Social Networks and Resulting Patterns and Dynamics of Employment and Wages*

Antoni Calvó-Armengol† and Matthew O. Jackson‡

Preliminary: October 2001
Draft: February 13, 2002

Abstract

We develop a simple 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 and wages are positively correlated across time and agents; and that the strength of the correlation depends on the proximity in the network. 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, 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: Network, Employment, Unemployment, Wages, Wage Inequality, Duration Dependence.
JEL Classification Numbers:.

---

*We thank Antonio Cabrales, David Pérez-Castrillo, and Eddie Lazear, for helpful conversations and discussion. We also thank seminar participants at Caltech, U.C. San Diego, Stanford, Venice, and Yale for their comments and suggestions. We gratefully acknowledge the financial support of Spain's Ministry of Education under grant SEC2001-0973.

†Department of Economics, Universidad Carlos III, C/ Madrid 126, 28903 Getafe (Madrid), Spain and CERAS-ENPC, 28 rue des Saints-Pères, 75007 Paris, France. Email: ancalvo@eco.uc3m.es, www.eco.uc3m.es/∼ancalvo/jobnet.pdf

‡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.

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 of their direct acquaintances 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 exit 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.

We show that this simple model exhibits a variety of characteristics that are consistent with previous empirical studies on employment and wages. There are four characteristics that we focus on:

- Employment and wages are both positively correlated across agents within and across periods. This results from the fact that information is passed to a direct acquaintance when an agent is employed, so that an agent is more likely to find employment if his or her direct acquaintances are employed. This is in turn linked to the employment status of indirect acquaintances, and so on.

  The clustering of employment and inequality of wages across various races and social groups is well-studied (see the discussion of inequality below). The correlation across likely social contacts is also documented in recent work by Topa (2001) who demonstrates geographic correlation in unemployment across neighborhoods tracts in Chicago, and founds a significantly positive amount of social interactions across such neighborhoods.

- Unemployment exhibits duration dependence and persistence. That is, when conditioning on 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. This happens since a longer history of unemployment is more likely to come when direct and indirect contacts are unemployed, which lowers the 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)).[1] 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.

[1] Heckman and Borjas (1980) (and some of the literature that followed) distinguish between duration dependence based on endogenous economic factors such as skills deterioration, and that based on statistical aspects such as unobserved heterogeneity. We come back to discuss how our model fits into this view.

- A poor social environment reinforces the incentives to withdraw from the labor force. If staying in the labor market is costly (in terms of opportunity costs, education costs, skills maintenance, etc.) and we compare two networks that are identical except that one network starts with a worse employment status than the other, then the first network's drop-out rate will be higher.

  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 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.

- Higher drop-out rates are consistent with persistent employment and wage inequality. Comparing across networks, if the initial starting state of the network is worse for one group, then 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 that group. Thus, inequality in wages and employment will persist. This inequality results from the higher drop-out rate, as those remaining in the labor force end up with a network with fewer direct and indirect acquaintances, and thus worse prospects for obtaining job information through their social network.

  The persistent inequality in wages between whites and blacks is one of the most extensively studied areas in labor economics. Smith and Welch (1989) provide statistics breaking the gap down a variety of dimensions and across time from census data. The gap varies from 25% to 40%, and can only be partly explained by differences in skill levels, quality of education, and other factors (e.g., see Card and Krueger (1992), Chandra (2000), Heckman, Lyons and Todd (2000)).

As mentioned above, the importance of social networks in labor markets has been recognized. However, while there are models that have taken transmission of job information seriously (e.g., Boorman (1975), Montgomery (1991, 1992, 1994), Calvó-Armengol (2000), Arrow and Borzekowski (2001), Topa (2001)), this is the first to study explicit networks and the resulting implications for the pattern and dynamics of employment, wages and inequality across races.

Also, for each of the stylized facts above (with the exception of the clustering) there are other models and explanations that have been offered in the literature. There are several things to say about the contribution of the networks model. First, the range of implications that it provides is quite wide, while most models are aimed at specific aspects of the labor market. Second, many of the predictions it makes are complementary to previous models. For instance, one (among a number of) explanations that has been offered for duration dependence is unobserved heterogeneity.[2] A simple variant of unobserved heterogeneity is that agents have idiosyncratic features that are relevant to their attractiveness as an employee and are unobservable to the econometrician but observed by employers. With such idiosyncratic features some agents will be quickly re-employed while others will have longer spells of unemployment, and so the duration dependence is due simply 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. Thirdly,

[2]Theoretical models predicting duration dependence, though, are scarce. 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.

it 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.[3] Fourth, it provides new angles on policy predictions 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's network, improves the outlook for that agent. This provides a sort of local increasing returns to subsidizing education, and other policies like affirmative action.[4] One implication is that it may be more efficient to concentrate subsidies or programs so that many agents who are interconnected in a network are affected, rather than spreading resources more broadly so that only a small fraction of agents in any part of a network are affected.

2 A Model of a Network of Labor Market Contacts

We emphasize at this point that the basic model we develop is a purely mechanical one. That is, there is no overtly strategic behavior on the part of agents. The network structure is fixed in place and information flows through it. We simply characterize the resulting stochastic process on employment and wages and discuss how this depends on initial conditions and the structure of the network. This does not mean that the model is inconsistent with rational behavior. Every part of the model can be rationalized, and we discuss the ideas behind each unmodelled feature of the network. We keep the model relatively simple and stark as we wish to emphasize how the network of information transmission alone can result in interesting dynamics and patterns in employment and wages. Later in the paper we come back to introduce some strategic choices: we introduce participation decisions on the part of agents and discuss an equilibrium.

$N = \{1, \ldots, n\}$ is a set of agents.

