Centre de Referència en Economia Analítica

Barcelona Economics Working Paper Series

Working Paper nº 55

Networks in Labor Markets: Wage and Employment Dynamics and Inequality

Antoni Calvó-Armengol and Matthew O. Jackson

October 2001. Revision: April, 2003

Networks in Labor Markets: Wage and Employment Dynamics and Inequality*

Antoni Calvó-Armengol† and Matthew O. Jackson‡
Barcelona Economics WP n 55

October 2001
Revision: April 21, 2003§

Abstract

We present a model of labor markets that accounts for the social network through which agents hear about jobs. We show that an improvement in the wage or 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 both wages and employment are positively associated (a strong form of correlation) across time and agents. We also analyze the decisions of agents regarding staying in the labor market or dropping out. If there are costs to staying in the labor market, and we compare two networks of agents that are identical except that one group starts with a worse wage status, then that group's drop-out rate will be higher than the other's and there will be a persistent difference in wages between the groups.

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

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 for their comments and suggestions. We gratefully acknowledge the financial support of Spain's Ministry of Science and Technology under grants SEC2001-0973 and BEC2002-02130 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

§This paper was formerly part of Calvó-Armengol and Jackson (2001). That paper was split into two parts, with the part under the former title looking at a special case of the networks outlined here and focusing on employment dynamics, and this part looking at a more general set of networks and exploring both wage and employment dynamics.

1 Introduction

One of the most extensively studied issues in labor economics is the persistent inequality in wages between whites and blacks. For instance, the gap between white and black males between the ages of twenty and fifty ranges from 25% to 40% depending on education level, and this has varied some over time (see Smith and Welch (1989) for statistics from U.S. census data, broken down across a variety of dimensions and time).[1] It is clear that part of the gap can be explained by differences in factors such as the quality of education and skill levels, but nonetheless, there is still a significant residual that is unexplained (for instance, see Card and Krueger (1992) and Chandra (2000)). Moreover, once one accounts for the fact that blacks drop out of the labor force at a higher rate than whites, the true gap in wages actually increases (e.g., see Heckman, Lyons, and Todd (2000)). The idea is that the wages for drop-outs are ignored when comparing only employed workers. The fact that participation in the labor force is different across groups such as whites and blacks is also well-documented. For instance, Card and Krueger (1992) quote drop-out rates for blacks that are 2.5 to 3 times higher for blacks compared to whites (roughly 12 to 13% for blacks compared to 4 to 5% for whites); and Chandra (2000) finds similar sized differences and provides a breakdown of differences in participation rates by education level and other characteristics.

Even if one believes any inequality in wages between social groups 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.[2]

The purpose of this paper is to provide a comprehensive theory accounting both for the observed patterns of wages and employment as well as differences in drop-out patterns and their roles in sustaining inequality. We do this by explicitly modeling how social networks are used in the transmission of job information. An analysis of social networks provides a basis for observing both higher drop-out rates in one race versus another and sustained inequality in wages and employment rates even among those remaining in the labor force.

Our model builds upon a well-known stylized fact: a significant fraction of all jobs are found through contacts. While estimates of the percentage of jobs found through social contacts vary across location and profession, they consistently range between 25 and 80% of jobs in a given profession.[3] We take as given that an important source of information about jobs is a social network, and we derive the implications for wage dynamics and drop-out decisions.

More precisely, we model the transmission of job information among individuals by a function that keeps track of who first heard about a job and who (if anyone) eventually ended up getting an offer for that job. The key condition that we impose on this function is that the expected number

[1] For a view on other ethnic groups, see Farley (1990), who provides data for 50 racial-ethnic groups in the U.S..

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

[3] There are many studies relating to this, from early work in labor economics by Rees (1966) to some of the seminal work in the sociology ("social networks") literature by Granovetter (1973, 1995). Montgomery (1991) provides a nice summary of some of the key studies and statistics from that literature.

of offers that a given agent ends up with is nondecreasing in the wage status of other agents. This embodies several things. First, as the status of other agents improves, they have better access to information about jobs. Second, as the status of other agents improves, fewer of the new jobs offer improvements for those agents relative to their current status, and so they are more likely to pass information on. Third, as the status of other agents improves, our given agent becomes relatively "needier". To the extent that agents pay attention to this in passing job information, this would also make it more likely that the given agent hears about jobs.

We emphasize that the reduced form approach we take here, and the parsimony of the assumptions being made, accommodates a large variety of situations, including heterogeneous agents, jobs and wages, repeated and selective passing of information, competing offers for employment, raises due to outside offers, decisions as to whether to switch jobs, etc.

