that this is true more generally, below.

Next let us examine some eight person networks, with the same information arrival and job breakup rates, $a = .100$ and $b = .015$.

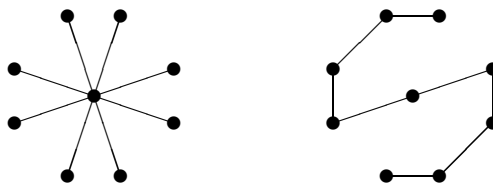
g	$\text{Prob}(s_1 = 0)$	$\text{Corr}(s_1, s_2)$	$\text{Corr}(s_1, s_3)$	$\text{Corr}(s_1, s_4)$	$\text{Corr}(s_1, s_5)$
	.060	.023	.003	.001	—
	.030	.014	.014	.014	.014

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 with a complete four person network rather than in an eight person circle. In this case, the direct connection is worth more than a number of indirect ones. More generally, the trade off between one direct connections and many 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 more 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, this tradeoff can be reversed.

Example 3 *Structure Matters: densely versus closely-knit networks*

We can also use the model to analyze how the structure matters. For instance, let us compare two different network structures that have the same number of links, but very different properties. “Stars” and “lines” are both minimally connected networks with $n - 1$ links each. Here, a star means that a center agent is connected to each other link, while a line is a sequence of agents each linked to the previous and following agents in the sequence (except for the end agents). The figure below shows a star and a line for $n = 9$ agents.



An easy way to see the contrast between these structures is to keep track of the average length of the paths connecting different agents. The average length of the minimum path between agents, denoted \bar{d} , is $2(n - 1) / n$ in the star and $(n + 1) / 3$ in the line. When $n \rightarrow \infty$, we have $\bar{d}(star) \rightarrow 2$

and $\bar{d}(line) \rightarrow \infty$. In this way we can think of stars as an example of densely-knit network and lines as loosely-knit networks.

Again, let $a = .100$ and $b = .015$. The following are the long-run average unemployment rates for stars and lines and for different values of n .

Star	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n \rightarrow \infty$
Average length \bar{d}	1.50	1.75	1.86	1.94	2.00
Unemployment	.076	.081	.094	.114	.132

Line	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n \rightarrow \infty$
Average length \bar{d}	1.67	3.00	5.67	11.00	∞
Unemployment	.069	.065	.062	.062	.061

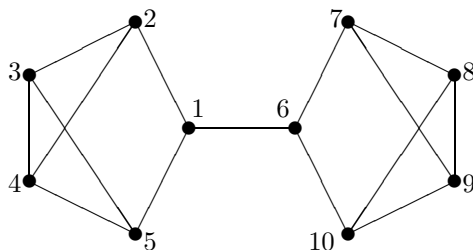
It is clear that the line dominates the star in terms of its average unemployment rate. In fact, the unemployment rate falls in the line while it increases in the star. In the star, the center agent benefits with the increase in n and her unemployment rate converges to b , as she will hear information about a job with a probability approaching one as the number of links increases, as then almost surely at least one connection will be employed and hear about a job. However, peripheral agents in the star are all connected to the same agent. In the limit, the chance that any single one of them hears about information from the center agent converges to 0, and so in the limit it is as if they were not connected at all.

Note, that this does not contradict our earlier claims (or results below) regarding positive correlation in employment. Here we are varying the network structure and comparing employment averages, rather than examining the correlation structure within a fixed network. While the star is an extreme network structure, this example does show that having a positive correlation with another agent's employment does not guarantee that the other agent is good to have around. Here the peripheral agents would like to see fewer other peripheral agents.

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

Example 4 *Bridges and Asymmetries*

Consider the following network, and again the same arrival and breakup rates.



In this network the steady state unemployment probabilities are 4.7% for agents 1 and 6, 4.8% for agents 2,5,7 and 10, and 5.0% for the rest. While these are fairly close, just some very simple

differences in the place in the network affects the unemployment rates of the agents, even though they 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: each of 3's connections are each on a path to each other with at most two links (and that does not contain 3). In a sense, they are not “well-diversified”. In contrast, some of the connections of agent 1 are not so closely tied to each other. For instance, agents 5 and 6 are not path connected (except through 1). In fact, 1 and 6 form what is referred to as a “bridge” in the social networks literature.⁹

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 eventually are looking at very short time periods– then we can begin to sort out the short and longer run effects. Essentially, in the limit we approximate a continuous time (Poisson) process, which is effectively the natural situation, where such temporary competition for jobs is short-lived and inconsequential while the overall status of indirect connections does tell one a great deal about the possible status of direct connections and the longer run effects come to dominate.

Proposition 1 *Under fine enough sub-divisions 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.*

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

Moreover, the positive correlations hold not only under the steady-state distribution, but also across any arbitrary time periods. That is, comparing i 's status at time t with j 's status at time t' .

Proposition 2 *Under fine enough sub-divisions of periods, starting under the steady-state distribution, the employment statuses of any path-connected agents are positively correlated across arbitrary periods.*

This follows, as having agents employed is beneficial both in helping their neighbors become employed in the future and in helping their neighbors pass to other people who need jobs.

⁹The lower unemployment (higher employment) rate of these agents is then consistent with ideas such as Burt's (1992) structural holes, although for different reasons than the power reasoning behind Burt's theory.

2.3 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 can be established in a networked model. To see this, let us examine some of the serial patterns of employment that emerge.

Again, consider a homogeneous network setting with an job arrival and breakup rates of $a = .100$ and $b = .015$.

Ask the following question: suppose that a person has been unemployed for at least X periods. What is the probability that they will become employed in this period? We can examine the answer to this question as we vary the number of periods of observed past unemployment:

g	1 period	2 periods	10 periods	limit
	.099	.099	.099	.099
	.176	.175	.170	.099
	.305	.300	.278	.099

Again, the patterns observed here are not particular to the example but can be shown to hold more generally.

Proposition 3 *Under fine enough sub-divisions 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.¹⁰ 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.

¹⁰Theoretical models predicting duration dependence, though, are a bit scarcer. In Blanchard and 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 Vishwanath (1989) for a model with stigma effect). In Pissarides (1992), duration dependence arises as a consequence of a decline in worker skills during the unemployment spell.

With such idiosyncratic features some agents will be quickly re-employed 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 other worker characteristic variables which would not be predicted by other models.¹¹

Some comments on stickiness in the dynamics of employment

Let us discuss some aspects of the resulting aggregate employment dynamics. In our model, the stochastic processes that regulate each individual employment history are interrelated. In particular, past employment within a close-knit set of connections breeds future employment for these connected individuals. Any shock to or change in employment has both a contemporaneous and a delayed impact on labor outcomes. In other words, duration dependence for individuals are also reflected in persistence for aggregate employment dynamics. This means that individual employment (viewed in isolation of the overall state) does not follow a Markov process, but exhibits the duration dependence documented in the above proposition. This also means that the process governing aggregate employment exhibits special features. The higher the overall employment rate, the faster unemployed vacancies are filled. So, the closer one comes to full employment, the harder it is to leave full employment. The converse also holds so that the lower the employment rate, the slower vacancies are filled. The process will oscillate between full employment and unemployment. But it exhibits a certain 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.¹²

Note also that, given an aggregate unemployment rate, filled jobs need not be evenly spread on the network, and this can even be amplified in cases where the network is asymmetric in some ways to begin with (as in Example 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.

¹¹We thank Eddie Lazear for pointing this out to us.

¹²We 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 effects pushes the process to have certain characteristics. See Wright (1986) for a search model that generates fluctuating dynamics in a proper market setting.

2.4 Dropping out and inequality in wages and 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 wage and employment status will depend in important ways on the status of those agents who are path connected to the agent in the network. This can lead to some heterogeneity across agents, as local conditions in their networks 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 have a choice 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 stay in the labor market network or to drop out. Staying in the labor market requires payment of 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.

In the following examples, 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. The theory of supermodular games then guarantees the existence of a *maximal* Nash equilibrium in pure strategies (again, see the next section for details). We restrict attention to such maximal equilibria.

This supermodular aspect of the drop-out decisions also means that there are contagion effects. An agent dropping out is bad news for the agents connections, which increases the chance that they drop out, and so forth. Thus, drop-out decisions are not i.i.d., even when the costs of staying in the labor force are i.i.d. across agents.

Example 5 *Drop Outs*

Consider a simple homogeneous network setting with a society of n individuals, where $a = .100$ and $b = .015$. Suppose that individual costs of staying in the network, c_i , are drawn at random from a uniform distribution with support on $[0, 1]$.

We compute the percentage of drop-outs for different values of n and w , where w is the wage of an employed worker and they simply compare c_i to their steady state expectation of employment times wage.¹³ The calculations are made for complete networks among the participating agents.

¹³The steady state expected employment here captures the present discounted value, except for the scaling of a discount factor δ_i . In particular, the present discounted value of expected wages is simply $w \frac{p_i}{1-\delta_i}$, where p_i is the

Dropout Percentage	$n = 1$	$n = 2$	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n \rightarrow \infty$
$w = 1$	11.5	8.3	4.6	2.8	2.1	1.7	1.5
$w = 0.8$	28.2	27.7	24.0	23.2	21.9	21.6	21.2
$w = 0.6$	47.2	46.6	44.5	45.5	41.8	41.6	40.9
$w = 0.4$	65.6	64.4	62.3	62.0	61.4	61.0	60.6
$w = 0.2$	83.7	82.9	82.2	82.0	81.1	80.7	80.3

To get a feeling for the importance of the variance of the c_i 's across the population, suppose that individual costs of staying c_i are instead drawn at random from a uniform distribution with support on some subset of $[0, 1]$. We see how this matters below, when w is fixed at 1.

Dropout Percentage	$n = 1$	$n = 2$	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n \rightarrow \infty$
$w = 1, c_i \sim U[0, 1]$	11.5	8.3	4.6	2.8	2.1	1.7	1.5
$w = 1, c_i \sim U[.8, 1]$	65.5	49.8	29.8	16.9	10.9	8.7	3.0
$w = 1, c_i \sim U[.9, 1]$	100	98.0	76.0	43.4	23.7	17.3	15.0

Drop-out percentages are clearly decreasing in wages and increasing in costs. These effects are obvious. The intuition behind the drop-out rates decreasing in the size of n is also fairly clear: with larger numbers of links, there is more of a chance of hearing about jobs and so the future prospects of employment are higher.

Next, let us consider how the initial condition of the network affects drop-out rates. This will be the place where slight starting differences can end up having lasting effects.

Take the c_i to be uniformly distributed on $[.8, 1]$ and fix $w = 1$. We compute the percent of drop-outs for different values of n . We do the calculation for two initial states: everybody employed, $s_0 = (1, \dots, 1)$, and everybody unemployed, $s_0 = (0, \dots, 0)$

The calculations are done for a discount rate of .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, be enough to see the effects. So, an agent gets a payoff of $.1s_i + .9p$, where p is the steady state employment probability in the maximal equilibrium of the dropout game and s_i is their starting state.

In calculating the dropout percentage, we can ask what amount is due to the "contagion" effect. That is, we consider the following. We can ask how many people would drop out without any equilibrium effect: so if they each did the calculation supposing that everyone else was going to stay in. Then we can calculate how many people will drop out in the equilibrium. Any extra

equilibrium steady state employment probability. The scale of δ_i can then be subsumed into either w or c_i .

people dropping out in the equilibrium, are dropping out as the result of somebody else dropping out –which we call the “contagion effect”.

So, for instance, in the lower table below, when $n = 16$ and everybody is initially unemployed, we have 68% 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% due to contagion means that about 1.5 ($= .087 \times 16$) 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.

$s_0 = (1, \dots, 1)$	$n = 1$	$n = 2$	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n \rightarrow \infty$
Dropout 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

$s_0 = (0, \dots, 0)$	$n = 1$	$n = 2$	$n = 4$	$n = 8$	$n = 16$	$n = 32$	$n \rightarrow \infty$
Dropout 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

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 non-monotonicity is a bit subtle. The possibility of contagion is initially non-existent. 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 one 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 left.

The previous examples show that different social groups with identical network relationships but differing by their starting employment 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.

For simplicity, but at some loss of generality, 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 drop-out’s employment status to be 0 forever. This does introduce a bias in terms of underestimating some gains of direct competitors to a drop-out for information from other agents. However, in

most networks this bias will be small.¹⁴ Without this simplifying assumption, keeping track of the impact of changes in the network for arbitrary structures becomes intractable. So, this approach provides for substantial simplifications in the analysis of the drop-out decisions and allows us to establish some analytical results.

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 between the two groups.*

The wage distribution and employment outcomes may thus differ among two social groups with identical economic characteristics that just differ in their starting state. In fact, many empirical studies illustrate how accounting for voluntary drop-outs from the labor force negatively affect the standard measures of black economic progress (e.g. Chandra (2000), Heckman, Lyons, and Todd (2000)).