2.1 Employment Status

Time evolves in discrete periods, $t \in \{0, 1, 2, \ldots\}$.

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 capitol letters and realizations by small letters. Thus, the sequence of random variables $\{S_0, S_1, S_2, \ldots\}$ comprise the stochastic process of employment status.

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

[4] In 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). Of course, our model is not meant to be comprehensive and so we are simply pointing out some implications of a network model.

2.2 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 $w_t = (w_{1t}, \ldots, w_{nt})$ represents a realization of the wage levels a time t.

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.

2.3 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 acquaintances in the network. Probabilities $p_{ij}(w)$ keep track of the probability that i directly hears about a job and then passes that information on to agent j, given the last period wage status.[5] 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.

2.4 Specifics of Information Transmission

There are many possible variations to consider on how information is transmitted and how information affects wages. There are two important dimensions that we consider.

One dimension to consider is whether or not an already employed agent can make use him or herself of information of a new job. In the case of completely homogeneous jobs (where jobs are fully equivalent and interchangeable, as in Example 1 below), 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

[5] Note 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 acquaintances.

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 acquaintances, or may instead select among all of his or her acquaintances 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 acquaintance. 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. It is also possible that the agent passes the job information to more than one acquaintance, and even that they indirectly pass it on to others. We discuss this in more detail below.

We begin with a simple example, as it helps to illustrate the basic structure before discussing the more general model.

Example 1 *Unskilled Labor: Homogeneous Job Networks*

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. So, there is a single value $\overline{w} > 0$ such that wages are 0 if an agent is unemployed, and $\overline{w}$, otherwise, regardless of the number of offers received in the past or past wages. So here, the variables S_t and W_t are equivalent in terms of the information that they convey.

Agents keep any news regarding job openings if they are unemployed and otherwise pass news to an unemployed acquaintance. In choosing who to pass the information to, they may favor some acquaintances (e.g., are more likely to pass information to an unemployed relative than an unemployed friend), but they do so in a consistent manner.

In particular, acquaintances are described by the network g, which is an $n \times n$ real valued matrix, where $g_{ij} > 0$ if 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, then i may tell j about the job. As formulated, the network may be directed and may also include intensities of relationships. In some cases of interest, one should expect such social relationships to be reciprocal. Such non-directed networks have $g_{ij} > 0$ if and only if $g_{ji} > 0$. Let $U(w) = \{j | w_j = 0\}$ denote the unemployed workers at state w.

The passing of information from i to j is described by the following probability $p_{ij}(w)$

$$p_{ij}(w) = \begin{cases} a_i(w) & \text{if } j = i \text{ and } w_i = 0, \\ \frac{a_i(w) g_{ij}}{\sum_{k \in U(w)} g_{ik}} & \text{if } w_i > 0,\ w_j = 0, \text{ and} \\ 0 & \text{otherwise,} \end{cases}$$

where $a_i(w)$ is the probability that agent i directly hears about a job in the given period as a function of last period's state w.

So, agents keep any information they hear about directly if they are unemployed. If they are already employed, then they pass the information along to their unemployed acquaintances. Here the weighting by which an agent decides who to pass a job to is proportional to the intensity of the relationship g_{ij}.

The General Model of Information Passing

In order to capture a much wider class of information passing possibilities, we model the job transmission in a general way that allows for a variety of special cases.

The job transmission and offer generation is described by the function $p_{ij} : I\!\!R_+^n \to [0,1]^n$. 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.

The function p_{ij} *is a reduced form that can accommodate a very 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 acquaintance at a time or they may tell several acquaintances about the job. These acquaintances 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 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) $1 \geq \sum_j p_{ij}(w)$, and

(2) $p_i(w)$ is nondecreasing in w_{-i} and nonincreasing in w_i.

(1) requires the probabilities to sum to less than one. Implicit in this is the requirement that i hears about at most one job directly in a period. As we can (and will) adjust periods to be arbitrarily small, this is effectively without loss of generality. Allowing the probabilities to sum to less than one allows for the possibility that i does not hear about a job in a period, or that that the information does not generate an offer for any agent.

(2) imposes two requirements. The first is that the probability 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 (2) 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,

[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 conditions (1) and (2).

etc., which are accounted for in the p_{ij}'s. Condition (2) only describes how changes in agents' current circumstances affect job transmission.

2.5 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

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 $w_i(W_{i,t-1}, O_{it}) \geq W_{i,t-1}$.

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 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.[7]

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 really viable ones.

For simplicity in what follows, we assume that w_i only 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 $w' = w_i(w, 1)$. This means that wages are delineated so that an agent may reach the next higher wage level with one offer.

We 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).

[7] 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.

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.

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) 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.

Throughout the paper we fix the starting state to be one of unemployment $W_0 = S_0 = (0, \ldots, 0)$. This has no effect on long run distributions of the processes, but is helpful in the proofs since then the first period distributions of wages and employment (in many cases of interest) are independent.

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 acquaintance 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 ise *acquainted* with j if $p_i(w) \neq p_i(w_{-j}, \widetilde{w}_j)$ for some w and $\widetilde{w}_j$.

Since we allow for the possibility of indirect passing of information in the p_{ij}'s, the term "acquaintance" might be a bit of a misnomer. Nevertheless, we use the term for lack of a better one.

Let

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

In many cases where p is fixed, we simply write N_i.

It is natural to focus on situations where acquaintance 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.[8]

We can also keep track of further levels of this "acquaintance" 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 the indirect acquaintance 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 $\Pi(p)$.

[8]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 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 such paths.