Our model takes such an information transmission network and embeds it into a dynamic setting where we keep track of job turnover. At the beginning of each period, each agent hears of new jobs according through the information transmission network. This results in new job offers to various agents, and possibly results in new wage levels through a well-defined process. Next, at the end of the period, agents randomly lose jobs according to some idiosyncratic breakup probabilities. Wages thus follows a Markov process, with state transitions depending on the information transmission network. Given a discrete wage space, we get an irreducible and aperiodic Markov process with a unique steady-state distribution on wages. We prove that this stationary distribution is strongly associated, that is, the wages of any path-connected agents are positively correlated under the steady-state distribution.[4] We also establish positive correlation of wages for path-connected agents across arbitrary time periods, meaning that, within a given social group, individual wage dynamics are (positively) correlated. These results are quite robust to the details of the actual information transmission protocol being used. The proof is not as easy as one might expect, as there is a countervailing effect that path-connected agents are sometimes in competition for information about certain jobs. This entails some within period negative correlation among the status of certain agents. So we have to prove that in the long run, the long run benefits of improved status of friends-of-friends outweighs the short run competition that they might represent.

In terms of establishing persistent inequality between wages of different types of agents, we need to analyze drop-out decisions. This is clear since the ergodicity of the Markov process governing wages implies that any short-run inequalities may disappear in the long run. However, once we allow agents to decide whether to enter the labor market or to drop out, inequalities between agents persist. We model the entry/drop-out decision as a simultaneous-move game. The decision is made by comparing the discounted expected flow of future wages stemming from entering the labor force with the corresponding discounted costs (such as education costs, opportunity costs, skills maintenance, etc.). Given that individual wages are positively associated across agents, entry decisions in this entry/drop-out game turn out to be strategic complements. Applying the theory of supermodular games, we deduce that two different social groups with identical job information

[4]To be more precise, strong association is displayed by the stationary distribution of a Markov process derived from the original one by shortening the time periods, whenever such period subdivisions are fine enough.

networks but differing in their starting wage and employment profile will have different drop-outs rates that can be strictly ranked. These differences in drop-out patterns in turn breed persistent differences in wages between the two groups. This theory thus highlights the role of collective employment history in persistent wage inequality across social groups.

This paper has a companion paper: Calvó-Armengol and Jackson (2001). In that paper we examine a specific case of the model considered here. In particular, we examine a situation where all jobs are identical and there is a single wage level, and job information is passed to at most one agent and only if that agent is unemployed. There we derive positive correlation between the employment of path-connected agents and show that differences in starting conditions can result in different drop-out decisions and persistent inequality in employment. We also provide a number of simulation results illustrating the model in the context of some simple networks. The main contributions here relative to the companion paper are twofold. First, we consider a much more general model both in the passing of job information and in the structure of wages and their relation to job offers. This is not a simple extension, but broadens things in important directions. In fact, it is important to see the results in the more general model because most any application would fit into the broader class examined here rather than the specific case examined in the companion paper. The added features here allow for things such as heterogeneity in jobs, heterogeneity in agents (skills, education, etc.), multiple offers, higher wages due to outside offers, switching of jobs, passing of information to more than one agent, competition between agents for jobs, and repeated passing of information. Second, we study wage dynamics in addition to employment and drop-out decisions. It is very important to derive implications for wages as much of the empirical evidence for inequality relates to wage differences between races, rather than employment. Moreover, we establish a strong form of association of the wages and employment of path-connected agents rather than simple correlation. Hopefully, the network-based labor model developed here can serve as a useful alternative to the competitive models and search models that are traditional to labor economics.

2 A Model of Networks in Labor Markets

We begin with a formal description of our model.

2.1 Wage and Employment Status

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, and more generally W_{it} takes on values in $I\!R_+$. The vector $w_t = (w_{1t}, \ldots, w_{nt})$ represents a realization of the wage levels a time t.

From an agent's wage status we can deduce the agent's employment status. Yet, for convenience, we introduce a binary random variable S_{it} that keeps track of the employment status of agent i at time t. We set $s_{it} = 1$ when i is employed, and $s_{it} = 0$ otherwise. So, the vector $s_t \in \{0, 1\}^n$ represents a realization of the employment status at time t. In general (with multiple wage levels),

S_t is a coarsening of W_t.

We follow the convention of representing random variables by capitol letters and realizations by small letters. Thus, the sequence of random variables $\{W_0, W_1, W_2, \ldots\}$ comprises the stochastic process of wage status.

We now discuss how wages and employment evolve over time.

2.2 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, agents hear about new jobs. 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. 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.

2.3 Specifics of Information Transmission

There are many possible variations to consider regarding 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 general, jobs have different characteristics and values to different agents. Whenever the new job constitutes an improvement for an already employed agent, this agent switches job, and so the information about the new job is not passed on.[5] 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 new job will not generate an improving offer and the agent will pass on the information about the job.

[5]It might also be that the new job leads the agent's current job to raise its wage to retain the agent.

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, depending on the current status of their 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 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.

