FEDERAL RESERVE BANK of ST. LOUIS

RESEARCH DIVISION

Working Paper Series

Network Search: Climbing the Job Ladder Faster

Marcelo Arbex,
Dennis O'Dea
and
David Wiczer

Working Paper 2016-009B
https://doi.org/10.20955/wp.2016.009

May 2016

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

Network Search: Climbing the Job Ladder Faster

Marcelo Arbex* **Dennis O'Dea†** **David Wiczer‡**

April 17, 2018

Abstract

We introduce an irregular network structure into a model of frictional, on-the-job search in which workers find jobs through their network connections or directly from firms. We show that jobs found through network search have wages that stochastically dominate those found through direct contact. In irregular networks, heterogeneity in the worker's position within the network leads to heterogeneity in wage and employment dynamics: better-connected workers climb the job ladder faster. Despite this rich heterogeneity from the network structure, the mean-field approach allows the problem of our workers to be formulated tractably and recursively. We then calibrate a quantitative version of our mechanism, showing it is consistent with several empirical findings regarding networks and labor markets: jobs found through networks have higher wages and last longer. Finally, we present new evidence consistent with our model that job-to-job switches at higher rungs of the ladder are more likely to use networks.

Keywords: Labor Markets; Social networks; Job search; Unemployment; Wages dispersion.

JEL Classification: D83; D85; E24; J31; J64.

*Department of Economics, University of Windsor. arbex@uwindsor.ca.; †Department of Economics, University of Washington. odea@uw.edu; ‡ Department of Economics, Stony Brook University. david.wiczer@stonybrook.edu.
We have benefited comments and suggestions from Manolis Galenianos, George-Levi Gayle, Aaron Hedlund, Henry Hyatt, Leo Kaas, Rasmus Lentz, Paulina Restrepo, Giorgio Topa, Weilong Zhang and participants at various seminars and conferences. Any errors are our own.

1 Introduction

How do workers find jobs and does it matter? In this paper, we study an equilibrium search model in which some jobs are found through network connections and workers differ in their position within the network. In our framework, jobs may be found by direct contact with a firm or through a searcher's network, but the latter method draw fundamentally different and better offers. There are two reasons. First, because we assume network connections pass along wage offers like their own, offers passed through network search come from workers who have already selected wages and climbed the ladder. The second reason comes endogenously from the interaction between network structure and on-the-job search: Job referrals come disproportionately from workers who are more central to the network and these better-connected workers have more access to better jobs and thus even higher wages. The heterogeneity presented in our model, workers' network position, implies both differences in workers' finding rate and differences in the distribution from which they sample wage offers. Finding rates differ because better-connected workers have more connections from whom to draw a referral. Draw distributions differ because better-connected workers samples more frequently from the network distribution. To understand this mechanism, we show analytically how network search affects equilibrium on-the-job search, then illustrate its features in a calibrated version and discuss how the model relates to common empirical findings associated with job referrals and network search.

The model is able to tractably incorporate two methods of search, direct contact and network referral, along with heterogeneity associated with the latter. This tractability comes from our ability to summarize a worker's network position into a scalar state variable, her number of ties. To do so, we employ the mean field approach, which amounts to a set of assumptions over the network and information structure to omit local correlation and neighborhood effects.[1] Rather, agents will use the global network structure to take expectations over their peers' types. This simplifies the state while still preserving the network's influence in an equilibrium we term a "*Sufficient Recursive Equilibrium*" because it uses a sufficient, scalar statistic to summarize the worker's network position.

Our principal theoretical result shows that the equilibrium wage offer distribution via social networks first-order stochastically dominates the wage offer distribution via direct search. Underlying this result, we showed that the reservation wage is decreasing and employment rate increasing in the number of connections. Therefore, well-connected individuals are more likely to be employed and the distribution of referrals disproportionately weights these network-central peers. The distribution of offers hence comes from a weighted average of wages of employed workers, one that

[1]The mean field approach is a technique to analyze the long run or average behavior of a complex system and it was first developed for use in statistical physics to analyze the Ising model of interacting charged particles (see Vega-Redondo (2007) for a summary). By replacing the effect of nearby particles with the "average" effect of the magnetic field, one can solve for the equilibrium field, without need to consider the potentially intractable interaction of local effects.

dominates the overall wage distribution of employed workers which dominates the distribution of offers. Introducing network search also affects the equilibrium wage and offer distributions by reducing firms' profits. This occurs because the lowest wage firm shrinks as its workers, who are more likely well connected, are more quickly poached away. With falling profits fall at the lowest wage level, free entry implies they also fall throughout.