3 The Distributions of Employment and Wages

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

Employment states and wage states follow a Markov process, where current states are the wage state. The following lemmas describe that process 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 $t > 0$, and two states $w \in I\!R^n_+$ and $w' \in I\!R^n_+$ and i such that $w_i = w'_i = 0$ (so i is unemployed in both states). If $w'_j \geq w_j$ for all $j \in N^2_i$, then the distribution of 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 $p_{ki}(w') \neq p_{ki}(w)$ for some k (possibly even i), then the first order stochastic dominance is strict.*

Lemma 1 says that improving the wage status of any of an agent's direct or indirect acquaintances 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 $k \in N_i(p)$.

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 acquaintances, the long run comparison between p and p' might differ from the one period comparison.

As an example, under homogeneous job networks (Example 1), 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 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 i's perspective for some i. The distributions of S_{it}, O_{it} and W_{it} conditional on $W_t = w$ under p' first order stochastically dominate the corresponding distributions under p. If $p'_{ki}(w) \neq p_{ki}(w)$ for some k (possibly i) and $w_i = 0$, then the first order stochastic dominance is strict.*[9]

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.

The lemmas above show how the one-period-ahead employment and wage status of an agent depend on their direct and indirect acquaintances. The results follow the clear intuition that having more employed direct acquaintances improves i's prospects, as does decreasing the competition for information from those acquaintances. The other indirect relationships in the network and status of other agents does not enter

[9]The first order stochastic dominance for O_{it} and W_{it} is strict even if $w_i > 0$, provided $w_i < \overline{w}_i$.

the calculation. However, once we take a longer time perspective, the evolution of employment and wages 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 direct and indirect acquaintances.

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.

We first present an example which makes it clear why a full analysis of the dynamics of wages and employment is more subtle than the results in Lemmas 1 and 2.

Example 2 *Negative Conditional Correlations*

Consider a homogeneous job network with three agents, $N = \{1, 2, 3\}$, with $g_{21} > 0$ and $g_{23} > 0$. 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, as are the wages W_{1t} and W_{3t}. 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, 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.

A nice way to sort out the short run competition from the longer run benefits of indirect acquaintances is to allow the periods to become shorter. As the periods shorten, the competitive effects become outweighed by the longer run benefits. We make this clear below. A natural way to analyze shortened periods is simply by dividing p and b by some T.[10]

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.

Before discussing the patterns and dynamics of employment and wages, we define some useful tools.

4.1 Dominance Relations

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 standard generalizations of first order stochastic dominance relationships that applies to random vectors.

Domination

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

μ *dominates* ν if

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

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

for every non-decreasing function $f : \mathbb{R}^n \to \mathbb{R}$.[11] 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.

Self-Domination

In many cases when we are comparing distributions of wages or employment, we will be looking at the distribution as it varies conditional on different histories. So, ν is simply μ conditional on different information. We name this *self-domination*.

μ *self-dominates* if

$$E_\mu [f \,|g \geq d] \geq E_\mu [f]$$

for every non-decreasing $f : \mathbb{R}^n \to \mathbb{R}$, non-decreasing $g : \mathbb{R}^n \to \mathbb{R}$ and scalar $d \in \mathbb{R}$.[12]

Self-domination 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.

Note that if X is a random vector described by a measure μ, then self-domination of μ implies that X_i and X_j are non-negatively correlated for any i and j. Essentially, self-domination is a way of saying that all dimensions of X are non-negatively interrelated. If X 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' employment status. The definition captures more general interactions between many agents, and says that good news in the sense of higher values of $X_i, i \in \{i_1, \ldots, i_\ell\}$ 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 Self-Domination

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

We offer the definition when the support of the distribution μ is finite.

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

A nondecreasing function $f : \mathbb{R}^n \to \mathbb{R}$ 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.

Given a nondecreasing function $g : \mathbb{R}^n \to \mathbb{R}$, we say that $g \geq d$ *reveals information about* $\pi \in \Pi$ (relative to μ) if there exist x and $\widetilde{x}_\pi$ such that $g(x) \geq d > g(x_{-\pi}, \widetilde{x}_\pi)$ and x and $x_{-\pi}, \widetilde{x}_\pi$ are in the support of μ.

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

$$E_\mu [f \,|g \geq d] > E_\mu [f]$$

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

[12] $E_\mu [f \,|g \geq d] = \dfrac{\int_{x\in\mathbb{R}^n : g(x)\geq d} f(x)d\mu(x)}{\int_{x\in\mathbb{R}^n : g(x)\geq d} d\mu(x)}$.

whenever f is sensitive to π and $g \geq d$ reveals information about π.

Strong self-domination captures the idea that better information about any of the dimensions in π leads to unabashedly higher expectations regarding all of the dimensions in π. One implication of this is that X_i and X_j are positively correlated for any i and j in π.

4.2 Employment Dynamics and Patterns

We are now ready to state the central theorems concerning the dynamics of employment and wages. We begin with employment.

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 3 *Consider any economy (N, p, b). There exists T' such that μ^T strongly self-dominates relative to $\Pi(p)$ for any $T \geq T'$, where μ^T is the (unique) steady state distribution on employment associated with the T-period subdivision of (N, p, b).*[13]

Corollary 4 *Consider any economy (N, p, b). There exists T' such that for any $T \geq T'$, and i and j, S_i and S_j are non-negatively correlated under the steady state distribution μ^T. If i and j are path connected, then the correlation is positive.*

Theorem 3 establishes the positive interrelationships between the employment of any collections of path connected agents under the steady state distribution. Despite the short run conditional negative correlations between competitors for jobs and information, in the longer run (with smaller subdivisions) any interconnected agents' employment is positively correlated. This is consistent with the sort of clustering observed by Topa (2001).

The proof of Theorem 3 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 comes to overwhelm 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 by Freidlin and Wentzel (1984) (as adapted to finite processes by Young (1993)) to verify that one can simply keep track of these 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 the example 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.