The job transmission and offer generation is described by a function $p_{ij} : \mathbb{R}^n_+ \to [0,1]^n$. 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.

As emphasized above, 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.

For simplicity, we assume that in a given period a given agent hears about at most one job directly. This is purely a simplifying assumption, and it is quite clear that it is inconsequential in any of the proofs. And it is still possible that an agent ends up hearing about several jobs, some of which are heard of indirectly.

Assumptions on the reduced form

Let $p_i(w) = \sum_j p_{ji}(w)$. So, $p_i(w)$ represent the expected number of offers that i will get depending on the wage state in the last period being w. Implicit in this is the assumption 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 $\overline{w}$ 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 $w_i < \overline{w}_i$

(3) if $p_i(w) > p_i(w_{-j}, \widetilde{w}_j)$ for $j \neq i$, then p_i is increasing in w_j whenever $w_i < \overline{w}_i$.

(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).[6]

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.

Let us briefly describe a few examples that fit into our model.

EXAMPLE **1** *Homogeneous jobs and agents connected by a graph.*

This is the situation considered in Calvó-Armengol and Jackson (2001).

The network of connections among agents is described by a non-directed graph g, which is an $n \times n$ symmetric matrix, that is $g_{ij} = g_{ji}$. 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 $g_{ij} = 1$, 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 this example jobs are all identical (e.g., unskilled labor) and wages depend only on whether a worker is employed or not. At the beginning of each period, agent i hears of some available job

[6]Note 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).

slot with probability $a \in (0,1)$. The passing of information is as follows. When i hears of a job and is unemployed, then i takes the job. If i is employed, then he randomly picks an unemployed (direct) acquaintance and passes the information along. If all direct acquaintances are employed, then the information is lost.

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

Example **2** *Heterogeneous Agents.*

Consider the above example with some variations. The arrival rate a is specific to agents and so is denoted a_i. In addition, the network relationships may be weighted and asymmetric, so that the g_{ij}'s are not necessarily in $\{0,1\}$.

Here $p_{ij}(s)$ is as follows:

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

Example **3** *Relaying information.*

Consider a variation on Example 2 where we now allow agents to relay information in the case that all of their acquaintances are employed. For instance, if all of i's direct acquaintances are employed, then he passes the information randomly to one of them, who then picks an unemployed direct acquaintance randomly (if one exists) and relays the information to her. Otherwise the job information is lost.

Here $p_{ij}(s)$ is as follows:

$$p_{ij}(s) = \begin{cases} a_i & \text{if } j = i \text{ and } s_i = 0; \\ a_i \frac{g_{ij}}{\sum_{k:s_k=0} g_{ik}} & \text{if } s_i = 1,\ g_{ij} = 1 \text{ and } s_j = 0; \\ a_i \sum_{k:g_{kj} \neq 0} \frac{g_{ik}}{\sum_{k' \neq i} g_{ik'}} \frac{g_{kj}}{\sum_{k'':s_{k''}=0} g_{kk''}} & \text{if } s_i = 1,\ s_k = 1 \text{ for all } k \text{ such that } g_{ik} \neq 0, \text{ and } s_j = 0; \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

Example **4** *Heterogeneous jobs*

Again, consider Example 2, but now where jobs are heterogeneous and wages may take on different values. Denote by $\overline{w}_i$ the highest wage attainable by agent i. At the beginning of each period, agent i hears about a job opening offering a wage w_i with probability $a_i^{w_i} \in (0,1)$.

If i directly hears about a new job that pays a higher wage than i's current job position, then agent i keeps that information.[7] If the new job does not offer any improvement, then agent i

[7] A given agent may end up with offers for several jobs, as we discuss next. So holding on to information does not necessarily imply that an agent takes that job.

randomly passes the information on to one of his direct acquaintances who currently has a wage lower than that of the new job, with weights reflecting the relative link intensities.

Here $p_{ij}(w)$ is as follows (now needing to be written as a function of the vector of beginning wages $w = (w_1, \ldots, w_n)$ rather than just the vector of employment s):

$$p_{ij}(w) = \begin{cases} \sum_{w_i' > w_i} a_i^{w_i'} & \text{if } j = i \ ; \\ \sum_{w_i' : w_i \geq w_i' > w_j} \left(a_i^{w_i'} \frac{g_{ij}}{\sum_{k: w_k < w_i'} g_{ik}} \right) & \text{if } w_i > w_j \text{ and } g_{ij} \neq 0 \text{; and} \\ 0 & \text{otherwise.} \end{cases}$$

These are just a few examples to give some idea of the kinds of things admitted. As should be clear from the conditions on the p_i's, one can also include things like passing information to several agents at once (the p_{ij}'s then represent the chance that an agent obtains an offer when competing with whomever else has also heard about a given job), multiple relays of information, situations where the weights depend on agents' current wages, situations where the arrival rates of job information (e.g., $a_i^{w_i}$ in Example 4) depend on the current state of wages,...