While this comparison is a bit stylized, the fact that we consider two completely identical networks except for their starting states emphasizes how important starting conditions can be. It points out that when combined with the network dynamics and drop-out decisions, differences in initial conditions can lead to sustained inequality in a network. Moreover, these conditions will feed on each other: as one agent decides to drop-out this worsens the prospects for the agent's connections, who then drop-out at a higher rate, and so forth. This means that slight initial variations can have drastic implications.

Classical theories of discrimination, such as that of Becker (1957) or Schelling (1971), postulate that individuals have an intrinsic preference for individuals in their own societal group.¹⁵ 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.¹⁶ Our model offers an alternative and novel explanation for inequality in wages and employment.^{17, 18} Two otherwise identical individuals embedded in two societal groups with different collective employment histories (or with

¹⁴A network structure where the bias might be more substantial is the star network, as we can see from the earlier example where peripheral agents actually suffer from the addition of other peripherals. Even so, in such situations the bias is small for large numbers of agents.

¹⁵There is also an important literature on “statistical” discrimination that follows Arrow (1972), McCall (1972), Phelps (1972), and others. Our work is quite complementary to that work as well.

¹⁶We 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 non-discriminating employers and elasticities of labor supply and demand.

¹⁷While 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.

¹⁸Two other important explanations for inequality can be found in Loury (1981) and Durlauf (1996). As in our

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.

Minority traps

The previous proposition makes the point that simple differences in starting conditions can lead to different decisions to drop-out, which can in turn lead to sustained gaps in employment and wages between two otherwise identical groups. Let us add another observation to this that indicates that differences in network structure, rather than starting conditions, can also lead to different drop-out decisions and sustained inequality.

Minority groups tend to be closer knit in terms of their network connections (e.g., see Portes and Sensenbrenner (1993)), which is partly due to the size of the group. This leads to different network dynamics as having more dispersed network connections is beneficial. 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% while it is 6% 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 magnified as follows. Even if there are just a few agents who face drop-out costs that are on this order, such an agent's decision could differ depending on which group they are in. Thus, they would drop out if part of the smaller network, but not if they are part of the larger network. Their decision to drop out of the smaller network, then has implications for other agents who then might also tend to drop out. This contagion effect can lead to drastically different drop-out rates in the two networks, thus amplifying the differences in the networks and the resulting wage and employment dynamics.

As this effect involves group size, we call such a drop-out cascade a "minority-trap". This complements our earlier result on drop-out rates based on initial conditions. Here it is not the initial conditions that matter, but instead the network structure.

2.5 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. One obvious lesson is that the dynamics of the model show that policies that affect current employment or wages will have both delayed and long-lasting effects.

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 modelled 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.

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 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.

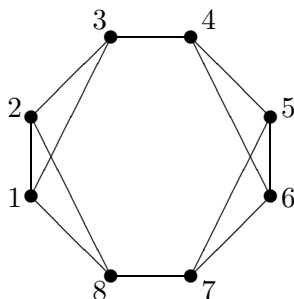
To get a better picture of this, consider the drop-out game. Suppose that we are in a situation where all agents decide to drop out. Consider two different subsidies: in the first, we pick agents distributed around the network to subsidize; while in the second we subsidize a group of agents that are clustered together. In the first case, other agents might now just have one (if any) connection who is in the market. This might not be enough to induce them to enter, and so nobody other than the subsidized agents enter the market. This hurts both their prospects and does not help the drop-out rate other than through the direct subsidy. In contrast in the second clustered case, a number of agents now have several connections who are in the market. This may induce them to enter. This can then have a contagion effect, carrying over to agents connected with them and so on. This decreases the drop-out rate beyond the direct subsidy, 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.

Concentration of subsidies

Let us again consider a society of eight individuals, again where $a = .100$ and $b = .015$. Suppose that individual costs of staying in the network, c_i , are drawn at random from a uniform distribution with support $[.8, 1]$. Initially, everybody is unemployed, so $s_0 = (0, \dots, 0)$. We work with drop out decisions when the discount rate is .9, as in the previous example.

The experiment we perform here is the following. In each case we subsidize two agents. 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. In particular the links in the network are: 12, 23, 34, 45, 56, 67, 78, 81; as well as 13, 28, 46, 57. This is pictured below.



The table 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.¹⁹

Agents Subsidized	Cost Range			
	.80 to 1	.82 to 1	.84 to 1	.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

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 the number of agents who stay in the market when the cost is at its lowest level.²⁰ When the cost is in the .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. Contagion effects are 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 is of the form 3 and 8. What matters to concentrate subsidies is the number of contacts subsidized agents have in common: the higher this number, the more concentrated the subsidies are. 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 return. We can see this not only in the choice of who to subsidize, but also how to spread subsidies, as the

¹⁹Note that the different cases of who are subsidized cover all possible configurations, up to a relabeling of the agents.

²⁰It is almost a tie with 1 and 6, but slightly ahead in the next decimal.

next calculation also shows.

Subsidies and increasing returns

The above example showed the importance of network and contagion effects in the choice of how to allocate subsidies given a fixed number of such subsidies. Suppose now that we have a fixed amount to give to the agents. We can then ask the question of whether it is better to give it all to one agent or to split it up among many agents.

Generally, the answer will depend on many factors. If the subsidy is very large, then spreading it around makes clear sense, as even a fraction of the subsidy is enough to encourage an agent (and perhaps some neighbors) to stay in. However, with a smaller sized subsidy, we can ask how this might optimally be split. Clearly, on the opposite extreme where the subsidy is very small, splitting it up makes no sense as that may spread it to thin so that it has no impact at all.

The following, however, shows that the contagion effect of the network can generally lead to some increasing returns. First let us do a primitive calculation, and then talk about the sense of it.

Consider a society of four individuals, again where $a = .100$ and $b = .015$. As before, individual costs of staying in the network, c_i , are drawn at random from a uniform distribution with support $[.8, 1]$, and everybody is initially unemployed, so $s_0 = (0, 0, 0, 0)$.

In this case, the threshold values of costs below which an individual decides to stay in are .855, .8424 and .8253 in a complete network of 4, 3 and 2 agents, respectively. The probability of various numbers of drop-outs can then be calculated, taking into account the contagion effects.

Number of agents subsidized	Percentage Staying In
0	7.1
1	26.6
2	52.3

Let $SI(x)$ be the percentage of agents who stay in the labor market given that x agents are subsidized to stay in. Thus,

$$\underbrace{\frac{1}{2} [SI(2) + SI(0)]}_{29.7} > \underbrace{SI(1)}_{26.6} .$$

What this shows, is that one would rather have 2 agents subsidized with probability one half, rather than one agent subsidized for sure. Note that this does not account for the cost of subsidizing. However, the cost of subsidizing an agent to stay in is even lower when another agent is subsidized to stay in, and so that will even further reduce the cost and increase the impact. This shows a sort of local increasing returns in the number of agents subsidized, even without accounting for the decrease in the per-capita cost of subsidy.

While this simple model is very highly stylized to the point that policy implications should be interpreted cautiously, the examples point out that the importance of network effects and their inherent contagion sorts of features can have intuitive and specific prescriptions for benefits of structuring policies with networks in mind. Moreover, as we shall show in the next section, the

types of effects noted in the model above are not unique to the homogeneous setting, but extend to a wide variety of situations and are thus representative of networked markets.

2.6 Some comparative statics and possible tests

The following tables show dropout rates and contagion effects for different arrival and breakup rates, and two different cost ranges, all for the case of four agents. The simulations assume that the starting state is one where all agents are employed, and the corresponding discount factor is .9.

The first figure in each entry is the drop-out rate, and the second is the amount attributable to contagion effects.

$c_i \sim U[.8, 1]$		b										
		.015	.045	.075	.105	.135	.165	.195	.225	.255	.285	
a	.05	69:27	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0
	.10	27:5	99:27	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0
	.15	17:2	76:27	99:12	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0
	.20	13:1	52:13	97:28	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0
	.25	11:1	42:10	83:26	100:16	100:0	100:0	100:0	100:0	100:0	100:0	100:0
	.30	10:1	37:9	68:18	98:24	100:2	100:0	100:0	100:0	100:0	100:0	100:0
	.35	9:1	29:3	61:15	88:24	100:12	100:0	100:0	100:0	100:0	100:0	100:0
	.40	9:1	27:2	53:11	84:23	98:18	100:0	100:0	100:0	100:0	100:0	100:0
	.45	8:0	25:2	46:7	76:18	96:20	100:6	100:0	100:0	100:0	100:0	100:0
	.50	7:0	25:2	45:6	69:14	91:20	100:11	100:0	100:0	100:0	100:0	100:0

$c_i \sim U[.6, 1]$		b										
		.015	.045	.075	.105	.135	.165	.195	.225	.255	.285	
a	.05	24:3	96:25	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0
	.10	11:1	44:8	90:26	100:0	100:0	100:0	100:0	100:0	100:0	100:0	100:0
	.15	8:1	28:3	56:12	85:21	99:16	100:0	100:0	100:0	100:0	100:0	100:0
	.20	6:0	22:3	45:10	66:15	90:23	98:16	99:3	100:0	100:0	100:0	100:0
	.25	5:0	18:2	34:5	54:12	71:16	92:22	99:16	100:0	100:0	100:0	100:0
	.30	4:0	16:2	28:3	44:8	60:11	75:14	94:20	100:15	100:3	100:0	100:0
	.35	4:0	14:1	27:3	37:5	52:8	65:11	82:16	95:18	100:12	100:1	100:1
	.40	4:0	13:1	24:2	34:4	46:6	58:9	71:10	87:16	96:15	100:8	100:8
	.45	4:0	12:1	21:1	32:3	42:5	56:9	69:12	80:14	90:14	99:13	99:13
	.50	4:0	11:1	20:1	30:2	37:1	52:7	63:9	75:12	84:12	96:15	96:15

The above tables provide some idea of various comparative statics. Some are quite obvious: (1) as the expected cost increases (relative to wages, which are fixed at 1 in the above), the dropout rate increases, (2) as the breakup rate increases the dropout rate increase, and (3) as the arrival

rate increases, the dropout rate increases. 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 dropout 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. The table below pulls out various rescalings of the arrival and breakup rates for the two cost ranges. As before, the first figure is the drop-out rate, and the second is the amount attributable to contagion effects.

Scaled by a and b	1	3	5	7	9
	.05, .015	.015, .045	.25, .075	.35, .105	.45, .135
$c_i \sim [.8, 1]$	69:27	76:27	83:26	88:24	96:20
$c_i \sim [.6, 1]$	24:3	28:3	34:5	37:5	42:5

As we can see, higher turnover rates (higher rescalings of a and b) lead to higher dropout rates. The intuition behind this is as follows. 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., 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 rate should also vary across these socio-economic groups. Moreover, even within a specific socio-economic group, our model predicts differences across separate components of the network as the local status of the connections changes.

3 The General Model of a Network of Labor Market Contacts

The results presented in the simple version of the homogenous job networks can also be established in a much more general model. This is important not only because of the potential applications beyond unskilled labor markets, but also because it shows that the reasoning behind the results is quite robust.

We now present the general model and a formal statement of the results.

3.1 Employment Status

There are several things that we keep track of over time.

The first is the employment status of agents. At time t , an agent $i \in N$ can either be employed (state $s_{it} = 1$) or unemployed (state $s_{it} = 0$). So, the vector $s_t \in \{0, 1\}^n$ represents a realization of the employment status at time t .

We follow the convention of representing random variables by capital letters and realizations by small letters. Thus, the sequence of random variables $\{S_0, S_1, S_2, \dots\}$ comprise the stochastic process of employment status.

3.2 Wage Status

In addition to employment status, we track wages over time.

The random variable W_{it} keeps track of the wage of agent i at time t . We normalize wages to be 0 if i is unemployed ($S_{it} = 0$), and more generally W_{it} takes on values in \mathbb{R}_+ . The vector $w_t = (w_{1t}, \dots, w_{nt})$ represents a realization of the wage levels at time t . With homogeneous job networks, the variables S_t and W_t are equivalent in terms of the information that they convey.

We allow (but do not require) the wage of an agent to depend on how many job opportunities they have come across. We now discuss how employment and wages evolve over time.

3.3 Labor Market Turnover

The labor market we consider is subject to turnover which proceeds repeatedly through two phases as follows.

- In one phase, each currently employed worker i is fired with probability $b_i \in (0, 1)$, which is referred as the *breakup rate*.
- In the other phase, each agent i is directly informed about at most one job vacancy with some probability (which may depend on the current state). If an agent directly hears about a job vacancy, then he or she either keeps that information or passes the job on to one of their direct connections in the network. Probabilities $p_{ij}(w)$ (as a function of the last period wage

status w) keep track of the probability that i first hears about a job and this job ultimately results in an offer for agent j .²¹ We discuss these p_{ij} functions in more detail below.