Example 3 *Correlation and Network Structure.*

[13] 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 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.

Consider a simple homogeneous network setting (Example 1) with $n = 4$ agents. Let $a_i(w) = b_i = \frac{1}{2}$ for all i and w, so the probability of any agent hearing (directly) about a job or losing a job in a given period is equal to one half.

If there is no network relationship at all, $Prob(S_i = 0) = 2/3$ under the steady state distribution.

If we move to a single link $g_{12} = g_{21} = 1$, then $Prob(S_i = 0) = .629$ and $Corr(S_i, S_j) = .037$, where $i, j \in \{1, 2\}$. So the probability of unemployment falls and the two linked agents employment statuses are correlated.

If we move to a "triangle," $g_{12} = g_{23} = g_{31} = 1$ (with reciprocal relationships where $g_{ki} = g_{ik}$), then $Prob(S_i = 0) = .621$ and $Corr(S_i, S_j) = .023$ for each of the linked agents. The probability of unemployment falls further, and the correlation between any two employed agents also falls.

If we move to a "square," $g_{12} = g_{23} = g_{34} = g_{41} = 1$ (again with reciprocal relationships), then $Prob(S_i = 0) = .617$ and $Corr(S_1, S_2) = .022$ while $Corr(S_1, S_3) = .010$. The probability of unemployment falls further, and the correlation between any two employed agents also falls.[14] Here the correlation between directly connected agents is higher than for indirectly connected agents.

Example 4 *Bridges and Asymmetries*

[Add other example from seminar slides with 14 agents to illustrate asymmetries and role of bridges.]

While Theorem 3 provides results on the steady state distribution, we can deduce similar statements about the relationships between employment at different times.

Theorem 5 *Consider any economy (N, p, b). There exists T' such for any $T \geq T'$, any nondecreasing functions f and g, and any times t and t'*

$$E^T \left[f(S_t) \left| g(S_{t'}) \geq d \right. \right] \geq E^T \left[f(S_t) \right],$$

with strict inequality whenever there is some $\pi \in \Pi(p)$ such that f is sensitive to π and $g \geq d$ reveals information about π, if we start under the steady state distribution μ^T where E^T is the expectation associated with the T-period subdivision of (N, p, b).

Although Theorem 5 is similar to Theorem 3 in its structure, it provides different implications. Theorem 3 addresses the steady state distribution, or the expected long run behavior of the system. Theorem 5 addresses any arbitrary dates in the system.[15]

It is important in Theorem 5 that we start from the steady state distribution. For instance, if we start from a given state, such as that in Example 2, we could end up with a negative correlation.

[14] While the changes across networks are marked, the magnitude diminishes quickly in the size of the network in this example. This is partly due to the choice of a=b=1/2, which makes calculations easy. For smaller a and b, employment changes are rarer and then the network status becomes more important as the passing of information becomes a more central way of obtaining jobs, while with high a and b, direct news plays a larger role.

[15] Theorem 3 almost seems to be a corollary of Theorem 5, since as we let t and t' become large, the distributions of S_t and $S_{t'}$ approach the steady state distribution. However, we cannot deduce Theorem 3 from Theorem 5 since it is not ruled out that the positive correlation vanishes in the limit under Theorem 5, while we know that this is not the case from Theorem 3.

4.3 Patterns and Dynamics of Wages

One might imagine that we can directly state a parallel result to Theorem 3 for wages instead of employment status. However, when trying to prove similar statements about wages, we run into a minor complication.

To see the difficulty, consider a situation where agents are more likely to pass job information on to acquaintances with lower wages than to acquaintances 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 acquaintances would be more likely to pass information to them, and they might have a positive probability of obtaining several offers at once.

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. That last case is ruled out by the following assumption.

(3) p is *weakly positive* if for any w and i such that $p_i(w) > 0$ it follows that $p_i(w_{-i}, w_i') > 0$ for any $w_i' < \overline{w}_i$.

Weak positivity requires that if there is some chance that i obtains an offer at one wage level, then that probability remains positive as long as i is not at the highest wage level (holding others wages fixed). This assumption is easily satisfied by homogeneous job networks, and is quite natural more generally.[16]

Theorem 6 *Consider a network economy (N, p, b) with a weakly positive p. There exists a large enough T' such that μ^T strongly self-dominates for any $T \geq T'$, where μ^T is the (unique) steady state distribution on wages associated with the T-period subdivision of (N, p, b).*

Corollary 7 *Consider a network economy (N, p, b) with a weakly positive p. There exists T' such that for any $T \geq T'$, and i and j, W_i and W_j are non-negatively correlated under the steady state distribution μ^T. If i and j are path connected, then the correlation is positive.*

At the steady state, there is a clustering of wage levels and employed workers tend to be acquainted with workers earning similar wages. We also have an analog of Theorem 5, stating that this positive interrelationships between agents' wages hold both under the steady distribution and at any time along the dynamics.

Theorem 8 *Consider any economy (N, p, b) with a weakly positive p. There exists T' such for any $T \geq T'$, any nondecreasing functions f and g, and any times t and t'*

$$E^T\left[f\left(W_t\right)|g\left(S_{t'}\right) \geq d\right] \geq E^T\left[f(W_t)\right],$$

with strict inequality whenever there is some $\pi \in \Pi(p)$ such that f is sensitive to π and $g \geq d$ reveals information about π, if we start under the steady state distribution μ^T where E^T is the expectation associated with the T-period subdivision of (N, p, b).

[16] Note also that any p is arbitrarily close to one satisfying this condition.

5 Duration Dependence and Persistence in Unemployment

We now turn to the theorem outlining the duration dependence that we discussed in the introduction.