2.4 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.[8]

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 $w_i : I\!R_+ \times \{0, 1, 2, \ldots\} \to I\!R_+$ 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 $w_i(W_{i,t-1}, O_{it}) \geq W_{i,t-1}$.

[8] So the number of job opportunities that some agent j ends up with, O_{jt}, is the sum of binomial random variables. These binomial random variables are that some agent i originally heard about a job and that information ended up in j's hands, which from our definitions has probability $p_{ij}(w_t)$ of occuring.

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. In general, the wage might be effectively increasing in the number of offers an agent for at least two different reasons. First, the best match of a larger set of offers is likely to be better; second, if an agent has several potential employers then competition between them will bid the wage up.[9]

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 $\overline{w}_i > 0$. 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 Examples 1 to 4, the network of relationships among individuals is modeled as a graph and the p function is determined by geometry of the graph. In general, the agents' interrelationships may be quite complicated and more easily described by the p function than any graph. 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 $p_i(w) \neq p_i(w_{-j}, \widetilde{w}_j)$ for some w and $\widetilde{w}_j$.

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

[9] One 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.

that their statuses directly or indirectly affect each other's probability of hearing about a job.[10]

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 $j \in N^n(p)$.

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 $\{S_1, S_2, \ldots\}$ and wages $\{W_1, W_2, \ldots\}$ 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.

3 Useful Observations Regarding Wage and Employment Distributions

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

[10]Note 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.

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 **1** *Consider any economy (N, p, b), time $t > 0$, two wage states $w \in I\!R^n_+$ and $w' \in I\!R^n_+$ and an agent i who is unemployed in both states ($w_i = w'_i = 0$). If $w'_j \geq w_j$ for all $j \in N^2_i$, 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 1 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 1 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 $w \in I\!R^n_+$ 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. In Example 4, this one period domination condition is satisfied at w for some i if $w_i < \overline{w}_i$ implies that, for each k, $g'_{ki} \geq g_{ki}$ and $g_{kj} \geq g'_{kj}$ for each $j \neq i$ such that $w_j < \overline{w}_j$.

LEMMA **2** *Consider an economy (N, p, b) and an alternative social structure p' that one-period dominates p at $w \in I\!R^n_+$ 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 $w_i < \overline{w}_i$, then the first order stochastic dominance is strict.*

Lemma 2 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.

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

4 The Dynamics and Patterns of Wage and Employment

We begin with patterns of wages as the results on employment have an added complication that we will discuss shortly.

4.1 Wage Patterns and Dynamics

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, as in Example 4. 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 Subvidisions

A natural way to analyze shortened periods is simply by dividing p and b by some T.[11]

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. 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.

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*.

A probability measure μ describing a random vector (say W defined on $\mathbb{R}^n$) is *associated* if

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

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

Association tells us that good news about the state (conditioning on $g(w) \geq d$) leads us to higher beliefs about the state in the sense of domination. If $W_1, \ldots, W_n$ are the random variables

[11] In the limit, this simply approximates a continuous time Poisson arrival process.

described by an associated measure μ, then we say that $W_1, \ldots, W_n$ 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 $W_i, i \in \{i_1, \ldots, i_\ell\}$ about any subset or combinations of agents (here, $\{i_1, \ldots, i_\ell\}$) 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 $\{1, \ldots, n\}$ that captures which random variables might be positively related.

A probability measure μ describing a random vector 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

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

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.

We are now ready to state our first theorem. 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 **1** *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 1 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 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, there are no short-run negative conditional correlations. 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 1 provides results on the steady state distribution, we can deduce similar statements about the relationships between wages at different times.

THEOREM **2** *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*

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

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 2 is similar to Theorem 1 in its structure, it provides different implications. Theorem 1 addresses the steady state distribution, or the expected long run behavior of the system. Theorem 2 addresses any arbitrary dates in the system.[12]

It is important in Theorem 2 that we start from the steady state distribution.

4.2 Employment Patterns and Dynamics

One might conjecture (as we initially did) that it would be a simple Corollary to Theorem 1 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

[12] Theorem 1 almost seems to be a corollary of Theorem 2, 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 1 from Theorem 2 since it is not ruled out that the positive correlation vanishes in the limit under Theorem 2, while we know that this is not the case from Theorem 1.

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.

EXAMPLE **5** *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).[13] 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 **3** *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.*[14]

[13] That 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.

[14] Having 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.

- *If p can be written as a function of S,*[15] *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).*

Theorem 3 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 5.

We also have an analog of Theorem 2, stating that the positive interrelationships between employment statuses hold both under the steady distribution and at any time along the dynamics.