As these phases occur repeatedly over time, it is irrelevant whether we index periods so that first the breakup phase occurs and then the hiring phase occurs, or vice-versa. It turns out to be more convenient to consider the hiring phase first and then the breakup phase. Thus, our convention is that S_t and W_t are the employment and wage status that occurs at the *end* of period t . So, in the beginning of period t the status is described by S_{t-1}, W_{t-1} . Next, agents hear about jobs, possibly transfer that information, and hiring takes place. This results in a new employment and wage pattern. Then, the breakup phase takes place and the period ends with an employment and wage status S_t, W_t .

3.4 Specifics of Information Transmission

There are many possible variations to consider how information is transmitted and how information affects wages. There are at least three important dimensions that we consider.

One dimension to consider is whether or not an already employed agent can make direct use of information about a new job. In the case of completely homogeneous jobs (as in the section above), information about a new opening is of no use to an employed agent, and so it will be passed on. In the case of heterogeneous jobs (where jobs may have different characteristics and values to different agents), the new job may be an improvement for an already employed agent and so that agent might wish to switch jobs, and so the information about the new job is not passed on. However, there may also be a probability that the new information is not valuable to the agent (e.g., the new job is worse than their current position) and so they wish to pass it on. Generally, the higher the current wage of the agent, the higher the probability that the current job will not generate an improving offer and so the agent will pass on information about a job that he or she hears of directly.

Another dimension for consideration is to whom an employed agent passes job information. The agent may pass the job information on only to unemployed connections, or may instead select among all of his or her connections in passing on the job information. In the case where jobs are all homogeneous, it makes sense for the agent to pass the job information on to an unemployed connection. However, in the case where jobs are heterogeneous, it may make still sense for the agent to pass the job on to another agent who is already employed. Also, the agent may choose to pass the job depending on the current status of the connections.

Third, it is also possible that the agent passes the job information to more than one connection, and even that they indirectly pass it on to others, and that a number of agents end up applying or being considered for the job.

In order to capture all of these variations on information passing, we model the job transmission in a general way that allows for a wide range of cases.

²¹Note that it is possible that an agent hears about more than one job vacancy in a given period, as the agent may hear about a job directly and also may indirectly hear about jobs from one or more connections.

The job transmission and offer generation is described by a function $p_{ij} : \mathbb{R}_+^n \rightarrow [0, 1]^n$. Here $p_{ij}(W_{t-1})$ is the probability that i originally hears about a job and then it is eventually j that ends up with an offer for that job. The case where $j = i$ (that is, $p_{ii}(W_{t-1})$) represents the situation where i hears about a job and is the one who eventually gets an offer for the job.

The function p_{ij} is a reduced form that can accommodate a very large variety of situations. All that is important for our analysis is to keep track of who first heard about a job and who (if anyone) eventually ended up getting an offer for the job. In the interim it might be that agents keep any job information they hear about or it may be that they pass the information on. When passing information, agents may pass it to just one connection at a time or they may tell several connections about the job. These connections might also pass the information on to others, and it could be that several agents end up in competition for the job. And of course, all of this can depend on the current state w . Regardless of this process, we simply characterize the end result through a probability that any given agent j ends up with an offer for a job that was first heard about by agent i .

Let $p_i(w) = \sum_j p_{ji}(w)$. So, $p_i(w)$ represents the expected number of offers that i will get depending on the wage state in the last period being w . We assume that the realizations under $p_{ji}(w)$ and $p_{ki}(w)$ are independent. Note that this is very different from the realizations under p_{ij} and p_{ik} , which will generally be negatively correlated. So we are just assuming that j and k do not coordinate on whether they pass i a job. We could allow agents to coordinate on whom they pass information to. This would complicate the proofs in the paper, but would not alter the qualitative conclusions. In fact, as we let the periods become small, the probability that more than one job appears in a given period will go to zero in any case, and so it will be clear that the results extend readily.

We let p denote the vector of functions across i and j . Let \bar{w} denote the maximum value in the range of wages. The functions p_{ij} are assumed to satisfy the following conditions for any w in the range of wages:

- (1) $p_i(w)$ is nondecreasing in w_{-i} and nonincreasing in w_i , and
- (2) $p_i(w) > 0$ for any w and i such that $w_i < \bar{w}_i$
- (3) if $p_i(w) > p_i(w_{-j}, \tilde{w}_j)$ for $j \neq i$, then p_i is increasing in w_j whenever $w_i < \bar{w}_i$.

(1) imposes two requirements. The first is that the expected number of jobs that i hears about is weakly increasing in the wages of agents other than i . This encompasses the idea that other agents are (weakly) more likely to directly or indirectly pass information on that will reach i if they are more satisfied with their own position, and also that they might have better access to such information as their situation improves. It also encompasses the idea that other agents are (weakly) less likely to compete with i for an offer if they are more satisfied with their own position. The second requirement is similar but keeps track of i 's wage. Note that this allows for i to be more

likely to directly hear about a job as i 's situation worsens (allowing for a greater search intensity).²²

We remark that (1) is not in contradiction with the fact that some agents might be more qualified than other agents for a given job. Such qualifications can be completely built into the agents' identities i, j , etc., which are accounted for in the p_{ij} 's. Condition (1) only describes how changes in agents' current circumstances affect job transmission.

(2) simply requires that if an agent is not at their highest wage level, then there is some probability that they will obtain an offer. This is clearly satisfied as long as there is some probability that they directly obtain an offer, and is a very weak requirement.

(3) is a simplifying assumption. This guarantees that if i 's probability of hearing about a job is sometimes sensitive to j 's status, then it is sensitive to j 's status whenever i is not at the highest wage level. This simply allows us to make statements about positive correlations that do

not need to be conditioned on particular circumstances. Without this assumption, some strict inequalities simply become weak ones in some special cases.

3.5 The Determination of Offers, Wages, and Employment

Determination of Offers

The above described process leads to a number of new job opportunities that each agent ends up at the end of the hiring process. Let O_{it} be the random variable denoting the number of new opportunities that i has in hand at the end of the hiring process at time t . Given $W_{t-1} = w$, the distribution of O_t is governed by the realizations of the $p_{ij}(w)$'s.

Determination of Employment

As before, the labor market we consider is subject to turnover which proceeds repeatedly through two phases: first, the breakup phase where each agent i loses job with probability b_i , then, the hiring phase where agent i gets offers. The employment status then evolves as follows. If agent i was employed at the end of time $t - 1$, so $S_{i,t-1} = 1$, then the agent remains employed ($S_{it} = 1$) with probability $(1 - b_i)$ and becomes unemployed ($S_{it} = 0$) with probability b_i . If agent i was unemployed at the end of time $t - 1$, so $S_{i,t-1} = 0$, then the agent becomes employed ($S_{it} = 1$) with probability $(1 - b_i)$ conditional on $O_{it} > 0$, and otherwise the agent stays unemployed ($S_{it} = 0$).

Determination of Wages

The evolution of wages is as follows. The function $w_i : \mathbb{R}_+ \times \{0, 1, 2, \dots\} \rightarrow \mathbb{R}_+$ describes the wage that i obtains as a function of i 's previous wage and the number of new job opportunities that i ends up with at the end of the hiring phase. This function is increasing in past wages and satisfies $w_i(W_{i,t-1}, O_{it}) \geq W_{i,t-1}$.

There may still be a loss of wages, but this occurs during the breakup phase when an agent becomes unemployed. It is also assumed that $w_i(W_{i,t-1}, O_{it})$ is nondecreasing in the number of new offers received, O_{it} , and that $w_i(0, 1) > 0$ so that a new job brings a positive wage.

²²Note that it is possible to have the probability that an employed agent directly hears about a job vacancy be higher or lower than the same probability for an unemployed agent, and still be consistent with the condition (1).

In the case of completely homogeneous jobs, the wage will simply depend on whether the agent is employed or not. But in the case of heterogeneous jobs, the wage might be increasing in the number of offers an agent has. This captures the fact that the best match of a larger set of offers is likely to be better, and also that if an agent has several potential employers then competition between them will bid the wage up.²³

We emphasize that this is not at all in contradiction with the previous

assumptions on the p_{ij} 's. Wages are increasing in the offers that an agent eventually obtains, which can be thought of as the “viable” offers. An agent might hear about a job that is a poor match for him or her (e.g., their current location or position dominates the new job) and would never lead to a viable offer. It is then perfectly rational for the agent to pass the job information on to other agents, as might happen under the p_{ij} 's. The important distinction is that the offers (O_{it} 's) that are kept track of in the model are only the viable ones.

For simplicity in what follows, we assume that w_i takes on a finite set of values and that these fall in simple steps so that if $w' > w$ are adjacent elements of the range of w_i , then $w'_i = w_i(w, 1)$. This means that wages are delineated so that an agent may reach the next higher wage level with one offer. We assume that the highest wage an agent may obtain is above 0, that is $\bar{w}_i > 0$. We also assume that $w_i(w', o) \geq w_i(w, o + 1)$ for any o and w' and w such that $w'_i > w_i$. This simply says that having a higher wage status is at least as good as having one additional offer (at least in expectations).

The wage of an agent then evolves according to the following

$$W_{it} = w_i(W_{i,t-1}, O_{it})S_{it}$$

Multiplying the expression by S_{it} keeps track of whether i loses his or her job during the breakup phase.

Networks

In the general model, the network through which information is passed is already completely embodied in the p function. Nevertheless it will still be useful for us to keep track of some connection relationships. In particular, it is helpful to keep track of agents i and j for which $p_i(w)$ is sensitive to changes in w_j for some w .

We will say that i is *connected* with j if $p_i(w) \neq p_i(w_{-j}, \tilde{w}_j)$ for some w and \tilde{w}_j .

Let us emphasize that the term “connected” does not necessarily mean that i and j pass information to each other. It might be that $p_{ij}(w) = p_{ji}(w) = 0$ for all w , and yet still $p_i(w)$ is sensitive to w_j . This would happen if $p_{ki}(w)$ depended on w_j , and hence the connection might be “indirect”. In words, two agents who are connected need not pass each other information; it is just that their statuses directly or indirectly affect each other’s probability of hearing about a job.²⁴

²³One can see the reasoning behind this in search models and, for instance, in Arrow and Borzekowski (2001) where firms compete for an agent and the best match must pay the value of the second highest match.

²⁴Note also that this definition can also have $p_{ij} > 0$, but i and j not be “connected” (if p_i does not depend on w_j). This is merely an issue of semantics, as for our results it is important how changes in one agent’s status affect another, and hence our definition of connected.

Let

$$N_i(p) = \{j \mid i \text{ is connected with } j\}$$

It is natural to focus on situations where connection relationships are at least minimally reciprocal, so that $i \in N_j(p)$ if and only if $j \in N_i(p)$. We maintain this assumption in what follows. In the absence of such an assumption, some of the statements in the results that follow need to be more carefully qualified. Generally, all of the nonnegative correlation results will still hold. However, for strictly positive correlations to ensue, it must be that information can have implications that travel sufficiently through the network to have one agent's status affect another, and so the definition of path connected would need to be carefully modified to account for directed paths.

We can also keep track of further levels of this "connection" relationship. Let

$$N_i^2(p) = N_i(p) \cup (\cup_{j \in N_i(p)} N_j(p)).$$

and inductively define

$$N_i^k(p) = N_i^{k-1}(p) \cup (\cup_{j \in N_i^{k-1}(p)} N_j(p)).$$

$N_i^n(p)$ then captures all of paths generated by the indirect connection relationships of an agent i . We say that i and j are *path connected* if $j \in N_i^n(p)$.

The sets $N_i^n(p)$ partition the set of agents, so that all the agents in any element of the partition are path connected to each other. We denote that partition by $\Pi(p)$.

We assume that any $\pi \in \Pi(p)$ contains at least two agents. Thus each agent is connected with at least one other agent. Completely isolated agents have dynamics of wages and employment that are trivial, and so we restrict our attention to non-isolated agents for whom network relationships matter.

An Economy

Given an initial distribution over states μ_0 and a specification of N , p_i 's, and b_i 's, the stochastic process of employment $\{S_1, S_2, \dots\}$ and wages $\{W_1, W_2, \dots\}$ is completely specified. We refer to the specification of (N, p, b) satisfying the properties that we have outlined as an *economy*. We discuss the dependence on the initial distributions over states when necessary.

We remark that keeping track of employment status is redundant given wages, but it is still useful to distinguish these in the discussion below.

4 The Dynamics and Patterns of Employment and Wages

Next, we turn to understanding the dynamics and patterns in both employment and wages, as we look across agents and/or across time.

4.1 Patterns of Wages

We begin with patterns of wages as the results on employment have an added complication that we will discuss shortly. Before stating a theorem on wage patterns, let us discuss an issue that arises that we need to address.

Consider a situation where agents are more likely to pass job information on to direct connections with lower wages than to direct connections with higher wages. In such a situation, an agent who has a low wage, but whose wage is still higher than some other agents who are competitors for information about a job, might end up with a next period expected wage that is lower than what they would expect if they quit their job. This can happen because if they were to quit their job, their direct connections would be more likely to pass information to them, and they might have a positive probability of obtaining several offers at once.