Because our model includes both network search and direct contact search, it is a useful framework for understanding the differences between a job found via network search and direct. Our model is consistent with many empirical findings about these differences.[2] In particular, jobs found through a worker's network have (i) higher wages and (ii) longer employment duration and (iii) workers experience shorter unemployment spells. These are all predictions of our model. Network search brings higher wages both because network offers come from a better distribution and also because those who find a job through their network tend to be better connected, a composition effect. Employment duration is longer because these higher wages are higher on the job ladder. Again because of composition, unemployment spells are shorter; those who find through networks are better connected and have faster finding rates. While many other studies attribute features these to information frictions,[3] we show that network heterogeneity and a job ladder are sufficient. We also contribute to the empirical literature by introducing findings from the Survey of Consumer Expectations (SCE). A prediction central to our model is that network search becomes increasingly important at higher rungs of the job ladder. In fact, with the SCE, we find direct evidence of this: higher paid workers are more likely to find their next job via a referral, and we use the rate this likelihood increases as quantitative discipline in our calibration.

There is an extensive literature about frictional labor markets and wage dispersion. The basis for our model, Burdett and Mortensen (1998), develops a wage-posting model in which firms offer high wages to attract workers from other firms and to reduce worker turnover. Their main result is an equilibrium with wage dispersion even if workers and firms are identical. Wage dispersion is produced by the possibility of on-the-job search. Wage offers generally exceed the reservation wages and firms can attract more workers from other employers. Under the assumption of equal profits for firms that offer different wages, it is possible to obtain an equilibrium where the wage distribution is not degenerate on the common reservation wage. While Burdett and Mortensen (1998) introduce a version with reservation heterogeneity, this has quite different features than our network heterogeneity and so for simplicity, our benchmark comparison will be the pure-dispersion version.

To match empirical wage distributions better than Burdett and Mortensen (1998), a list of

[2]A large empirical literature has shown that labor market networks play an important role in matching workers to employers. Even though estimates of the percentage of jobs found through social contacts vary across location and profession, they consistently range between 25% and 80% of jobs in a given profession (see, for instance, Holzer (1988), Ioannides and Datcher Loury (2004), Bayer et al. (2008), Hellerstein et al. (2011)).

[3]Many principally empirical papers ascribe their findings, network search leads to longer-lasting matches at higher wages, to network connections ameliorating information frictions. See Granovetter (1995); Kugler (2003); Datcher (1983); Marmaros and Sacerdote (2002); Simon and Warner (1992); Dustmann et al. (2015)

papers including this one introduces additional heterogeneity. Postel-Vinay and Robin (2002) distinguish three sources of wage dispersion: productive heterogeneity of workers, productive heterogeneity of firms and search frictions. All three sources are quantitatively important, and worker and employer heterogeneity are both required to explain the observed shape of earnings distributions. This is echoed in other search models capturing wage dispersion with worker and potentially firm-level heterogeneity; equilibrium on-the-job search models can match the wage distribution and provide a sizable role for person effects, individual-level heterogeneity, that interact with search frictions (see, e.g. Bontemps et al. (2000), Postel-Vinay and Turon (2010), or Bagger et al. (2014)). Contributing to this literature, our model introduces a dimension that increases wage dispersion and was previously treated as unobserved heterogeneity.

Theoretical models of of labor market networks generally assume that there are information frictions that hinder the job search behavior of unemployed workers and/or firms, and that information flows through networks. In models such as Calvo-Armengol and Jackson (2007) and Ioannides and Soetevent (2006) job searchers can learn about job vacancies either directly from employers or indirectly via employed individuals among their network contacts, but unemployed workers do not have full information about job vacancies. In equilibrium, better-connected job searchers are more likely to find employment and to have higher wages. In Montgomery (1991), the information imperfection is on the employer side. Employers do not have full information about the quality of job applicants and firms learn about a potential worker's ability if the firm employs individuals from the potential worker's network. In equilibrium, individuals are more likely to receive and accept wage offers from businesses that employ others in their network.