THEOREM **4** *Consider any economy (N, p, b) such that p is a function of employment status.*[16] *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 $T \geq T'$*

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

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 Dropping Out and Long-Run Inequality

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.

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.[17] 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 $i \in \{1, \ldots, n\}$, $d_i = 1$ implies

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

[15](3) is relaxed to hold relative to S rather than W.

[16]The result also holds for connected agents without this assumption.

[17]This 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.

and $d_i = 0$ implies the reverse inequality.

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

Lemma 3 *Consider any economy* (N, p, b)*, state* $W_0 = w$*, and vector of costs* $c \in I\!R^n_+$*. 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 3 as the *maximal* equilibrium.

Theorem 5 *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 stochastic 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* $f : I\!R^n_+ \to I\!R$ *and any* t

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

with strict inequality for some specifications of c *and* δ*.*

Theorem 5 shows how persistent inequality can arise between two otherwise similar groups with different initial employment conditions.

References

[1] Aliprantis C.D. and K.C. Border (2000) *Infinite Dimensional Analysis*, Springer-Verlag.

[2] Arrow, K. and R. Borzekowski (2000) "Limited Network Connections and the Distribution of Wages," mimeo: Stanford University.

[3] Calvó-Armengol, A. (2000) "Job Contact Networks," mimeo, http://www.eco.uc3m.es/∼ancalvo/jobnet.pdf

[4] Calvó-Armengol, A. and M.O. Jackson (2001) "Social Networks in Determining Employment: Patterns, Dynamics, and Inequality," mimeo: Caltech and Universitat Autònoma de Barcelona, http://www.hss.caltech.edu/∼jacksonm/dynet.pdf

[5] Card, D. and Krueger (1992) "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics*, 151-200.

[6] Chandra, A. (2000) "Labor-Market Dropouts and the Racial Wage Gap: 1940-1990," *American Economic Review*, Papers and Proceedings, 333-338.

[7] 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.

[8] Esary, J.D., F. Proschan, and D.W. Walkup (1967) "Association of Random Variables, with Applications," *Annals of Mathematical Statistics*, 38, 1466-1474.

[9] Farley, R. (1990) "Blacks, Hispanics, and White Ethnic Groups: Are Blacks Uniquely Disadvantaged?," *American Economic Review*, Papers and Proceedings, 80, 237-241.

[10] Freidlin, M. and A. Wentzel (1984) *Random Perturbations of Dynamical Systems*, Springer-Verlag, New-York.

[11] Granovetter, M. (1973) "The Strength of Weak Ties," *American Journal of Sociology*, 78, 1360-1380.

[12] Granovetter, M. [1974](1995) *Getting a Job: A Study of Contacts and Careers*, 2^{nd} edition, University of Chicago Press, Chicago.

[13] Heckman, J.J. (1998) "Detecting Discrimination," *Journal of Economic Perspectives*, 12, 101-116.

[14] 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.

[15] Montgomery, J. (1991) "Social Networks and Labor Market Outcomes: Toward and Economic Analysis," *American Economic Review*, 81, 1408-1418

[16] Rees, A. (1966) "Information Networks in Labor Markets," *American Economic Review*, 56, 559-566.

[17] Smith and Welch (1989) "Black Economic Progress after Myrdal," *Journal of Economic Literature*, 27, 519-564.

[18] Szekli, R. (1995) *Stochastic Ordering and Dependence in Applied Probability*, Springer-Verlag (Lecture Notes in Statistics # 97), New York.

[19] Topkis, D. (1979) "Equilibrium points in nonzero-sum $n-$person submodular games," *SIAM Journal of Control and Optimization* 17, 773-787.

[20] Young, H.P. (1993) "The Evolution of Conventions," *Econometrica* 61, 57-84.

Appendix

Proof of Lemmas 1 and 2: 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

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

Simplifying, we obtain

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

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 $\text{Prob}\left(\{O_{it} \geq h\} \mid p', W_{t-1} = w'\right)$ is at least as large $\text{Prob}\left(\{O_{it} \geq h\} \mid p, W_{t-1} = w\right)$ for each h. Strict dominance follows if there is a strict inequality for any h.

Note that from (1) we can write $\text{Prob}\left(\{O_{it} \geq h\} \mid p, W_{t-1} = w\right)$ 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 $\text{Prob}\left(\{O_{it} \geq h\} | \, p, W_{t-1} = w\right)$, 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

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

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 3 and 1.

Consider two probability measures μ and ν on a state space that is a subset of $I\!R^n$.