While this might be an unusual case, it is one that we have not precluded under the assumptions on p . This difficulty is overcome when we look at fine enough subdivisions of a period, as then the probability of obtaining more than one offer becomes negligible compared to the probability of obtaining one offer, provided the probability of obtaining at least one offer is not zero, which is assured under (3). This is captured in the following definition.

T -period Subdivisions

A natural way to analyze shortened periods is simply by dividing p and b by some T .²⁵

More formally, starting from some economy (N, p, b) , the T -period subdivision, denoted (N, p^T, b^T) , is such that $b_i^T = \frac{b_i}{T}$ and $p_{ij}^T = \frac{p_{ij}}{T}$ for each i and j .

T -period subdivisions are also the natural way to sort out the short run competition from the longer run benefits of indirect connections (see Example 1). As the periods shorten, the competitive effects become outweighed by the longer run benefits. Again, this is the natural approximation of the underlying Poisson arrival process.

Recall that $\Pi(p)$ is the partition of the agents so that all the agents in any element of the partition are path connected to each other under p .

Theorem 5 *Consider any economy (N, p, b) . There exists T' such that for any $T \geq T'$, the wages of any path connected agents are positively correlated under the (unique) steady state distribution on wages corresponding to the T -period subdivision of (N, p, b) . Moreover, the limit of the steady state distributions is strongly associated relative to $\Pi(p)$.*

The theorem states that any path connected agents have positively correlated wage levels, and in fact exhibit strong association, which is a property that provides for positive interrelationships between all different subgroups. We provide a detailed definition of strong association in the appendix.

We emphasize that the limit of the steady state distributions as T becomes large is a very natural thing to consider, as it is a Poisson birth/death process which would naturally describe the job search. The reason we work with a discrete time approximation is purely for tractability.

The proof of Theorem 5 is long and appears in the appendix. The proof can be broken down into several steps. The first step shows that for large enough T the steady state distribution is approximately the same as one for a process where the realizations of $p_{ij}(w)$ across different j 's is

²⁵In the limit, this simply approximates a continuous time Poisson arrival process.

independent. Essentially, the idea is that for large enough T , the probability that just one job is heard about overwhelms the probability that more than one job is heard about. This is also true under independence. The proof then uses a characterization of steady state distributions of Markov processes by Freidlin and Wentzel (1984) (as adapted to finite processes by Young (1993)). We use the characterization to verify that one can simply keep track of the probabilities of just a single job event to get the approximate steady state distribution for large enough T . Next, note that under independence of job hearing, the negative correlation effects of Example 1 are no longer an issue. So we can then establish that the conclusions of the theorem are true under the independent process. Finally, we come back to show that the same still holds under the true (dependent) process, for large enough T .

While Theorem 5 provides results on the steady state distribution, we can deduce similar statements about the relationships between wages at different times.

Theorem 6 *Consider any economy (N, p, b) . For fine enough sub-divisions and starting under the steady state distribution, there is a strictly positive relationship between the wage statuses of any path connected agents and at any times. That is, for any any times t and t' there exists T' such that for any $T \geq T'$ and*

$$\text{Cov}^T [W_{it}W_{jt'}] > 0,$$

where i and j are path connected, where Cov^T is the covariance associated with the T -period sub-division of (N, p, b) starting at time 0 under the steady state distribution μ^T .

Although Theorem 6 is similar to Theorem 5 in its structure, it provides different implications. Theorem 5 addresses the steady state distribution, or the expected long run behavior of the system. Theorem 6 addresses any arbitrary dates in the system.²⁶

It is important in Theorem 6 that we start from the steady state distribution. For instance, if we start from a given state, such as that in Example 1, we could end up with a negative correlation.

4.2 Employment Patterns and Dynamics

One might conjecture (as we initially did) that it would be a simple Corollary to Theorem 5 for employment to exhibit the same positive correlation structure as wages. It turns out to directly follow from the positive correlation (in fact the strong association) of wages that employment is nonnegatively correlated (in fact, weakly associated) across agents. However, positive correlation of wages does not always translate into positive correlation of employment status. That is, it is possible for two agents to have positively correlated wages and yet have their employment status be independent.

This is illustrated in the following example.

²⁶Theorem 5 almost seems to be a corollary of Theorem 6, since as we let t and t' become large, the distributions of W_t and $W_{t'}$ approach the steady state distribution. However, we cannot deduce Theorem 5 from Theorem 6 since it is not ruled out that the positive correlation vanishes in the limit under Theorem 6, while we know that this is not the case from Theorem 5.

Example 6 *Positive Correlation of Wages but Independence of Employment.*

Let agent i 's wages take on three values $\{0, 1, 2\}$ and agent j 's wages take on two values $\{0, 1\}$. Let i and j be path connected (but say not connected).²⁷ Consider a limiting steady state distribution which has the following marginal distribution on W_i and W_j :

	$w_j = 0$	$w_j = 1$
$w_i = 2$	$\frac{1}{12}$	$\frac{1}{4}$
$w_i = 1$	$\frac{1}{4}$	$\frac{1}{12}$
$w_i = 0$	$\frac{1}{6}$	$\frac{1}{6}$

Under this marginal distribution, W_i and W_j are positively correlated. That is easily checked from the above table. Note, however, that S_i and S_j are independent. That is easily seen since the above distribution reduces to the following distribution on employment:

	$s_j = 0$	$s_j = 1$
$s_i = 1$	$\frac{1}{3}$	$\frac{1}{3}$
$s_i = 0$	$\frac{1}{6}$	$\frac{1}{6}$

This type of distribution cannot arise if p is a function of S rather than of W . Thus, with this added condition we can establish positive correlation in employment.

Theorem 7 *Consider any economy (N, p, b) .*

- *There exists T' such that for any $T \geq T'$ the employment of any connected agents is positively correlated under the (unique) steady state distribution on employment corresponding to the T -period subdivision of (N, p, b) .*
- *The limit (as the subdivisions become finer) of the (unique) steady state distributions on employment status is associated.*²⁸
- *If p can be written as a function of S ,²⁹ then the limit of the (unique) steady state distributions on employment is strongly associated relative to $\Pi(p)$. Thus, exists T' such that for any $T \geq T'$ the employment of any path-connected agents is positively correlated under the (unique) steady state distribution on employment corresponding to the T -period subdivision of (N, p, b) .*

²⁷That is, i and j wage statuses *do not* influence each other, but i and j are connected through a chain of agents whose wages statuses *do* influence each other.

²⁸Having fixed an initial state W_0 , an economy induces a Markov chain on the state W_t . Note that this does not correspond to a Markov chain on the state S_t , as the probability of transitions from S_t to S_{t+1} can still depend on W_t (rather than just S_t) and hence on t for a given starting distribution. Nevertheless, as the wage states do form a Markov chain, there is a steady state distribution induced on the wage state W . As S is a coarsening of W , there is a corresponding steady state distribution on S .

²⁹(3) is relaxed to hold relative to S rather than W .

Theorem 7 establishes the positive interrelationships between the employment of any collections of path connected agents under the steady state distribution.

The role of the assumption that p is dependent only on S is important in establishing the strong association of employment of agents who are path connected (rather than connected), as was shown in Example 6.

We also have an analog of Theorem 6, stating that the positive interrelationships between employment statuses hold both under the steady distribution and at any time along the dynamics.

Theorem 8 *Consider any economy (N, p, b) such that p is a function of employment status.³⁰ For fine enough sub-divisions and starting under the steady state distribution, there is a strictly positive relationship between the employment statuses of any path connected agents and at any times. That is, for any times t and t' there exists T' such that for any $T \geq T'$*

$$\text{Cov}^T [S_{it} S_{jt'}] > 0$$

for any path connected i and j , where Cov^T is the covariance associated with the T -period subdivision of (N, p, b) starting at time 0 under the steady state distribution μ^T .

5 Duration Dependence and Persistence in Unemployment

We now present a theorem outlining the duration dependence that we discussed in the introduction and in the previous section for the simple case of homogeneous job networks.

Theorem 9 *Consider an economy (N, p, b) such that p is a function of employment status. For fine enough sub-divisions, every agent's employment exhibits duration dependence.³¹ That is, for any t there exists T' such that starting from the steady state distribution at time 0, for all i and $t > 0$,*

$$\text{Prob}^T (S_{i,t+1} = 1 | S_{it} = \dots = S_{i,0} = 0) < \text{Prob}^T (S_{i,t+1} = 1 | S_{it} = \dots = S_{i,1} = 0),$$

for all T -period subdivisions of (N, p, b) where $T \geq T'$.

An implication of Theorem 9 is that longer histories of unemployment (simply iteratively applying the theorem) lead to lower expectations of obtaining a job offer in the future.

6 Dropping Out and Inequality in Wages and Employment

Consider the following game endogenizing the network structure. Let $d_i \in \{0, 1\}$ denote i 's decision of whether to stay in the labor market. Each agent discounts future wages at a rate $0 < \delta_i < 1$ and pays an expected discounted cost $c_i \geq 0$ to stay in. Agents dropping out get a payoff of zero.

³⁰The result also holds for connected agents without this assumption.

³¹Recall that we have assumed that each agent is connected to at least one other, so that $N_i(p) \neq \emptyset$ for each i . Isolated agents would not exhibit any duration dependence.

An augmented economy is a specification (N, p, b, c, δ) , where c is a vector of costs and δ is a vector of discount rates.

When an agent i exits the labor force, we reset the p 's so that $p_{ij}(w) = p_{ji}(w) = 0$ for all j and w , but do not alter the other p_{kj} 's. The agent who drops out has his or her wage set to zero.³² Therefore, when an agent drops out, it is as if the agent disappeared from the economy.

Fix an augmented economy (N, p, b, c, δ) and a starting state $W_0 = w$. A vector of decisions d is an *equilibrium* if for each $i \in \{1, \dots, n\}$, $d_i = 1$ implies

$$E \left[\sum_t \delta_i^t W_{it} \mid W_0 = w, d_{-i} \right] \geq c_i,$$

and $d_i = 0$ implies the reverse inequality.

The “drop-out” game is supermodular (see Topkis (1979)) which leads to the following lemma.

Lemma 10 *Consider any economy (N, p, b) , state $W_0 = w$, and vector of costs $c \in \mathbb{R}_+^n$. There exists T' such that for any T -period subdivision of the economy ($T \geq T'$), there is a unique equilibrium $d^*(w)$ such that $d^*(w) \geq d$ for any other equilibrium d .*

We refer to the equilibrium $d^*(w)$ in Lemma 10 as the *maximal equilibrium*.

Theorem 11 *Consider any augmented economy (N, p, b, c, δ) . Consider two starting wages states, $w' \geq w$ with $w \neq w'$. There exists T' such that the set of drop-outs under the maximal equilibrium following w' is a subset of that under w that for any T -period subdivision ($T \geq T'$); and for some specifications of the costs and discount rates the inclusion is strict. Moreover, if $d^*(w)_i = d^*(w')_i = 1$, then the distributions of i 's wages and employment W_{it} and S_{it} for any t under the maximal equilibrium following w' first order stochastically dominate those under the maximal equilibrium following w , with strict dominance for large enough t if $d^*(w)_j \neq d^*(w')_j$ for any j who is path connected to i . In fact for any increasing $f : \mathbb{R}_+^n \rightarrow \mathbb{R}$ and any t*

$$E^T [f(W_t) \mid W_0 = w', d^*(w')] \geq E^T [f(W_t) \mid W_0 = w, d^*(w)],$$

with strict inequality for some specifications of c and δ .

Theorem 11 shows how persistent inequality can arise between two otherwise similar groups with different initial employment conditions.

³²This choice is not innocuous, as we must make some choice as to how to reset the function p_{kj} when i drops out, as this is a function of w_i . How we set this has implications for agent j if agent j remains in the economy.

7 Concluding Discussion

As we have mentioned several times, we treat the network structure as largely given, except to the extent that we consider drop-outs in the last section. Of course, people do have some important control over whom they socialize with both in controlling through direct friendships they undertake as well as through making 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.³³

In addition to further endogenizing the network, we can also look deeper behind the p_{ij} 's. 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 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. To some extent that can be captured in the way we have set up the p_{ij} 's to depend on the full status of all workers. So we do allow the strength of a relationship between two agents to depend, for instance, on their employment status. But beyond this, the history of how long one has been at a current status might also affect the strength of connections. Long unemployment spells can generate a de-socialization process leading to a progressive removal from labor market opportunities and to the formation of unemployment traps. This is worth further investigation.

Finally, as we have mentioned at several points, we have not formally modeled the job arrival process or an equilibrium wage process. Extending the model to endogenize the labor market equilibrium so that probability of hearing about a job depends on current overall employment and wages are equilibrium ones, is an important next step in developing a network-labor market model. This would begin to give insights into how network structure influences equilibrium structure.