Say that i's job contacts are non-degenerate, if there exists $j \in N_i(p)$ with $j \neq i$.

Theorem 9 *Consider an economy (N, p, b). If i's job contacts are non-degenerate, then there exists T' such that starting from the steady state distribution at time 0,*

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

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

The idea behind Theorem 9 is as follows. Longer past unemployment histories lead to worse inferences regarding 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 acquaintances (direct and indirect) are themselves also unemployed. This makes it more likely that i's acquaintances 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 worsens one's social environment, which in turn harms individual future employment prospects.

Example 5 *Duration Dependence.*

Consider a simple homogeneous network setting (Example 1) with $n = 2$ agents. Let $a_i(w) = b_i = \frac{1}{2}$ for all i and w, so the probability of any agent hearing directly about a job or losing a job in a given period is equal to one half.

Consider the simple network: $g_{12} = g_{21} = 1$.

Here, starting under the steady state distribution, for any t: $E(S_{it}) = .371$, $E(S_{it} = 1|S_{it-1} = 0) = .294$, and $E(S_{it}|S_{it-1} = 0, S_{it-2} = 0) = .286$. The same claims hold for large t starting from any initial distribution. As we increase the length of unemployment, the probability of employment continues to drop.

[[add in other examples.]]

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 closed-knit of acquaintances breeds future employment for these acquainted 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 individual dynamics generates persistence for aggregate employment dynamics. This means that individual employment does not follow a Markov process, but exhibits the duration dependence documented in Theorem 9. This also means that the process governing aggregate employment exhibits special features. The higher the overall employment rate, the faster rate at which 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 lower the chance that 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.[17]

[17] We have not explicitly modeled equilibrium wages and the job arrival process. Incorporating these effects might mitigate some of the effects our model exhibits. However, taking the arrival process as exogenous helps us show how the network effects pushes the process to have certain characteristics.

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.

6 Dropping Out, Inequality in Wages and Employment

We now turn to the discussion of agents' decisions of whether or not to stay in the labor force. This effectively endogenizes the network or function p.

Staying in the labor market requires payment of a expected present value of costs $c_i \geq 0$. These include costs of education, skills maintenance, opportunity costs, etc. Let $x_i \in \{0, 1\}$ denote i's decision of whether to stay in the labor market. Each agent discounts future wages at a rate $0 < \delta_i < 1$. Effectively, 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.

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.[18] Therefore, when an agent drops out, it is as if the agent disappeared from the economy.

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

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

and $x_i = 0$ implies the reverse inequality.

In our model, 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. This results in the "drop-out" game being supermodular (see Topkis (1979)) which leads to the following lemma.

Lemma 10 *Consider any economy (N, p, b) with a weakly positive p, 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 $x^*(w)$ such that $x^*(w) \geq x$ for any other equilibrium x.*

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

Theorem 11 *Consider an economy (N, p, b) with a weakly positive p and where at least two agents are path connected, and for each i a costs $c_i \in I\!R_+$ and discount rate $0 < \delta_i < 1$. Consider two starting wages states, $w' \geq w$ with $w_i \neq w'_i$. There exists T' such that for any T-period subdivision ($T \geq T'$), the set of dropouts following w' is a subset of that under w under the maximal equilibria; and for some specifications of the costs and discount rates the inclusion is strict. Moreover, if $x^*(w)_i = x^*(w')_i = 1$, then the distributions of W_{it} and S_{it} for any t under the maximal equilibrium following w' first order stochastic dominate those under*

[18] 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.

the maximal equilibrium following w, with strict dominance for large enough t if $x^*(w)_j \neq x^*(w')_j$ for any j who is path connected to i. In fact for any increasing $f : I\!R^n_+ \to I\!R$ and any t

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

with strict inequality for some specifications of c and δ.

Theorem 11 shows how persistent inequality can arise between two otherwise similar groups. If two different social groups (even with identical network relationships) differ in their starting wage state, the resulting drop-out rates will differ. If the starting state is higher for one group, then that group will have fewer drop-outs (a subset) than the other group. Because dropping out hurts the prospects for the group further, differences in drop out rates end up generating persistent inequality 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 volunteer drop-outs from the labor force affects negatively the standard measures of black economic progress (e.g. Chandra (2000), Heckman et al. (2000)).

[[add in examples.]]

7 Concluding Discussion

While this model is too stylized to provide specific policy implications, we mention some lessons that still can be learned about policy in the presence of network effects. One 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.

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 acquainted with these individuals. Moreover, concentrated improvements lead to a greater improvement of the status acquaintances then 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) acquaintance 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 acquaintances who are in the market. This may induce them to enter. This can then have a contagion effect, carrying over to agents acquainted 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.

Let us also mention some possible empirical tests of the model. To the extent that direct data on network relationships is available, one can directly test the model. However, even without such detailed information, there are 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 and duration dependence 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 acquaintances changes.

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.[19]

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 acquaintances. 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.

[19] See 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. See Dutta and Jackson (2001) for a brief overview and references.

References

[1] Akerlof, G.A. (1997) "Social Distance and Social Decisions," *Econometrica*, 65, pp 1005-1028.

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

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

[4] Blanchard, O.J. and P. Diamond (1994) "Ranking, Unemployment Duration and Wages", *Review of Economic Studies* 61, 417-434.

[5] Boorman, S.A. (1975) "A Combinatorial Optimization Model for Transmission of Job Information Through Contact Networks," *Bell Journal of Economics* 6, 216-249.