μ *dominates* ν if

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

for every non-decreasing function $f : I\!R^n \to I\!R$.[18] 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 4 *Consider two measures μ and ν on $I\!R^n$ which have supports that are a subset of a finite set $W \subset I\!R^n$. μ dominates ν if and only if there exists a Markov transition function $\phi : W \to \mathcal{P}(W)$ 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 4: 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 5 *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

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

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

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 $\pi \in \Pi$.[19]

Proof of Lemma 5: First, suppose that for every $E \in \mathcal{E}$:

$$\mu(E) \geq \nu(E). \tag{2}$$

Consider any non-decreasing f. Let the elements in its range be enumerated $r_1, \ldots, r_K$, with $r_K > r_{k-1} \ldots > r_1$. 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)). ∎

Lemma 6 *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 $Cov_\mu(W_i, W_j) > 0$, then $Cov_\mu(f, g) > 0$.*

[19] E is sensitive to π if its indicator function is. A nondecreasing function $f : I\!R^n \to I\!R$ is *sensitive to* $\pi \in \Pi$ (relative to μ) if there exist x and $\widetilde{x}_\pi$ such that $f(x) \neq f(x_{-\pi}, \widetilde{x}_\pi)$ and x and $x_{-\pi}, \widetilde{x}_\pi$ are in the support of μ.

Proof of Lemma 6: 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 $Cov_\mu (W_i, W_j) > 0$, then $Cov_\mu (f(W), g(W)) > 0$.*

Proof of Claim 1: We write

$$\text{Cov}_\mu (W_i, W_j) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \text{Cov}_\mu \left(I_{W_i}(s), I_{W_j}(t) \right) dsdt,^{20}$$

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, $\text{Cov}_\mu \left(I_{W_i}(\overline{s}), I_{W_j}(\overline{t}) \right) > 0$ for a set of $\overline{s}, \overline{t}$'s. Also,

$$\text{Cov}_\mu (f(W_i), g(W_J)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \text{Cov}_\mu (I_f(s), I_g(t)) \, dsdt, \tag{3}$$

where $I_f(s) = 1$ if $f(W_i) > s$, and $I_f(s) = 0$, otherwise. For each $\overline{s}$ as described above, there exists some s' such that $I_{W_i}(\overline{s}) = 1$ if and only if $I_f(f(s')) = 1$, and similarly for $\overline{t}$, 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 $\overline{s}, \overline{t}$'s and corresponding s', t''s are unions of closed intervals with nonempty interiors. By association also we know that $\text{Cov}_\mu (I_f(f(\overline{s})), I_g(g(\overline{t}))) \geq 0$ 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 $w_i^K = \overline{w}_i$. 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 $f'(w_i^k) = f(0, \ldots, 0) + k\frac{\gamma}{2}$. 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

$$Cov(f, g) = Cov(f'', g'') + Cov(f'', g') + Cov(f', g'') + Cov(f', g').$$

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 $W_{t-1} = w$.
Let $P_{wE}^T = \sum_{w' \in E} P_{ww'}^T$.

LEMMA **7** *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.

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

Proof of Lemma 7: 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

$$P^T_{w'E} \geq P^T_{wE}.$$

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

$$P^T_{w'E} = \sum_o Prob^T_{w'}(W_t \in E | O_t = o) Prob^T_{w'}(O_t = o)$$

and similarly

$$P^T_{wE} = \sum_o Prob^T_{w}(W_t \in E | O_t = o) Prob^T_{w}(O_t = o),$$

where $Prob^T_w$ 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 $Prob^T_{w'}(O_{-\ell,t})$ dominates $Prob^T_w(O_{-\ell,t})$. 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, $Prob^T_{w'}(O_{-\ell,t} \geq a) \geq Prob^T_w(O_{-\ell,t} \geq a+1)$. 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).[21]

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 **8** *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.*

Proof of Lemma 8:

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

By Lemma 4 we rewrite this as

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

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

We rewrite this as

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

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

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

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

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

Lemma 7 implies that for large enough T, $P^T_{wE} \geq P^T_{w'E}$ 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 4 that there exists some $w \neq w'$ such that $\phi_{w'w} > 0$. By Lemma 7, there exists some $E \in \mathcal{E}$ such that $P^T_{wE} > P^T_{w'E}$. Then $[\mu P^T](E) > [\nu P^T](E)$ for such E, implying (by Lemma 5) that μP^T strictly dominates νP^T. ∎

We prove Theorem 1 and then Theorem 3, as the latter makes use of the proof of the former.

Proof of Theorem 1: 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 $\overline{P}^T$ 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 $\overline{\mu}^T$ be the associated (again unique) steady state distribution, and $\overline{\mu}^* = \lim_T \overline{\mu}^T$ (which is well-defined as shown in the proof of Claim 2 below).

The following claims establish the theorem.

Claim **2** $\overline{\mu}^* = \mu^*$.