³³See Holland and Leinhardt (1977) for an early model of network co-evolution. There is a growing literature on the formation of networks that now provides a ready set of tools for analyzing this problem. An incomplete list of some of the literature includes Aumann and Myerson (1988), Jackson and Wolinsky (1996), Dutta and Mutuswami (1997), Tesfatsion (1997), Dutta, van den Nouweland, and Tijs (1998), Jackson and Watts (1998, 1999), Bala and Goyal (2000), Skyrms and Pemantle (2000), Dutta and Jackson (2000) Kranton and Minehart (2001), Currarini and Morelli (2001), Johnson and Gilles (2000), Droste, Gilles, and Johnson (2000), Slikker and van den Nouweland (2001), Jackson and van den Nouweland (2001), Mutuswami and Winter (2001), and Watts (2001). See Dutta and Jackson (2002) for an overview and further references.

References

- [1] Aizer, A. and J. Currie (2002) “Networks, Neighborhoods, and the Utilization of Publicly-Funded Prenatal Care in California,” mimeo UCLA.
- [2] Akerlof, G.A. (1997) “Social Distance and Social Decisions,” *Econometrica*, 65, pp 1005-1028.
- [3] Aliprantis C.D. and K.C. Border (2000) *Infinite Dimensional Analysis*, Springer-Verlag.
- [4] Arrow, K. (1972a) “Models of Job Discrimination,” in A. Pascal *Racial Discrimination in Economic Life*, Lexington Books: Lexington MA.
- [5] Arrow, K. (1972b) “Some Mathematical Models of Race in the Labor Market,” in A. Pascal *Racial Discrimination in Economic Life*, Lexington Books: Lexington MA.
- [6] Arrow, K. and R. Borzekowski (2000) “Limited Network Connections and the Distribution of Wages,” mimeo: Stanford University.
- [7] Aumann, R. and R. Myerson (1988) “Endogenous Formation of Links Between Players and Coalitions: An Application of the Shapley Value,” *The Shapley Value*, A. Roth, Cambridge University Press, 175–191.
- [8] Bala, V. and S. Goyal (2000) “A Non-Cooperative Model of Network Formation,” *Econometrica*, 68, pp 1181–1229.
- [9] Becker, G. (1957) *The Economics of Discrimination*, University of Chicago Press, Chicago.
- [10] Bertrand, M., E.F.P. Luttmer, and S. Mullainathan (2000) “Network Effects and Welfare Cultures,” *Quarterly Journal of Economics* 115, pp 1019-1056.
- [11] Blanchard, O.J. and P. Diamond (1994) “Ranking, Unemployment Duration and Wages”, *Review of Economic Studies* 61, 417-434.
- [12] Boorman, S.A. (1975) “A Combinatorial Optimization Model for Transmission of Job Information Through Contact Networks,” *Bell Journal of Economics* 6, 216-249.
- [13] Brock, W. and S. Durlauf (2001) “Interaction-Based Models,” forthcoming in *The Handbook of Econometrics*, vol. 5, J.J. Heckman and E. Leamer, Editors.
- [14] Burt, R.S. (1992) *Structural Holes: The Social Structure of Competition*, Harvard University Press.
- [15] Calvó-Armengol, A. (2000) “Job Contact Networks,” mimeo, [http:// www.eco.uc3m.es/~ancalvo/jobnet.pdf](http://www.eco.uc3m.es/~ancalvo/jobnet.pdf)
- [16] Card, D. and Krueger (1992) “School Quality and Black-White Relative Earnings: A Direct Assessment,” *Quarterly Journal of Economics*, 151-200.

- [17] Chandra, A. (2000) “Labor-Market Dropouts and the Racial Wage Gap: 1940-1990,” *American Economic Review*, Papers and Proceedings, 333-338.
- [18] Conley, T.G., and G. Topa (2001) “Socio-Economic Distance and Spatial Patterns in Unemployment,” mimeo: NYU.
- [19] Currarini, S. and M. Morelli (2000) “Network Formation with Sequential Demands,” *Review of Economic Design*, 5, 229–250.
- [20] Darity, W.A. and P.L. Mason (1998) “Evidence on Discrimination in Employment: Codes of Color, Codes of Gender,” *Journal of Economic Perspectives*, 12, 63-90.
- [21] Droste, E., R. Gilles, and C. Johnson (2000) “Evolution of Conventions in Endogenous Social Networks,” mimeo: Virginia Tech.
- [22] Durlauf, S. (1996) “A Theory of Persistent Income Inequality,” *Journal of Economic Growth*, 1, 75-93.
- [23] Dutta, B., and M.O. Jackson (2000) “The Stability and Efficiency of Directed Communication Networks,” *Review of Economic Design*, 5, 251–272.
- [24] Dutta, B., and M.O. Jackson (2002) “On the Formation of Networks and Groups,” Introductory chapter forthcoming in *Models of the Formation of Networks and Groups*, edited by B. Dutta and M.O. Jackson, Springer-Verlag: Heidelberg. <http://www.hss.caltech.edu/~jacksonm/Jackson.html>
- [25] Dutta, B., and S. Mutuswami (1997) “Stable Networks,” *Journal of Economic Theory*, 76, 322–344.
- [26] Dutta, B., A. van den Nouweland, and S. Tijs (1998) “Link Formation in Cooperative Situations,” *International Journal of Game Theory*, 27, 245–256.
- [27] Esary, J.D., F. Proschan, and D.W. Walkup (1967) “Association of Random Variables, with Applications,” *Annals of Mathematical Statistics*, 38, 1466-1474.
- [28] Farley, R. (1990) “Blacks, Hispanics, and White Ethnic Groups: Are Blacks Uniquely Disadvantaged?,” *American Economic Review*, Papers and Proceedings, 80, 237-241.
- [29] Flinn, C. and J. Heckman (1982) “New Methods for Analyzing Structural Models of Labor Force Dynamics,” *Journal of Econometrics*, 18, 115-168.
- [30] Freidlin, M. and A. Wentzel (1984) *Random Perturbations of Dynamical Systems*, Springer-Verlag, New-York.
- [31] Granovetter, M. (1973) “The Strength of Weak Ties,” *American Journal of Sociology*, 78, 1360-1380.

- [32] Granovetter, M. [1974](1995) *Getting a Job: A Study of Contacts and Careers*, 2nd edition, University of Chicago Press, Chicago.
- [33] Heckman, J.J. (1998) "Detecting Discrimination," *Journal of Economic Perspectives*, 12, 101-116.
- [34] Heckman, J.J. and G. Borjas (1980) "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence," *Economica* 47, 247-283.
- [35] Heckman, J., T.M. Lyons and P.E. Todd (2000) "Understanding Black-White Wage Differentials, 1960-1990," *American Economic Review*, Papers and Proceedings, 344-349.
- [36] Holland, P.W. and S. Leinhardt (1977) "A Dynamic Model for Social Networks," *Journal of Mathematical Sociology* 5, 5-20.
- [37] Jackson, M.O. (2002) "The Stability and Efficiency of Economic and Social Networks," in *Advances of Economic Design*, edited by Murat Sertel, Springer-Verlag.
- [38] Jackson, M.O. and A. van den Nouweland (2001) "Strongly Stable Networks," working paper: Caltech and University of Oregon.
- [39] Jackson, M.O. and A. Watts (1998) "The Evolution of Social and Economic Networks," forthcoming: *Journal of Economic Theory*.
- [40] Jackson, M.O. and A. Watts (1999) "On the Formation of Interaction Networks in Social Coordination Games," forthcoming: *Games and Economic Behavior*
- [41] Jackson, M.O. and A. Wolinsky (1996) "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44-74.
- [42] Johnson, C. and R.P. Gilles (2000) "Spatial Social Networks," *Review of Economic Design*, 5, 273-300.
- [43] Kranton, R. and D. Minehart (2001) "A Theory of Buyer-Seller Networks," *American Economic Review*.
- [44] Lin, N. and M. Dumin (1986) "Access to Occupations through Social Ties," *Social Networks*, 8, 365-385.
- [45] Loury, G.C. (1981) "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49, pp 843-867.
- [46] Lynch, L. (1989) "The Youth Labor Market in the Eighties: Determinants of Re-employment Probabilities for Young Men and Women," *Review of Economics and Statistics*, 71, 37-54.

- [47] McCall, J. (1972) "The Simple Mathematics of Information, Job Search, and Prejudice," in A. Pascal *Racial Discrimination in Economic Life*, Lexington Books: Lexington MA.
- [48] Montgomery, J. (1994) "Weak Ties, Employment, and Inequality: An Equilibrium Analysis," *American Journal of Sociology*, 99, 1212-1236
- [49] Montgomery, J. (1992) "Job Search and Network Composition: Implications of the Strength-of-Weak-Ties Hypothesis," *American Sociological Review*, 57, 586-596
- [50] Montgomery, J. (1991) "Social Networks and Labor Market Outcomes: Toward and Economic Analysis," *American Economic Review*, 81, 1408-1418
- [51] Mutuswami, S. and E. Winter (2000) "Subscription Mechanisms for Network Formation," mimeo: CORE and Hebrew University in Jerusalem.
- [52] Neal, D.A. and W.R. Johnson (1996) "The Role of Premarket Factors in Black- White Wage Differences," *Journal of Political Economy*, 104, 869-895.
- [53] Pissarides, C.A. (1992) "Loss of Skill During Unemployment and the Persistence of Employment Shocks," *Quarterly Journal of Economics* 107, 1371-1397.
- [54] Portes, A. and J. Sensenbrenner (1993) "Embeddedness and Immigration: Notes on the Social Determinants of Economic Action," *American Journal of Sociology*, 98, pp 1320-1350.
- [55] Rees, A. (1966) "Information Networks in Labor Markets," *American Economic Review*, 56, 559-566.
- [56] Schelling, T. (1971) "Dynamics Models of Segregation," *Journal of Mathematical Sociology* 1, 143-186.
- [57] Phelps, E. (1972) "The Statistical Theory of Racism and Sexism," *American Economic Review*, 62, 659-661.
- [58] Smith and Welch (1989) "Black Economic Progress after Myrdal," *Journal of Economic Literature*, 27, 519-564.
- [59] Schweitzer, S. and R. Smith (1974) "The Persistence of the Discourage Worker Effect," *Industrial and Labor Relations Review*, 27, 249-260
- [60] Skyrms, B. and R. Pemantle (2000) "A Dynamic Model of Social Network Formation," *Proceedings of the National Academy of Sciences*, 97, 9340-9346.
- [61] Slikker, M. and A. van den Nouweland (2001) "A One-Stage Model of Link Formation and Payoff Division," *Games and Economic Behavior*, 34, 153-175.
- [62] Szekli, R. (1995) *Stochastic Ordering and Dependence in Applied Probability*, Springer-Verlag (Lecture Notes in Statistics # 97), New York.

- [63] Tassier, T. (2001) "Labor Market Implications of the Breadth of Weak Ties," mimeo: University of Iowa.
- [64] Tesfatsion, L. (1997) "A Trade Network Game with Endogenous Partner Selection," in *Computational Approaches to Economic Problems*, H. Amman et al (eds.), Kluwer Academic Publishers, 249- 269.
- [65] Topa (2001) "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies*, 68, 261-295.
- [66] Topkis, D. (1979) "Equilibrium points in nonzero-sum n -person submodular games," *SIAM Journal of Control and Optimization* 17, 773-787.
- [67] Vishwanath, T. (1989) "Job Search, Stigma Effect, and Escape Rate from Unemployment," *Journal of Labor Economics* 7, 487-502.
- [68] Watts, A. (2001) "A Dynamic Model of Network Formation," *Games and Economic Behavior*, 34, 331-341.
- [69] Wright, R. (1986) "Job Search and Cyclical Unemployment," *Journal of Political Economy*, 94, 38-55.
- [70] Young, H.P. (1993) "The Evolution of Conventions," *Econometrica* 61, 57-84.

Appendix

We begin our analysis with two straightforward results that present intuitive observations regarding employment and wage status. These are useful later on.

The following lemmas describe the Markov process governing the evolution of employment and wages as it depends on two features: the current state of the process (w_t) and the transition probabilities (p_{ij} 's).

Lemma 12 *Consider any economy (N, p, b) , time $t > 0$, two wage states $w \in \mathbb{R}_+^n$ and $w' \in \mathbb{R}_+^n$ and an agent i who is unemployed in both states ($w_i = w'_i = 0$). If $w'_j \geq w_j$ for all $j \in N_i^2$, then the distribution of i 's employment, offers, and wages (S_{it} , O_{it} , and W_{it}) conditional on $W_{t-1} = w'$ first order stochastically dominate the corresponding distributions conditional on $W_{t-1} = w$. If $p_i(w') \neq p_i(w)$, then the first order stochastic dominance is strict.*

Lemma 12 says that improving the wage status of any of an agent's connections leads to an increase (in the sense of stochastic dominance) in the probability that the agent will be employed and the agent's expected wages. The proof of Lemma 12 follows from the fact that for any i and j the function p_{ji} is nondecreasing in w_k for $k \neq i$ (condition (2)). The proof appears in the appendix.

We offer a parallel result where the state is fixed but the network (p_{ij} 's) improves.

Fix an economy (N, p, b) and consider an alternative social structure p' . We say that p' one-period dominates p at $w \in \mathbb{R}_+^n$ from i 's perspective if $p'_{ki}(w) \geq p_{ki}(w)$ for all k .