[6] Calvó-Armengol, A. (2001) "Job Contact Networks," mimeo, www.eco.uc3m.es/∼ancalvo/jobnet.pdf

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

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

[9] Dutta, B., and M.O. Jackson (2001) "On the Formation of Networks and Groups," Introductory chapter in *Models of the Formation of Networks and Groups,* edited by B. Dutta and M.O. Jackson, forthcoming from Springer-Verlag: Heidelberg. http://www.hss.caltech.edu/∼jacksonm/Jackson.html

[10] Flinn, C. and J. Heckman (1982) "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, 18, 115-168.

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

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

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

[14] Heckman, 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.

[15] Heckman, Lyons and Todd (2000) "Understanding Black-White Wage Differentials, 1960-1990," *American Economic Review*, Papers and Proceedings, 344-349.

[16] Holland, P.W. and S. Leinhardt (1977) "A Dynamic Model for Social Networks," *Journal of Mathematical Sociology* 5, 5-20.

[17] Lin, N. and M. Dumin (1986) "Access to Occupations through Social Ties," *Social Networks*, 8, 365-385.

[18] 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.

[19] Montgomery, J. (1994) "Weak Ties, Employment, and Inequality: An Equilibrium Analysis," *American Journal of Sociology*, 99, 1212-1236

[20] Montgomery, J. (1992) "Job Search and Network Composition: Implications of the Strenght-of-Weak-Ties Hypothesis," *American Sociological Review*, 57, 586-596

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

[22] 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.

[23] Pissarides, C.A. (1992) "Loss of Skill During Unemployment and the Persistence of Employment Shocks," *Quarterly Journal of Economics* 107, 1371-1397.

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

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

[26] Schweitzer, S. and R. Smith (1974) "The Persistence of the Discourage Worker Effect," *Industrial and Labor Relations Review*, 27, 249-260

[27] Topa (2001) "Social Interactions, Local Spillovers and Unemployment" *Review of Economic Studies*, 68, 261-295.

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

[29] Vishwanath, T. (1989) "Job Search, Stigma Effect, and Escape Rate from Unemployment," *Journal of Labor Economics* 7, 487-502.

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

Appendix

(still partly incomplete.)

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 $1-b$ and 0 with probability b, and similarly $S_{it} = 1$ when $O_{it} > 0$ with probability $1-b$, 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

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

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. ∎

Next we prove Theorem 6. Theorem 3 can then be proven as a special case noting that any nondecreasing function f and g on employment can be written as a corresponding nondecreasing function of wages (which maintains the sensitivity and information revealing characteristics).[20]

The following lemmas are useful in the proof of Theorem 6.

Lemma 12 *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 for each $w \in W$ there exists $\phi_w : W \to [-1, 1]$ such that*

$$\mu(w') = \nu(w') + \sum_w \phi_w(w'),$$

where each ϕ_w satisfies

$$\begin{cases} \phi_w(w') \geq 0 \text{ if } w' = w \\ \phi_w(w') \leq 0 \text{ if } w' \leq w,\ w' \neq w \text{ and} \\ \phi_w(w') = 0 \text{ otherwise.} \end{cases}$$

and

$$\phi_w(w) + \sum_{w' \neq w} \phi_w(w') = 0.$$

Strict domination holds if $\phi_w(w) > 0$ for some 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 12: This can be established 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.

Lemma 13 *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 $E \in \mathcal{E}$.

Proof of Lemma 13: 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

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

[20] Theorem 3 does not need the extra assumption of weak positivity. So Lemma 14 can be proven without this assumption when dealing with employment states rather than wage states. Thus, Theorem 3 is not quite a corollary, but can be proven by following the same steps and noting that this assumption is not necessary to reach the same conclusions for employment states.

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. ∎

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

Let $P^T_{wE} = \sum_{w' \in E} P^T_{ww'}$.

Lemma 14 *Consider an economy (N, p, b) with a weakly positive p. 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 $T > T'$ and $E \in \mathcal{E}$*

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

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

Proof of Lemma 14: Let us say that two states w' and w'' are adjacent if $w'_j \neq w''_j$ for only one j and then w'_j and w''_j take on adjacent values in the range of j's wage function.

We show this for w and w' adjacent, as the statement then follows from a chain of comparisons across such w' and w. Let ℓ 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

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

Note that $p_{\ell j}(w') \geq p_{\ell j}(w)$ for all $j \neq \ell$. Also since $w'_k = w_k$ for all $k \neq \ell$ it follows 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.

However, given that $w'_\ell > w_\ell$, 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, $Prob^T_{w'}(O_{-\ell,t} \geq a) \geq Prob^T_w(O_{-\ell,t} \geq a+1)$. For large enough T, given the weak positivity of p, this holds.

To see the strict domination, consider $E = \{w | w_\ell \geq w'_\ell\}$, and note that then, since there is a 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 15 *Consider an economy (N, p, b) with a weakly positive p and two measures μ and ν on W. There exists T' such that for all $T > T'$, 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 15:

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

By Lemma 12

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

Reordering the summations, and noting the properties of ϕ, this becomes

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

Lemma 14 implies that for large enough T, $P^T_{w''E} \geq P^T_{wE}$ whenever $w'' \geq w$. Thus since $\phi_{w''}(w'') \geq 0$ and $\phi_{w''}(w'') + \sum_{w \leq w'', w \neq w''} \phi_{w''}(w) = 0$, the result follows.

Suppose that μ strictly dominates ν. It follows from Lemma 12 that there exists some w'' such that $\phi_{w''}(w'') > 0$. Let $w \leq w''$, $w \neq w''$. By Lemma 14, there exists some $E \in \mathcal{E}$ such that $P^T_{w''E} > P^T_{wE}$. With such E, $[\mu P^T](E) > [\nu P^T](E)$, implying by Lemma 13 that $P^T \mu$ strictly dominates $P^T \nu$. ∎

Proof of Theorem 6: 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 j *independently* hears about a new job (and at most one) with probability $\sum_i \frac{p_{ij}(w)}{T}$, 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 below).