Claim **3** $\overline{\mu}^*$ *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

$$p_z = \sum_{\tau \in \mathcal{T}_z} \left[\times_{z',z'' \in \tau} P_{z'z''} \right]. \tag{4}$$

LEMMA **9** Freidlin and Wentzel (1984)[22]: *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 $P^T_{w',w''}$ goes to 0 at the rate $1/T$ when w' and w'' are adjacent,[23] and other transition probabilities go to 0 at a rate of at least $1/T^2$, it follows from Lemma 9 that $\mu^T(w)$ may be approximated for large enough T by

$$\frac{\sum_{\tau \in \mathcal{T}_w^*} \left[\times_{w',w'' \in \tau} P^T_{w'w''} \right]}{\sum_{\widehat{w}} \sum_{\tau \in \mathcal{T}^*_{\widehat{w}}} \left[\times_{w',w'' \in \tau} P^T_{w'w''} \right]}.$$

Moreover, note that for large T and adjacent w' and w'', $P^T_{w'w''}$ 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 $\widetilde{P}^T_{w'w''} = \frac{b_i}{T}$ when $w'_i > w''_i$, and $\frac{p_i(w')}{T}$ when $w'_i < w''_i$.[24] It then follows that

$$\mu^*(w) = \lim_{T \to \infty} \frac{\sum_{\tau \in \mathcal{T}_w^*} \left[\times_{w',w'' \in \tau} \widetilde{P}^T_{w'w''} \right]}{\sum_{\widehat{w}} \sum_{\tau \in \mathcal{T}^*_{\widehat{w}}} \left[\times_{w',w'' \in \tau} \widetilde{P}^T_{w'w''} \right]}. \tag{5}$$

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

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

$$\overline{\mu}^*(w) = \lim_{T \to \infty} \frac{\sum_{\tau \in \mathcal{T}_w^*} \left[\times_{w',w'' \in \tau} \widetilde{P}^T_{w'w''} \right]}{\sum_{\widehat{w}} \sum_{\tau \in \mathcal{T}^*_{\widehat{w}}} \left[\times_{w',w'' \in \tau} \widetilde{P}^T_{w'w''} \right]}.$$

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

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

where $\widehat{P}^T$ is set as follows. For adjacent w' and w'' (letting i be the agent for whom $w'_i \neq w''_i$) $\widehat{P}^T_{w'w''} = b_i$ when $w'_i > w''_i$, and $p_i(w')$ when $w'_i < w''_i$,[25] and $\widehat{P}^T_{w'w''} = 0$ for non-adjacent w' and w''.

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

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

[23]Note that under property (3) of p, since w' and w'' are adjacent, it must be that $P^T_{w',w''} \neq 0$.

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

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

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

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

Proof of Step 1: We show that for any T and any associated μ, $\mu\overline{P}^T$ is associated. From this, it follows that if we start from an associated μ_0 at time 0 (say an independent distribution), then $\mu_0(\overline{P}^T)^k$ is associated for any k. Since $\overline{\mu}^T = \lim_k \mu_0(\overline{P}^T)^k$ for any μ_0 (as $\overline{\mu}^T$ is the steady-state distribution), and association is preserved under (weak) convergence,[26] this implies that $\overline{\mu}^T$ is associated for all T. Then again, since association is preserved under (weak) convergence, this implies that $\lim_T \overline{\mu}^T = \overline{\mu}^*$ is associated.

So, let us now show that for any T and any associated μ, $\nu = \mu\overline{P}^T$ is associated. By Lemma 5, we need to show that

$$\nu(EE') - \nu(E)\nu(E') \geq 0 \tag{7}$$

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

$$\nu(EE') - \nu(E)\nu(E') = \sum_w \mu(w)\left(\overline{P}^T_{wEE'} - \overline{P}^T_{wE}\nu(E')\right).$$

Since W_t is independent conditional on $W_{t-1} = w$, it is associated.[27] Hence,

$$\overline{P}^T_{wEE'} \geq \overline{P}^T_{wE}\overline{P}^T_{wE'}.$$

Substituting into the previous expression we find that

$$\nu(EE') - \nu(E)\nu(E') \geq \sum_w \mu(w)\left(\overline{P}^T_{wE}\overline{P}^T_{wE'} - \overline{P}^T_{wE}\nu(E')\right).$$

or

$$\nu(EE') - \nu(E)\nu(E') \geq \sum_w \mu(w)\overline{P}^T_{wE}\left(\overline{P}^T_{wE'} - \nu(E')\right). \tag{8}$$

Under the properties of the p_{ij}'s, both $\overline{P}^T_{wE}$ and $\left(\overline{P}^T_{wE'} - \nu(E')\right)$ are non-decreasing functions of w. Thus, since μ is associated, it follows from (8) that

$$\nu(EE') - \nu(E)\nu(E') \geq \left[\sum_w \mu(w)\overline{P}^T_{wE}\right]\left[\sum_w \mu(w)\left(\overline{P}^T_{wE'} - \nu(E')\right)\right].$$

Then since $\sum_w \mu(w)\left(\overline{P}^T_{wE'} - \nu(E')\right) = 0$ (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,

$$Cov_{\overline{\mu}^*}(f,g) > 0.$$

By Lemma 6 it suffices to verify that

$$Cov_{\overline{\mu}^*}(W_i, W_j) > 0$$

[26] See, for instance, P5 in Section 3.1 of Szekli (1995).