We refer to the above as "one-period domination" since i 's perceived status will improve for the next period under p' compared to p . However, since p' and p might differ beyond i 's connections, the long run comparison between p and p' might differ from the one period comparison.

As an example, under homogeneous job networks (Example ??), this one period domination condition is satisfied at w for some i if $w_i = 0$ implies that for each k : $g'_{ki} \geq g_{ki}$ and $g_{kj} \geq g'_{kj}$ for each $j \neq i$ such that $w_j = 0$.

Lemma 13 *Consider an economy (N, p, b) and an alternative social structure p' that one-period dominates p at $w \in \mathbb{R}_+^n$ from some agent i 's perspective. The distributions of i 's employment, offers and wages (S_{it} , O_{it} and W_{it}) conditional on $W_{t-1} = w$ under p' first order stochastically dominate the corresponding distributions under p . If $p'_i(w) \neq p_i(w)$, and $w_i < \bar{w}_i$, then the first order stochastic dominance is strict.*

Lemma 13 states that an agent's probability of being employed, expected number of offers and wages all go up (in the sense of stochastic dominance) if the agent's probability of hearing job information through the network improves. Again, the straightforward proof appears in the appendix.

Proof of Lemmas 12 and 13: We prove the statements for the distribution of O_{it} . The first order stochastic dominance statements for W_{it} and S_{it} then follow easily, since W_{it} is simply $w(0, O_{it})$

with probability $1 - b_i$ and 0 with probability b_i , and similarly $S_{it} = 1$ when $O_{it} > 0$ with probability $1 - b_i$, and is 0 otherwise. We remark on the strict first order stochastic dominance for W_{it} and S_{it} at the end of the proof.

Fix some w and p . Consider i such that $w_i = 0$. Fix any agent $k \neq i$ and consider any $C \subset N \setminus \{k\}$. Let

$$P_C^k(w) = (\times_{j \in C} p_{ji}(w)) (\times_{j \in N \setminus (C \cup k)} (1 - p_{ji}(w))).$$

Thus, $P_C^k(w)$ is the probability that i hears of job offers from each agent in C and none of the agents in $N \setminus (C \cup k)$. We can then write the probability that i obtains at least h offers as

$$\text{Prob}(\{O_{it} \geq h\} \mid p, W_{t-1} = w) = \sum_{C \subset N \setminus k: |C| \geq h} (1 - p_{ki}(w)) P_C^k(w) + \sum_{C \subset N \setminus k: |C| \geq h-1} p_{ki}(w) P_C^k(w).$$

Simplifying, we obtain

$$\text{Prob}(\{O_{it} \geq h\} \mid p, W_{t-1} = w) = \sum_{C \subset N \setminus k: |C| \geq h} P_C^k(w) + \sum_{C \subset N \setminus k: |C| = h-1} p_{ki}(w) P_C^k(w). \quad (1)$$

To establish first order stochastic dominance of a distribution of O_{it} conditional on $W_{t-1} = w'$ over that conditional on $W_{t-1} = w$ (and/or similarly comparing p' and p), we need only show that $\text{Prob}(\{O_{it} \geq h\} \mid p', W_{t-1} = w')$ is at least as large $\text{Prob}(\{O_{it} \geq h\} \mid p, W_{t-1} = w)$ for each h . Strict dominance follows if there is a strict inequality for any h .

Note that from (1) we can write $\text{Prob}(\{O_{it} \geq h\} \mid p, W_{t-1} = w)$ as a function of the p_{ki} 's, which are in turn functions of w . Since $P_C^k(w)$ is independent of $p_{ki}(w)$ for any $k \in N$, it follows from equation (1), that $\text{Prob}(\{O_{it} \geq h\} \mid p, W_{t-1} = w)$, viewed as a function of the p_{ki} 's, is non-decreasing in the p_{ki} 's. Moreover, it is increasing in p_{ki} whenever there is some h such that $P_C^k(w) > 0$ for some $C \subset N \setminus \{k\} : |C| = h - 1$.

Thus, if $p'_{ji}(w') \geq p_{ji}(w)$ for each $j \in N$, then we have first order stochastic dominance, and that is strict if the inequality is strict for some k such that there is some h such that $P_C^k(w) > 0$ for some $C \subset N \setminus \{k\} : |C| = h - 1$. Note that since $p_{ji}(w) < 1$ for all $j \in N$, it follows that $1 - p_{ji}(w) > 0$ for all $j \in N$. This implies that when $h = 1$, $P_C^k(w) > 0$ for $C = \emptyset$ corresponding to $|C| = h - 1 = 0$. Thus, we get strict first order stochastic dominance if we have $p'_{ji}(w') \geq p_{ji}(w)$ for each $j \in N$ with strict inequality for any j . Therefore, any changes which lead all p_{ji} 's to be at least as large (with some strictly larger), will lead to the desired conclusions regarding (strict) first order stochastic dominance.

To establish the strict part of first order stochastic dominance for S_{it} and W_{it} , it is sufficient to conclude first order stochastic dominance and additionally that

$$\text{Prob}(\{O_{it} \geq 1\} \mid p', W_{t-1} = w') > \text{Prob}(\{O_{it} \geq 1\} \mid p, W_{t-1} = w).$$

As argued above (the case of $h = 1$), this holds whenever $p'_{ji}(w') \geq p_{ji}(w)$ for all j with strict inequality for some j ; as in the premise of the results. ■

The following definitions and lemmas are useful in the proof of Theorems 7 and 5.

We first define some useful tools.

7.1 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. There is a generalization of first order stochastic dominance relationships that applies to random vectors, that was introduced into the statistics literature by Esary, Proschan, and Walkup (1967) under the definition of *association*.

μ is *associated* if

$$\text{Cov}_\mu(f, g) \geq 0$$

for all pairs of non-decreasing functions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}$, where $\text{Cov}(f, g)$ is the covariance $E_\mu[fg] - E_\mu[f]E_\mu[g]$.

Association tells us that good news about the state (conditioning on $g(x) \geq d$) leads us to higher beliefs about the state in the sense of domination. If W_1, \dots, W_n are the random variables described by an associated measure μ , then we say that W_1, \dots, W_n are associated. Note that independent random variables are associated by definition.

Note that if W is a random vector described by a measure μ , then association of μ implies that W_i and W_j are non-negatively correlated for any i and j . Essentially, association is a way of saying that all dimensions of W are non-negatively interrelated. If W were just a two dimensional vector (e.g., there were just two agents), then this would reduce to saying that there was non-negative correlation between the agents' wage levels. The definition captures more general interactions between many agents, and says that good news in the sense of higher values of $W_i, i \in \{i_1, \dots, i_\ell\}$ about any subset or combinations of agents (here, $\{i_1, \dots, i_\ell\}$) is good (not bad) news for any other set or combinations of agents. This concept is useful in describing clustering and general forms of positive correlations in employment and wages in what follows.

Strong Association

As we often want to establish strictly positive relationships, and not just non-negative ones, we also 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.

Given is a partition Π of $\{1, \dots, n\}$ that captures which random variables might be positively related.

A measure μ on \mathbb{R}^n is *strongly associated* relative to the partition Π if it is associated, and for any $\pi \in \Pi$ and nondecreasing functions f and g

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

whenever there exist i and j such that f is increasing in w_i for all w_{-i} , g is increasing in w_j for all w_{-j} , and i and j are path connected under Π .

Strong association captures the idea that better information about any of the dimensions in π leads to unabashedly higher expectations regarding every other dimension in π . One implication of this is that W_i and W_j are positively correlated for any i and j in π .

Domination

Consider two probability measures μ and ν on a state space that is a subset of \mathbb{R}^n .

μ dominates ν if

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

for every non-decreasing function $f : \mathbb{R}^n \rightarrow \mathbb{R}$.³⁴ The domination is *strict* if strict inequality holds for some non-decreasing f .

Domination captures the idea that “higher” realizations of the state are likely under μ than under ν . In the case where $n = 1$ it reduces to first order stochastic dominance.

Lemma 14 *Consider two measures μ and ν on \mathbb{R}^n which have supports that are a subset of a finite set $W \subset \mathbb{R}^n$. μ dominates ν if and only if there exists a Markov transition function $\phi : W \rightarrow \mathcal{P}(W)$ such that*

$$\mu(w') = \sum_w \phi_{ww'} \nu(w),$$

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

Thus, μ is derived from ν by a shifting of mass “upwards” under the partial order \geq over states $w \in W$.

Proof of Lemma 14: This follows from Theorem 18.40 in Aliprantis and Border (2000). ■

Let

$$\mathcal{E} = \{E \subset W \mid w \in E, w' \geq w \Rightarrow w' \in E\}.$$

\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 wages (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, Proschan and Walkup (1967)). We include a proof of this version for completeness.

Lemma 15 *Consider two measures μ and ν on W .*

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

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 measure μ is associated if and only if

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

³⁴We can take the probability measures to be Borel measures and $E_\mu[f]$ simply represents the usual $\int_{\mathbb{R}^n} f(x)d\mu(x)$.

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 $\pi \in \Pi$.³⁵

Proof of Lemma 15: First, suppose that for every $E \in \mathcal{E}$:

$$\mu(E) \geq \nu(E). \quad (2)$$

Consider any non-decreasing f . Let the elements in its range be enumerated r_1, \dots, r_K , with $r_K > r_{K-1} \dots > r_1$. Let $E_K = f^{-1}(r_K)$. By the non-decreasing assumption on f , it follows that $E_K \in \mathcal{E}$. Inductively, define $E_k = E_{k+1} \cup f^{-1}(r_{k-1})$. It is also clear that $E_k \in \mathcal{E}$. Note that

$$f(w) = \sum_k (r_k - r_{k-1}) I_{E_k}(w).$$

Thus,

$$E_\mu(f(W_t)) = \sum_k (r_k - r_{k-1}) \mu(E_k)$$

and

$$E_\nu(f(W_t)) = \sum_k (r_k - r_{k-1}) \nu(E_k).$$

Thus, by (2) it follows that $E_\mu(f(W_t)) \geq E_\nu(f(W_t))$ for every non-decreasing f . This implies the dominance.

Note that if $\mu(E) > \nu(E)$ for some E , then we have $E_\mu(I_E(W_t)) > E_\nu(I_E(W_t))$, and so strict dominance is implied.

Next let us show the converse. Suppose that μ dominates ν . For any $E \in \mathcal{E}$ consider $f(w) = I_E(w)$ (the indicator function of E). This is a non-decreasing function. Thus, $E_\mu(I_E(W_t)) \geq E_\nu(I_E(W_t))$ and so

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

To see that strict dominance implies that $\mu(E) > \nu(E)$ for some E , note that under strict dominance we have some f for which

$$E_\mu(f(W_t)) = \sum_k (r_k - r_{k-1}) \mu(E_k) > E_\nu(f(W_t)) = \sum_k (r_k - r_{k-1}) \nu(E_k).$$

Since $\mu(E_k) \geq \nu(E_k)$ for each E_k , this implies that we have strict inequality for some E_k .

The proof for association (and strong association) is a straightforward extension of the above proof that we leave to the reader (or see Esary,

Proschan and Walkup (1967)). ■

³⁵ E is sensitive to π if its indicator function is. A nondecreasing function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is sensitive to $\pi \in \Pi$ (relative to μ) if there exist x and \tilde{x}_π such that $f(x) \neq f(x_{-\pi}, \tilde{x}_\pi)$ and x and $x_{-\pi}, \tilde{x}_\pi$ are in the support of μ .

Lemma 16 *Let μ be associated and have full (finite) support on values of W . If f is nondecreasing and is increasing in W_i for some i , and g is a nondecreasing function which is increasing in W_j for some j , and $\text{Cov}_\mu(W_i, W_j) > 0$, then $\text{Cov}_\mu(f, g) > 0$.*

Proof of Lemma 16: We first prove the following Claim.