The following claims establish the theorem.

Claim **1** $\overline{\mu}^* = \mu^*$.

Claim **2** $\overline{\mu}^*$ *strongly self-dominates.*

Claim **3** *There exists T' such that μ^T strongly self-dominates for all $T \geq T'$.*

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

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

$$p_z = \sum_{\tau \in \mathcal{T}_z} \left[\times_{z',z'' \in \tau} P_{z'z''} \right]. \tag{3}$$

Lemma 16 Freidlin and Wentzel (1984)[21]: *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 (3) above.

Proof of Claim 1: 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,[22] and other transition probabilities go to 0 at a rate of at least $1/T^2$, it follows from Lemma 16 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 $w'_i < w''_i$. It then follows that

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

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

Proof of Claim 2:

We use the following observation, for which we omit the proof.

Claim 4 *If the W_i's are independent under a measure μ then for any $E \in \mathcal{E}$ and $E' \in \mathcal{E}$*

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

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

[22] This assumes that $P^T_{w',w''} \neq 0$. To handle other cases (where some adjacent transition probabilities are 0), we can argue as follows. For any non-adjacent transition which has a non-zero probability, we can break it into some sequence of adjacent transitions which (by weak positivity) must be positive. The sum of those adjacent transition probabilities will go to zero at the rate of $1/T$, while the non-adjacent transition probability goes at a rate of at least $1/T^2$. We can thus replace any non-adjacent transition via a sequence of adjacent transitions and reduce the resistance of the tree.

Starting from an independent initial distribution over wage states, we show that for $E_t \in \mathcal{E}$ and $E'_{t'} \in \mathcal{E}$,

$$\overline{Prob}^T (E'_{t'}|E_t) \geq \overline{Prob}^T (E'_{t'}) .$$

The proof is by induction. The case $t = 0$ follows from Claim 4. Let $A \in \mathcal{E}$ be an event independent from time t (for instance, A is an event from $t-1$ or before, or A is a $t = 0$ event). We show that if for all $E_t \in \mathcal{E}$

$$\overline{Prob}^T (E_t|A) \geq \overline{Prob}^T (E_t)$$

then,

$$\overline{Prob}^T (E_{t+1}|A) \geq \overline{Prob}^T (E_{t+1})$$

for any $E_{t+1} \in \mathcal{E}$.

We have,

$$\overline{Prob}^T (E_{t+1}|A) - \overline{Prob}^T (E_{t+1}) = \sum_{w_t} \overline{Prob}^T (E_{t+1}|w_t) \left[\overline{Prob}^T (w_t|A) - \overline{Prob}^T (w_t) \right] .$$

By Lemma 12 and the definition of domination, it follows that there exists a ϕ_{w_t} for each w_t such that

$$\overline{Prob}^T (E_{t+1}|A) - \overline{Prob}^T (E_{t+1}) = \sum_{w'_t} \overline{Prob}^T (E_{t+1}|w_t) \left[\sum_{w_t} \phi_{w_t}(w'_t) \right] .$$

Therefore,

$$\overline{Prob}^T (E_{t+1}|A) - \overline{Prob}^T (E_{t+1}) = \sum_{w_t} \left[\sum_{w'_t} \phi_{w_t}(w'_t) \overline{Prob}^T (E_{t+1}|w'_t) \right] .$$

We rewrite this as

$$\overline{Prob}^T (E_{t+1}|A) - \overline{Prob}^T (E_{t+1}) \tag{4}$$
$$= \sum_{w_t} \left[\phi_{w_t}(w_t) \overline{Prob}^T (E_{t+1}|w_t) + \sum_{w'_t \leq w_t, w'_t \neq w_t} \phi_{w_t}(w'_t) \overline{Prob}^T (E_{t+1}|w'_t) \right] .$$

Lemma 14 implies that

$$\overline{Prob}^T (E_{t+1}|w_t) \geq \overline{Prob}^T (E_{t+1}|w'_t)$$

for each $w'_{t \geq w_t}, w'_t \neq w_t$. Also, we know from the definition of domination that $\phi_{w_t}(w_t) + \sum_{w'_t \neq w_t, w'_t \leq w_t} \phi_{w_t}(w'_t) = 0$. Thus, equation (4) implies that $\overline{Prob}^T (E_{t+1}|A) \geq \overline{Prob}^T (E_{t+1})$. Therefore, by induction using Lemma 15 (using $\overline{Prob}^T$ in place of P^T) it follows that for every $E \in \mathcal{E}$ and for any $t' > t$,

$$\overline{Prob}^T (E_{t'}|A) \geq \overline{Prob}^T (E_{t'}) .$$

Moreover, if we have strict domination at time t then we also have strict domination at any $t' > t$.

We now deal with the case $t' = t$. We want to show that for any $E_t, E'_t \in \mathcal{E}$, and for any t

$$\overline{Prob}^T (E'_t|E_t) \geq \overline{Prob}^T (E'_t) .$$

This is equivalent to

$$\sum_{w_{t-1}} \overline{Prob}^T (E'_t|w_{t-1}, E_t) \overline{Prob}^T (w_{t-1}|E_t) \geq \sum_{w_{t-1}} \overline{Prob}^T (E'_t|w_{t-1}) \overline{Prob}^T (w_{t-1}) .$$

Claim 4 implies that

$$\overline{Prob}^T \left(E_t'|w_{t-1}, E_t\right) \geq \overline{Prob}^T \left(E_t'|w_{t-1}\right),$$

for all w_{t-1}. We just need to show that $\overline{Prob}^T \left(w_{t-1}|E_t\right) \geq \overline{Prob}^T \left(w_{t-1}\right)$, for all w_{t-1}, which follows from the previous induction argument. ▮