Several very relevant papers consider the contribution of network search to wage inequality. Mortensen and Vishwanath (1994) considers a Burdett and Mortensen (1998) model but in which sometimes a worker draws from the earnings rather than the offer distribution. Our concept of a referral is also a draw from the distribution of other workers' earnings, however by adding an irregular network we show how network draws are actually from a different distribution: Earnings are weighted differently because one's peers are not representative of the whole population. In other words, our structure shows how network search adds another layer of inequality on top of that in Mortensen and Vishwanath (1994). Galenianos (2014) presents a search model with job referrals through a network and rich predictions about business cycle fluctuations and cross-sector differences in measured vacancy yield. In it, networks also contribute to inequality and worker dynamics depend on the network. But unlike our work, he considers a homogeneous network, so differences in network location cannot contribute to wage dispersion. Fontaine (2008) presents a matching model à la Pissarides (2000) where wage dispersion arises endogenously as the consequence of the joined dynamics of networks, firms' strategies and wage bargaining.

A great deal of empirical literature has studied the role networks play in the labor market, and we see their findings largely consistent with our model's implications. Here we briefly summarize some of the empirical support for network search leading to (i) higher wages, (ii) longer employment duration, and (iii) shorter unemployment spells. Using a direct measure of network quality based on the employment status of close friends, Cappellari and Tatsiramos (2015) provide robust evidence that a higher number of employed contacts increases wages for high-skilled workers forming networks with non-familial contacts. Bentolila et al. (2010) analyze surveys from both the US and Europe including information on job finding through contacts and find that network contacts reduce unemployment duration by 1-3 months on average. Hellerstein et al. (2015) present evidence that workers who are more residentially networked to their co-workers at the time of hire have lower rates of turnover. They also provide evidence that labor market network strength is linked to more rapid re-employment and re-employment at neighbors' employers. Defining networks by "co-displaced" workers, Cingano and Rosolia (2012) present evidence for Italy that employment (re-employment) of other co-displaced workers in the network reduces unemployment duration. There are related papers that study the effect of coworker-based networks and employee's referrals on individual labor market outcomes (see, for instance, Brown et al. (2013), Glitz (2013), and Burks et al. (2015))

Our main contribution is to incorporate irregular networks into the labor market search and matching process. In particular, by acknowledging the role of heterogeneity in network connections we provide one reason that workers are not equally likely to locate an opening nor find the same quality of job. We believe that this paper provides a framework to analyze wage dispersion and unemployment as driven by (non-observable) worker heterogeneity in social connections. The paper proceeds as follows. In Section 2 we introduce the main features and assumptions regarding workers' network and job information. Section 3 presents a model of labor markets with two channels of information about job opportunities: direct search and social networks. In Section 4, we discuss the steady-state equilibrium and the associated distributions. In this section, we also compare our equilibrium conditions to those in Burdett and Mortensen (1998). Section 5 presents the properties of the equilibrium and, in particular, we show that the earnings distribution of a peer stochastically dominates the earnings distribution in the economy. Section 7 presents the results from a calibrated version of our model where we explore the quantitative implications of network search for wage and employment dynamics. Section 6 analyzes important empirical results through the prism of our model. Section 8 offers concluding comments.

2 Demography, Network and Information Structure

2.1 Labor market characteristics

There are large fixed numbers of workers and employers participating in the labor market, formally a continuum of each. The measure of workers is normalized to one. At a moment in time, each worker is either unemployed (state $i = 0$) or employed (state $i = 1$). Workers have a wage $w \in \mathbb{W}$, where $\mathbb{W} \subset \mathbb{R}_+$ is the firms posted wage offers set. The unemployed all have the same benefit, b. We also define a worker's history of wage offers through referral $k \in \mathbb{W} \times \mathbb{R}$.

A job offer represents a new position at a firm. Employed workers can hear about new positions either at their own firms or at other firms. In the first case, the employed worker passes the job information to her peers. This new job is identical to the job of the employed worker, so that the worker who hears of it will not have any interest in taking the offer herself. Hence, this is a model of *referral*. All job offers stand for one period only. If a job offer is declined, it expires and the position is lost. Hence, referrals also last only one period, during which they are instantly transmitted to other workers. Because search is random, each of those postings may be filled by a currently employed or an unemployed worker.

An employed worker can only refer direct connections to jobs in her own firm. This relatively simple transmission process precludes a multi-step job transmission and simplifies our analysis. We focus instead on the role of the structure of the social network, rather than the technology of the transmission process. Rather than passing along the referral, employed workers can also move from lower- to higher-paying jobs if they find an offer from a higher-paying job.