CLAIM 1 *Let μ be associated and have finite support. If f is an increasing function of W_i which depends only on W_i , and g is an increasing function of W_j which depends only on W_j , and $\text{Cov}_\mu(W_i, W_j) > 0$, then $\text{Cov}_\mu(f(W), g(W)) > 0$.*

Proof of Claim 1: We write

$$\text{Cov}_\mu(W_i, W_j) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \text{Cov}_\mu(I_{W_i}(s), I_{W_j}(t)) ds dt, \quad (36)$$

where $I_{W_i}(s) = 1$ if $W_i > s$, and $I_{W_i}(s) = 0$, otherwise. By assumption, $\text{Cov}_\mu(W_i, W_j) > 0$. Therefore, $\text{Cov}_\mu(I_{W_i}(\bar{s}), I_{W_j}(\bar{t})) > 0$ for a set of \bar{s}, \bar{t} 's. Also,

$$\text{Cov}_\mu(f(W_i), g(W_j)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \text{Cov}_\mu(I_f(s), I_g(t)) ds dt, \quad (3)$$

where $I_f(s) = 1$ if $f(W_i) > s$, and $I_f(s) = 0$, otherwise. For each \bar{s} as described above, there exists some s' such that $I_{W_i}(\bar{s}) = 1$ if and only if $I_f(f(s')) = 1$, and similarly for \bar{t} , g , and t' . Therefore, $\text{Cov}_\mu(I_f(f(s')), I_g(g(t')) > 0$. Given the finite support of W , the sets of such \bar{s}, \bar{t} 's and corresponding s', t' 's are unions of closed intervals with nonempty interiors. By association also we know that $\text{Cov}_\mu(I_f(f(\bar{s})), I_g(g(\bar{t}))) \geq 0$ for any s, t . Since this expression is positive on a set with positive measure, and everywhere nonnegative, it follows from (3) that $\text{Cov}_\mu(f, g) > 0$. ■

Next consider f that is increasing in W_i , but might also depend on W_{-i} . Label the possible wage levels of i by w_i^k where $w_i^1 = 0$ and $w_i^K = \bar{w}_i$. Let $\gamma = \min_{K \geq k > 1, w_{-i}} f(w_i^k, w_{-i}) - f(w_i^{k-1}, w_{-i})$. By the increasing property of f it follows that $\gamma > 0$. Define $f'(w_i^k) = f(0, \dots, 0) + k \frac{\gamma}{2}$. Let $f''(w) = f(w) - f'(w_i)$. It is easily checked that f'' is non-decreasing. Similarly define g' and g'' for g relative to W_j . Then

$$\text{Cov}(f, g) = \text{Cov}(f'', g'') + \text{Cov}(f'', g') + \text{Cov}(f', g'') + \text{Cov}(f', g').$$

By association, each expression is nonnegative. By Claim 1 the last expression is positive. ■

Fix the economy (N, p, b) . Let P^T denote the matrix of transitions between different w 's under the T -period subdivision. So $P_{ww'}^T$ is the probability that $W_t = w'$ conditional on $W_{t-1} = w$.

$$\text{Let } P_{wE}^T = \sum_{w' \in E} P_{ww'}^T.$$

³⁶See, for instance, Corollary B in Section 3.1 of Szekli (1995). As μ has finite support, these integrals trivially exist.

Lemma 17 Consider an economy (N, p, b) . Consider $w' \in W$ and $w \in W$ such that $w' \geq w$, and any $t \geq 1$. Then there exists T' such that for all $T \geq T'$ and $E \in \mathcal{E}$

$$P_{w'E}^T \geq P_{wE}^T.$$

Moreover, if $w' \neq w$, then the inequality is strict for at least one E .

Proof of Lemma 17: Let us say that two states w' and w are adjacent if there exists ℓ such that $w'_{-\ell} = w_{-\ell}$ and $w'_\ell > w_\ell$ take on adjacent values in the range of ℓ 's wage function.

We show that

$$P_{w'E}^T \geq P_{wE}^T.$$

for large enough T and adjacent w and w' , as the statement then follows from a chain of comparisons across such w' and w . Let ℓ be such that $w'_\ell > w_\ell$. By definition of two adjacent wage vectors, $w'_i = w_i$, for all $i \neq \ell$.

We write

$$P_{w'E}^T = \sum_o \text{Prob}_{w'}^T(W_t \in E | O_t = o) \text{Prob}_{w'}^T(O_t = o)$$

and similarly

$$P_{wE}^T = \sum_o \text{Prob}_w^T(W_t \in E | O_t = o) \text{Prob}_w^T(O_t = o),$$

where Prob_w^T is the probability conditional on $W_{t-1} = w$. Note that by property (1) of p , $p_{\ell j}(w') \geq p_{\ell j}(w)$ for all $j \neq \ell$. Also since $w'_k = w_k$ for all $k \neq \ell$ property (1) also implies that $p_{ij}(w') \geq p_{ij}(w)$ for all $j \neq \ell$ and for all i . These inequalities imply that $\text{Prob}_{w'}^T(O_{-\ell,t})$ dominates $\text{Prob}_w^T(O_{-\ell,t})$. It is only ℓ , whose job prospects may have worsened.

Since $w'_\ell > w_\ell$, given our assumption on wages (that $w_i(w', o) \geq w_i(w, o+1)$ for any o and w' and w such that $w'_i > w_i$), it is enough to show that for any a , $\text{Prob}_{w'}^T(O_{-\ell,t} \geq a) \geq \text{Prob}_w^T(O_{-\ell,t} \geq a+1)$. This holds for large enough T , given the independence of different realizations of $p_{j\ell}$ and $p_{i\ell}$ for $i \neq j$ and property (2) of p , as then the probability of some number of offers is of a higher order than that of a greater number of offers (regardless of the starting state).³⁷

To see the strict domination, consider $E = \{w | w_\ell \geq w'_\ell\}$. Since (for large enough T) there is a positive probability that ℓ hears 0 offers under w , the inequality is strict. ■

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

Lemma 18 Consider an economy (N, p, b) and two measures μ and ν on W . There exists T' such that for all $T \geq T'$, if μ dominates ν , then μP^T dominates νP^T . Moreover, if μ strictly dominates ν , then μP^T strictly dominates νP^T .

³⁷This holds provided $w'_\ell < \bar{w}_\ell$, but in the other case, the agent is already at the highest wage state and so the claim is verified.

Proof of Lemma 18:

$$[\mu P^T](E) - [\nu P^T](E) = \sum_w P_{wE}^T (\mu_w - \nu_w).$$

By Lemma 14 we rewrite this as

$$[\mu P^T](E) - [\nu P^T](E) = \sum_w P_{wE}^T \left(\sum_{w'} \nu_{w'} \phi_{w'w} - \nu_w \right).$$

We rewrite this as

$$[\mu P^T](E) - [\nu P^T](E) = \sum_w \sum_{w'} \nu_{w'} \phi_{w'w} P_{wE}^T - \sum_w \nu_w P_{wE}^T.$$

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

$$[\mu P^T](E) - [\nu P^T](E) = \sum_{w'} \left(\sum_w \nu_{w'} \phi_{w'w} P_{wE}^T - \nu_{w'} P_{w'E}^T \right).$$

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

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

Lemma 17 implies that for large enough T , $P_{wE}^T \geq P_{w'E}^T$ whenever $w \geq w'$. Thus since $\phi_{w'w} \geq 0$ and $\sum_{w \geq w'} \phi_{w'w} = 1$, the result follows.

Suppose that μ strictly dominates ν . It follows from Lemma 14 that there exists some $w \neq w'$ such that $\phi_{w'w} > 0$. By Lemma 17, there exists some $E \in \mathcal{E}$ such that $P_{wE}^T > P_{w'E}^T$. Then $[\mu P^T](E) > [\nu P^T](E)$ for such E , implying (by Lemma 15) that μP^T strictly dominates νP^T . ■

We prove Theorem 5 and then Theorem 7, as the latter makes use of the proof of the former.

Proof of Theorem 5: Recall that P^T denotes the matrix of transitions between different w 's. Since P^T is 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 Young (1993)), which we denote μ^* .

Let \bar{P}^T be the transition matrix where the process is modified as follows. Starting in state w , in the hiring phase each agent i hears about a new job (and at most one) with probability $\frac{p_i(w)}{T}$ and this is *independent* of what happens to other agents, while the breakup phase is as before with independent probabilities $\frac{b_i}{T}$ 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 2 below).

The following claims establish the theorem.

CLAIM 2 $\bar{\mu}^* = \mu^*$.

CLAIM 3 $\bar{\mu}^*$ is strongly associated.

The following lemma is useful in the proof of Claim 2.

Let P be a transition matrix for an aperiodic, irreducible Markov chain on a finite state space Z .

For any $z \in Z$, let a z -tree be a directed graph on the set of vertices Z , with a unique directed path leading from each state $z' \neq z$ to z . Denote the set of all z -trees by \mathcal{T}_z .

Let

$$p_z = \sum_{\tau \in \mathcal{T}_z} [\times_{z', z'' \in \tau} P_{z' z''}]. \quad (4)$$

Lemma 19 Freidlin and Wentzel (1984)³⁸: *If 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*

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

where p_z is as in (4) above.

Proof of Claim 2: Given $w \in W$, we consider a special subset of the set of \mathcal{T}_w , which we denote \mathcal{T}_w^* . This is the set of w -trees such that if w' is directed to w'' under the tree τ , then w' and w'' are adjacent. As $P_{w', w''}^T$ goes to 0 at the rate $1/T$ when w' and w'' are adjacent,³⁹ and other transition probabilities go to 0 at a rate of at least $1/T^2$, it follows from Lemma 19 that $\mu^T(w)$ may be approximated for large enough T by

$$\frac{\sum_{\tau \in \mathcal{T}_w^*} [\times_{w', w'' \in \tau} P_{w' w''}^T]}{\sum_{\hat{w}} \sum_{\tau \in \mathcal{T}_{\hat{w}}^*} [\times_{w', w'' \in \tau} P_{w' w''}^T]}.$$

Moreover, note that for large T and adjacent w' and w'' , $P_{w' w''}^T$ is either $\frac{b_i}{T} + o(1/T^2)$ (when $w'_i > w''_i$) or $\frac{p_i(w')}{T} + o(1/T^2)$ (when $w'_i < w''_i$), where $o(1/T^2)$ indicates a term that goes to zero at the rate of $1/T^2$. For adjacent w' and w'' , let $\tilde{P}_{w' w''}^T = \frac{b_i}{T}$ when $w'_i > w''_i$, and $\frac{p_i(w')}{T}$ when $w'_i < w''_i$.⁴⁰ It then follows that

$$\mu^*(w) = \lim_{T \rightarrow \infty} \frac{\sum_{\tau \in \mathcal{T}_w^*} [\times_{w', w'' \in \tau} \tilde{P}_{w' w''}^T]}{\sum_{\hat{w}} \sum_{\tau \in \mathcal{T}_{\hat{w}}^*} [\times_{w', w'' \in \tau} \tilde{P}_{w' w''}^T]}. \quad (5)$$

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

Proof of Claim 3: Equation 5 and Claim 2 imply that

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

³⁸See Chapter 6, Lemma 3.1; and also see Young (1993) for the adaptation to discrete processes.

³⁹Note that under property (3) of p , since w' and w'' are adjacent, it must be that $P_{w', w''}^T \neq 0$.

⁴⁰We take T high enough such that all coefficients of the transition matrix \tilde{P} are between 0 and 1.

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

$$\bar{\mu}^*(w) = \frac{\sum_{\tau \in \mathcal{T}_w^*} \left[\times_{w', w'' \in \tau} \hat{P}_{w'w''}^T \right]}{\sum_{\hat{w}} \sum_{\tau \in \mathcal{T}_{\hat{w}}^*} \left[\times_{w', w'' \in \tau} \hat{P}_{w'w''}^T \right]}, \quad (6)$$

where \hat{P}^T is set as follows. For adjacent w' and w'' (letting i be the agent for whom $w'_i \neq w''_i$) $\hat{P}_{w'w''}^T = b_i$ when $w'_i > w''_i$, and $p_i(w')$ when $w'_i < w''_i$,⁴¹ and $\hat{P}_{w'w''}^T = 0$ for non-adjacent w' and w'' .

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

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 , $\mu \bar{P}^T$ is associated. From this, it follows that if we start from an associated μ_0 at time 0 (say an independent distribution), then $\mu_0 (\bar{P}^T)^k$ is associated for any k . Since $\bar{\mu}^T = \lim_k \mu_0 (\bar{P}^T)^k$ for any μ_0 (as $\bar{\mu}^T$ is the steady-state distribution), and association is preserved under (weak) convergence,⁴² 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 , $\nu = \mu \bar{P}^T$ is associated. By Lemma 15, we need to show that

$$\nu(E E') - \nu(E) \nu(E') \geq 0 \quad (7)$$

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

$$\nu(E E') - \nu(E) \nu(E') = \sum_w \mu(w) \left(\bar{P}_{w E E'}^T - \bar{P}_{w E}^T \nu(E') \right).$$

Since W_t is independent conditional on $W_{t-1} = w$, it is associated.⁴³ Hence,

$$\bar{P}_{w E E'}^T \geq \bar{P}_{w E}^T \bar{P}_{w E'}^T.$$

Substituting into the previous expression we find that

$$\nu(E E') - \nu(E) \nu(E') \geq \sum_w \mu(w) \left(\bar{P}_{w E}^T \bar{P}_{w E'}^T - \bar{P}_{w E}^T \nu(E') \right).$$

or

$$\nu(E E') - \nu(E) \nu(E') \geq \sum_w \mu(w) \bar{P}_{w E}^T \left(\bar{P}_{w E'}^T - \nu(E') \right). \quad (8)$$

Under the properties of the p_{ij} 's, both $\bar{P}_{w E}^T$ and $\left(\bar{P}_{w E'}^T - \nu(E') \right)$ are non-decreasing functions of w . Thus, since μ is associated, it follows from (8) that

$$\nu(E E') - \nu(E) \nu(E') \geq \left[\sum_w \mu(w) \bar{P}_{w E}^T \right] \left[\sum_w \mu(w) \left(\bar{P}_{w E'}^T - \nu(E') \right) \right].$$

⁴¹If $p_i(w') > 1$ for some i and w' , we can divide top and bottom through by some fixed constant to adjust, without changing the steady state distribution.