Proof of Claim 3: Claims 1 and 2 imply that μ^* strongly self-dominates. Next note that $E_{\mu^*}(f|g \geq k) = E_{\mu^*}(f)$ implies that there does not exist i and j who are path connected, for which f is sensitive to i and $g \geq k$ reveals information about j. This implies that none of the agents for whom g reveals any information are in any way connected to any of the agents for whom f is sensitive. This implies that $E_{\mu^T}(f|g \geq k) = E_{\mu^T}(f)$. For any other f and g, $E_{\mu^*}(f|g \geq k) > E_{\mu^*}(f)$, and so for large enough T, $E_{\mu^T}(f|g \geq k) > E_{\mu^T}(f)$. As given the finite W, we need only check this for a finite set of functions f and g (see the proof of Lemma 13), the result then follows. ▮

That completes the proof of Theorem 6. ▮

Proof of Theorem 9: We need to show that for large enough T

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

Let E^t refer to $S_t \in E$, and E^t_{i0} be the event that $S_{it} = 0$.

First, we claim that $\text{Prob}^T(\cdot|E^{t-1}_{i0})$ as viewed as a measure on time t strongly self-dominates. To see this, not that from Lemma 15 we need only show that this is true when $\text{Prob}^T(\cdot|E^{t-1}_{i0})$ viewed as a measure on time $t-1$. The conclusion that $\text{Prob}^T(\cdot|E^{t-1}_{i0})$ strongly self-dominates viewed as a measure on time $t-1$ follows from an extension of Claim 2.

This implies that

$$\text{Prob}^T(E^t_{i0}|E^{t-1}_{i0}) \geq \text{Prob}^T(E^t_{i0}|E^t, E^{t-1}_{i0}),$$

with strict inequality for any $E^t \in \mathcal{E}$ which is sensitive to $N^n_i(p)$.[23] By repeated use of Bayes' rule, this implies that

$$\text{Prob}^T(E^t|E^{t-1}_{i0}) \geq \text{Prob}^T(E^t|E^t_{i0}, E^{t-1}_{i0}),$$

with strict inequality for any $E^t \in \mathcal{E}$ which is sensitive to $N^n_i(p)$. Then again applying Lemma 15 we can deduce the same strict inequalities at time $t+1$.[24] ▮

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 15 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). ▮

[23] We defined sensitivity for functions, but as E^t can be represented by an indicator function, this directly translates.

[24] While Lemma 15 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.

Proof of Theorem 11: Let $w \geq w'$ and $x \in \{0,1\}^n$. We first show that

$$E\left[f(W_t)\,|W_0 = w', x\right] \geq E\left[f(W_t)\,|W_0 = w, x\right].$$

We deduce from Lemma 14 that for a fine enough $T-$period subdivision,

$$E\left[f(W_t)\,|W_0 = w', x\right] \geq E\left[f(W_t)\,|W_0 = w, x\right]$$

for $t = 1$ and for every non-decreasing f. Lemma 15 and a simple induction argument then establish the inequality for all $t \geq 1$. The inequality is strict whenever f is increasing and $w' > w$.

Let now $x, x' \in \{0,1\}^n$, with $x' \geq x$. Following the same lines than the proof of Lemma 14 we can show that, for a fine enough T-period subdivision,

$$E\left[f(W_t)\,|W_0 = w, x'\right] \geq E\left[f(W_t)\,|W_0 = w, x\right]$$

for $t = 1$ and for every non-decreasing f. The same induction argument than before extends the result to all $t \geq 1$. Again, f increasing and $x' > x$ imply a strict inequality.

Let $w' \geq w$. We now show that $x_i^*(w') = 0$ implies $x_i^*(w) = 0$. Suppose not. Let i such that $x_i^*(w') = 0$ and $x_i^*(w) = 1$. We distinguish two cases.

Suppose first that $x_{-i}^*(w') \geq x_{-i}^*(w)$. From $x_i^*(w') = 0$, we deduce that

$$E\left[\sum_t \delta_i^t W_{it}\,|W_0 = w', x_{-i}^*(w')\right] < c_i.$$

As $x_{-i}^*(w') \geq x_{-i}^*(w)$, it is also true that

$$E\left[\sum_t \delta_i^t W_{it}\,|W_0 = w', x_{-i}^*(w)\right] < c_i.$$

Finally, $w' \geq w$ imply that

$$E\left[\sum_t \delta_i^t W_{it}\,|W_0 = w, x_{-i}^*(w)\right] < c_i,$$

in contradiction with $x_i^*(w) = 1$.

Suppose now that $x_{-i}^*(w) > x_{-i}^*(w')$. Then, $x^*(w) > x^*(w')$. We construct an equilibrium y for the wage w' the following way. For all i such that $x_i^*(w) = 1$, let $y_i(w') = 1$. By definition, $x_i^*(w) = 1$ is equivalent to

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

implying that

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

Note that this inequality also implies that

$$E\left[\sum_t \delta_i^t W_{it}\,|W_0 = w', y_{-i}\right] \geq c_i,$$

for all $y \geq x^* (w)$. Now, for all i such that $x_i^* (w) = 0$ do the following. If

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

let $y_i = 1$, otherwise, let $y_i = 0$. Iterate the process. That is, for all i such that $y_i = 0$, let $y_i = 1$ whenever

$$E\left[\sum_t \delta_i^t W_{it} \left| W_0 = w', y_{-i} \right.\right] \geq c_i,$$

and keep $y_i = 0$, otherwise. After a finite number of iterations, this iterative process stops. By construction, y is an equilibrium for w', and $y \geq x^* (w) > x^* (w')$, in contradiction with $x^* (w')$ being a maximal equilibrium. ∎