[27] See, for instance, P2 in Section 3.1 of Szekli (1995).

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

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

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

$$Cov_{ij}^T > 0.$$

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

Let

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

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

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

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,

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

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

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

It follows that

$$\lim_{T\to\infty} Cov_{ij}^T \geq \sum_w \overline{\mu}^*(w)\widetilde{P}_{wi}\widetilde{P}_{wj} - \sum_w \overline{\mu}^*(w)\widetilde{P}_{wi} \sum_{w'} \overline{\mu}^*(w')\widetilde{P}_{w'j}.$$

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

$$\sum_w \overline{\mu}^*(w)\widetilde{P}_{wi}\widetilde{P}_{wj} > \sum_w \overline{\mu}^*(w)\widetilde{P}_{wi} \sum_{w'} \overline{\mu}^*(w')\widetilde{P}_{w'j}. \tag{9}$$

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

First let us consider the case where $j \in N_i(p)$.[28] We know that $\widetilde{P}_{wi}$ is increasing in w_i, and also, given the assumptions on p, that $\widetilde{P}_{wi}$ is increasing ni w_j for $j \in N_i(p)$. Similarly, $\widetilde{P}_{wj}$ is increasing in w_j. (9) then follows from Lemma 6 (where we apply it to the case where $W_i = W_j$), as both $\widetilde{P}_{wi}$ and $\widetilde{P}_{wj}$ are increasing in w_j.

Next, consider any $k \in N_j(p)$. Repeating the argument above, since $\widetilde{P}_{wj}$ is increasing w_j we apply Lemma 6 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 1 now follows from Claim 3 since $\mu^T \to \overline{\mu}^*$. ▮

Proof of Theorem 3: For the case where p depends only on S, the proof is an analog of the proof of Theorem 1. For the more general case, the association of the limiting distribution follows directly from the proof of Theorem 1. 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 $\overline{\mu}^*$ defined on W, let

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

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

Next, let

$$E\widetilde{P}_{s_ij} = \sum_w \overline{\mu}^*(w|s_i) \sum_{w'} \widetilde{P}_{ww'} S_j(w') \tag{10}$$

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

$$E[S_k] = \sum_{s_i} \overline{\mu}^*(s_i) E\widetilde{P}_{s_ik}.$$

Then, following steps similar to those in Step 2 we can write[29]

$$Cov_{\overline{\mu}^*}(S_i, S_j) \geq \sum_{s_i} \overline{\mu}^*(s_i) E\widetilde{P}_{s_ii} E\widetilde{P}_{s_ij} - \sum_{s_i} \overline{\mu}^*(s_i) E\widetilde{P}_{si} \sum_{s_i'} \overline{\mu}^*(s_i') E\widetilde{P}_{s_i'j}.$$

The remainder of the proof then follows the same lines as that of Theorem 1. $E\widetilde{P}_{s_ii}$ is clearly increasing in s_i.[30] There we need to employ (10). We note that under association $\overline{\mu}^*(w|s_i)$ is

[28] If 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.

[29] We 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 $E\widetilde{P}$ allows us to transition once.

[30] It 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.

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 $E\widetilde{P}_{s_ij}$ is increasing in s_i. ∎

Proof of Theorems 2 and 4: We show Theorem 2, as the other then follows from a similar argument. We know from Claim 3 that $\overline{\mu}^*$ is strongly

associated. The result then follows by induction using Lemma 8,[31] and then taking a large enough T so that μ^T is close enough to $\overline{\mu}^*$ for the desired strict inequalities to hold. ∎

Proof of Lemma 3: 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 8 for large enough T, the next period wage distribution over other agents when the agent drops out is dominated by 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 5: Let $w \geq w'$ and $d \in \{0,1\}^n$. We first show that for large enough T

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

Lemma 7 implies that for a fine enough T-period subdivision and for every non-decreasing f,

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

Lemma 8 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

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

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 $d' \geq d$

$$E^T\left[f(W_t)\,|W_0 = w', d'\right] \geq E^T\left[f(W_t)\,|W_0 = w, d\right] \tag{11}$$

Consider the maximal equilibrium $d^*(w)$. By (11), for large enough T and all t

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

[31] While Lemma 8 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.

Thus,

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

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

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

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

$$\sum_t \delta_i^t E^T \left[W_{it} \left| W_0 = w', d' \right. \right] \geq c_i. \tag{12}$$

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 (12), 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 $d^*(w') \geq d' \geq d^*(w)$. ∎