⁴²See, for instance, P5 in Section 3.1 of Szekli (1995).

⁴³See, for instance, P2 in Section 3.1 of Szekli (1995).

Then since $\sum_w \mu(w) (\bar{P}_{wE'}^T - \nu(E')) = 0$ (by the definition of ν), the above inequality implies (7). ■

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 w_i and w_j respectively, where i and j are path connected,

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

By Lemma 16 it suffices to verify that

$$\text{Cov}_{\bar{\mu}^*}(W_i, W_j) > 0$$

For any transition matrix P , let $P_{wij} = \sum_{w'} P_{ww'} w'_i w'_j$, and similarly $P_{wi} = \sum_{w'} P_{ww'} w'_i$. Thus these are the expected values of the product $W_i W_j$ and the wage W_i conditional on starting at w in the previous period, respectively.

Let

$$\text{Cov}_{ij}^T = \sum_w \bar{\mu}^T(w) \bar{P}_{wij}^T - \sum_w \bar{\mu}^T(w) \bar{P}_{wi}^T \sum_{w'} \bar{\mu}^T(w') \bar{P}_{w'j}^T.$$

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

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

The matrix \bar{P}^T has diagonal entries \bar{P}_{ww}^T which tend to 1 as $T \rightarrow \infty$ 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

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

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

$$\text{Cov}_{ij}^T = \sum_w \bar{\mu}^T(w) \underline{P}_{wij}^T - \sum_w \bar{\mu}^T(w) \underline{P}_{wi}^T \sum_{w'} \bar{\mu}^T(w') \underline{P}_{w'j}^T.$$

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

$$\underline{P}_{wij}^T \geq \underline{P}_{wi}^T \underline{P}_{wj}^T.$$

Therefore,

$$\text{Cov}_{ij}^T \geq \sum_w \bar{\mu}^T(w) \underline{P}_{wi}^T \underline{P}_{wj}^T - \sum_w \bar{\mu}^T(w) \underline{P}_{wi}^T \sum_{w'} \bar{\mu}^T(w') \underline{P}_{w'j}^T.$$

Note that \underline{P}_{wi}^T converges to \tilde{P}_{wi} , where \tilde{P}_{wi} is the rescaled version of \hat{P} (defined in the proof of Claim 2),

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

It follows that

$$\lim_{T \rightarrow \infty} \text{Cov}_{ij}^T \geq \sum_w \bar{\mu}^*(w) \tilde{P}_{wi} \tilde{P}_{wj} - \sum_w \bar{\mu}^*(w) \tilde{P}_{wi} \sum_{w'} \bar{\mu}^*(w') \tilde{P}_{w'j}.$$

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

$$\sum_w \bar{\mu}^*(w) \tilde{P}_{wi} \tilde{P}_{wj} > \sum_w \bar{\mu}^*(w) \tilde{P}_{wi} \sum_{w'} \bar{\mu}^*(w') \tilde{P}_{w'j}. \quad (9)$$

Viewing \tilde{P}_{wi} as a function of w , this is equivalent to showing that $\text{Cov}(\tilde{P}_{wi}, \tilde{P}_{wj}) > 0$. From Step 1 we know that $\bar{\mu}^*$ is associated. We also know that \tilde{P}_{wi} and \tilde{P}_{wj} are both non-decreasing functions of w .

First let us consider the case where $j \in N_i(p)$.⁴⁴ We know that \tilde{P}_{wi} is increasing in w_i , and also, given the assumptions on p , that \tilde{P}_{wi} is increasing in w_j for $j \in N_i(p)$. Similarly, \tilde{P}_{wj} is increasing in w_j . (9) then follows from Lemma 16 (where we apply it to the case where $W_i = W_j$), as both \tilde{P}_{wi} and \tilde{P}_{wj} are increasing in w_j .

Next, consider any $k \in N_j(p)$. Repeating the argument above, since \tilde{P}_{wj} is increasing in w_j we apply Lemma 16 again to find that W_i and W_k are positively correlated. Repeating this argument inductively leads to the conclusion that W_i and W_k are positively correlated for any i and k that are path connected. ■

The Theorem 5 now follows from Claim 3 since $\mu^T \rightarrow \bar{\mu}^*$. ■

Proof of Theorem 7: For the case where p depends only on S , the proof is an analog of the proof of Theorem 5. For the more general case, the association of the limiting distribution follows directly from the proof of Theorem 5. The remaining item is to show that in the general case, there is a large enough T so that any two indirectly connected agents have positively correlated employment under the steady state.

Consider i and $j \in N_i(p)$. We can write S as a function of W . For $\bar{\mu}^*$ defined on W , let

$$\bar{\mu}^*(s_i) = \sum_{w: S_i(w)=s_i} \bar{\mu}^*(w).$$

Note that $\bar{\mu}^*$ viewed as a measure on S_i is associated since $\bar{\mu}^*$ viewed as a measure on W is associated, and since $S_i(w)$ is non-decreasing (see Esary, Proschan and Walkup (1967)).

Next, let

$$E \tilde{P}_{s_i j} = \sum_w \bar{\mu}^*(w|s_i) \sum_{w'} \tilde{P}_{ww'} S_j(w') \quad (10)$$

So, (recalling that S_i takes on values in $\{0, 1\}$) this is the expected value of S_j conditional on the last period $S_{it-1} = s_i$, under the distribution $\bar{\mu}^*$. Note that under the steady state distribution $\bar{\mu}^*$, for any k

$$E[S_k] = \sum_{s_i} \bar{\mu}^*(s_i) E \tilde{P}_{s_i k}.$$

⁴⁴If i is such that $N_i(p) = \emptyset$, then strong association is trivial. So we treat the case where at least two agents are path connected.

Then, following steps similar to those in Step 2 we can write⁴⁵

$$\text{Cov}_{\bar{\mu}^*}(S_i, S_j) \geq \sum_{s_i} \bar{\mu}^*(s_i) E\tilde{P}_{s_i i} E\tilde{P}_{s_i j} - \sum_{s_i} \bar{\mu}^*(s_i) E\tilde{P}_{s_i} \sum_{s'_i} \bar{\mu}^*(s'_i) E\tilde{P}_{s'_i j}.$$

The remainder of the proof then follows the same lines as that of Theorem 5. $E\tilde{P}_{s_i i}$ is clearly increasing in s_i .⁴⁶ There we need to employ (10). We note that under association $\bar{\mu}^*(w|s_i)$ is nondecreasing in s_i (write $s_i = 1$ as the indicator function which is nondecreasing). Finally, since $j \in N_i(p)$ we know that $E\tilde{P}_{s_i j}$ is increasing in s_i . ■

Proof of Theorems 6 and 8: We show Theorem 6, as the other then follows from a similar argument. We know from Claim 3 that $\bar{\mu}^*$ is strongly

associated. The result then follows by induction using Lemma 18,⁴⁷ 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 Theorem 9: 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 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

$$P(S_{i,t+1} = 1 | h_{i0}^{0t}) < P(S_{i,t+1} = 1 | h_{i0}^{1t}). \quad (11)$$

Since (paying close attention to the subscripts and superscripts in the definition of h_i^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 $P(S_{i,t+1} = 1 | h_{i1}^{0t})$, (11) is equivalent to showing that

$$P(S_{i,t+1} = 1 | h_{i0}^{0t}) < P(S_{i,t+1} = 1 | h_{i1}^{0t}). \quad (12)$$

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 (12) as

$$\frac{P(S_{i,t+1} = 1, h_{i0}^{0t})}{P(S_{i,t+1} = 1, h_{i0}^{0t}) + P(S_{i,t+1} = 0, h_{i0}^{0t})} < \frac{P(S_{i,t+1} = 1, h_{i1}^{0t})}{P(S_{i,t+1} = 1, h_{i1}^{0t}) + P(S_{i,t+1} = 0, h_{i1}^{0t})}. \quad (13)$$

⁴⁵We remark that this still holds even though S does not follow a Markov process (past information about W matters and is not fully coded in the current value of S), provided we start from the steady state distribution and given that our definition of $E\tilde{P}$ allows us to transition once.

⁴⁶It is essentially $1 - b_i$ when $s_i = 1$ and is $p_i(s)$ otherwise. Without loss of generality, starting with a large enough T this is increasing.

⁴⁷While Lemma 18 does not state that the strict inequalities are preserved on given elements of the partition $\Pi(p)$, it is easy extension of the proof to see that this is true.

Rearranging terms, (13) is equivalent to

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

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 s_τ 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 (11) it is enough to show that

$$\frac{P(S_{i,t+1} = 1, h_{i0}^{0t})}{\mu^*(E_{i0}^0)} < \frac{P(S_{i,t+1} = 1, h_{i1}^{0t})}{\mu^*(E_{i1}^0)} \quad (14)$$

and

$$\frac{P(S_{i,t+1} = 0, h_{i1}^{0t})}{\mu^*(E_{i1}^0)} < \frac{P(S_{i,t+1} = 0, h_{i0}^{0t})}{\mu^*(E_{i0}^0)}. \quad (15)$$

Let us show (14), as the argument for (15) is analogous.

Then,

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

Which we rewrite as

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

Similarly

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

Note that by Theorem 7, $\mu^*(s^0 | E_{i1}^0)$ strictly dominates $\mu^*(s^0 | E_{i0}^0)$ (with some strict inequalities since i is connected to at least one other agent). Then, by the above equations, and Lemma 18 applied iteratively,⁴⁸ we derive the desired conclusion that (14) is satisfied. ▀

Proof of Lemma 10: Consider what happens when an agent i drops out. The resulting w' is dominated by the w if that agent does not drop out. Furthermore, from Lemma 18 for large enough T , the next period wage distribution over other agents when the agent drops out is dominated by

⁴⁸To be careful, at each stage we are applying the lemma to P where $P_{ss'}$ only has positive probability on s' where $s'_i = 0$, except at time $t+1$ when $s'_i = 1$. It is easy to see that Lemma 18 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 18 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 $t+1$.

that when the agent stays in, if one were to assume that the agent were still able to pass job information on. This domination then easily extends to the case where the agent does not pass any job information on. Iteratively applying this, the future stream of wages of other agents is dominated when the agent drops out relative to that where the agent stays in. This directly implies that the drop-out game is supermodular. The lemma then follows from the theorem by Topkis (1979). ■

Proof of Theorem 11: Let $w \geq w'$ and $d \in \{0, 1\}^n$. We first show that for large enough T

$$E^T [f(W_t) | W_0 = w', d] \geq E^T [f(W_t) | W_0 = w, d].$$

Lemma 17 implies that for a fine enough T -period subdivision and for every non-decreasing f ,

$$E^T [f(W_1) | W_0 = w', d] \geq E^T [f(W_1) | W_0 = w, d].$$

Lemma 18 and a simple induction argument then establish the inequality for all $t \geq 1$. The inequality is strict whenever f is increasing and $w' > w$.

Next, let $d \geq d'$. For a fine enough T -period subdivision and for every non-decreasing f , given that drop-outs have wages set to the lowest level it follows that

$$E^T [f(W_1) | W_0 = w, d'] \geq E^T [f(W_1) | W_0 = w, d]$$

As before, the inequality extends to all $t \geq 1$ by induction. Again, f increasing and $d' > d$ imply a strict inequality.

Combining these observations, we find that for large enough T when $w' \geq w$ and $d' \geq d$

$$E^T [f(W_t) | W_0 = w', d'] \geq E^T [f(W_t) | W_0 = w, d] \quad (16)$$

Consider the maximal equilibrium $d^*(w)$. By (16), for large enough T and all t

$$E^T [W_{it} | W_0 = w', d^*(w)] \geq E^T [W_{it} | W_0 = w, d^*(w)]$$

Thus,

$$\sum_t \delta_i^t E^T [W_{it} | W_0 = w', d^*(w)] \geq \sum_t \delta_i^t E^T [W_{it} | W_0 = w, d^*(w)]$$

If $d^*(w)_i = 1$, then

$$\sum_t \delta_i^t E^T [W_{it} | W_0 = w', d^*(w)] \geq \sum_t \delta_i^t E^T [W_{it} | W_0 = w, d^*(w)] \geq c_i$$

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

$$\sum_t \delta_i^t E^T [W_{it} | W_0 = w', d'] \geq c_i. \quad (17)$$

Set $d'_i = d^*(w)_i$ for any i such that $d^*(w)_i = 1$. Fixing d' for such i 's, find a maximal equilibrium at w' for the remaining i 's, and set d' accordingly. By (17), it follows that d' is an equilibrium when considering all agents. It follows that $d' \geq d^*(w)$. Given the definition of maximal equilibrium, it then follows that $d^*(w') \geq d' \geq d^*(w)$. ■