2.2 Network Structure and Position

A network is a collection of *nodes* and *links* (sometimes called vertex and edges, respectively), where each worker is a node, and an edge exists between two workers if job information may pass between them. The number of edges a node has is called its *degree*. In general, networks may be very complex, ranging from highly structured to highly random. One way to describe a network is through its *degree distribution*, which gives the proportion of nodes that have each possible degree. Thus, workers are heterogeneous in the number of peers to whom they are connected in a social network. We assume that the network is described by the degree distribution $\Omega(z)$ and we will focus on large, complex networks, so that $z \in [1, \infty)$.[4] As we will show, this is sufficient to summarize her *type*, given a set of assumptions.

The number of links will be centrally important and so to clarify notation, we will refer to the number of links with several names depending on the role of the worker in the network when we refer to her. When we use z to denote her number of peers, the worker is a generic agent. When we denote the type as s it will refer to the number of links belonging to a peer of one of these agents. When we refer to type t, we mean the number of links of an employee. Of course, any agent plays all these roles.

To summarize the position of a worker in this network, her number of peers is, generically not sufficient. In any generic network structure, we also would need to make explicit the connections of these peers. Two workers with z peers may experience the network differently if these peers have more or fewer connections, and so of course, this also applies to the connections of those connections. Hence, in the next section, we discuss a worker's state and how her number of peers summarizes the relevant information to the worker in the network.

[4] This is common to approximate the discrete number of network connections with a continuous variable, so rather than $z \in 1, ..., \infty$ we use this half-closed interval.

2.3 Workers' state

In our environment, the workers' labor-force status, wage and history of wage offers (i, w, k) is not enough to characterize a worker's state. To fully define a worker's state, we will need to include all of the attributes of the peers with whom she may interact. These peers may pass along job openings and may themselves be passed job openings and the wage of these referrals will depend on referrer's wage. Let $\chi \in \mathbb{X}$ summarize the relevant information to the worker in the network. It includes all of the labor market information along paths. This is the worker's collection of peers, their labor market-relevant state, this information for the peers' collection of peers, and so on. Clearly, it is convenient to define χ recursively. Consider an individual whose number of direct connections is $z \in [1, \infty)$. Then χ is a $z \times 4$ object, which contains, for each connection indexed c a quadruple $(i(c), w(c), k(c), \kappa(c)) \in \chi$, i.e., the connection's employment status $i(c) \in \{0, 1\}$, wage $w(c) \in \mathbb{W}$, history $k(c) \in \mathbb{W} \times \mathbb{R}$ and that connection's own position in the network $\kappa(c)$. For that connection c with s connections of her own, $\kappa(c)$ is an $s \times 4$-dimensional object.

2.4 Network and information assumptions

In what follows, we present and discuss network and information assumptions under which it is possible to ensure that z is actually sufficient to summarize a worker's state. Our approach relies on an assumption about the class of network structure and two key assumptions on the information structure.

Network Structure Assumption

We do not assume any particular degree distribution for our results; it may be finite or continuous, and our results apply to arbitrary degree distributions. We require, however, that the network described by the degree distribution $\Omega(z)$, $z \in [1, \infty)$, be *complex*, in the following sense: workers must not be able to infer too much information about the network and their peers from their own position in the network. Assumption A.1 formalizes this.

Assumption (A.1, Network Structure)**.** *The network is connected, infinitely large and generated by a configuration model consistent with a degree distribution* $\Omega(z)$*.*

A well known property of networks generated according to assumption A.1 is that they are locally tree-like. Equivalently, if we define the number of nodes as N, as it grows $(N \to \infty)$ then for any path of finite length, the probability it is a loop approaches zero. With an infinitely large network the probability of a loop is almost surely zero. This is the same as saying that any finite-length path will not loop back to the original node and therefore there are no clusters in this network. The following example illustrates this requirement.

Consider a highly structured network in which all workers are part of dense clusters of 6 workers, connected to each one of 5 peers. Two of these 6 peers have a connection to another dense cluster of 6 agents, and this continues *ad infinitum*. This may be visualized as clusters on a line, as in Figure 1. Such a network has a discrete degree distribution; all workers will have either five or six peers, with the proportion of workers of type 6 being $\frac{2}{6}$. Knowing one's own type, therefore, conveys a great deal of information; a workers of type 5 knows that she is *not* one of the "bridge" workers, and thus knows that each of her 5 peers is connected to one another, and two will be bridges. Thus she may infer that their employment statuses will be highly correlated, and with a long enough memory, she may learn which peers of hers are bridges. A worker of type 6 knows she is a bridge, and thus knows 5 of her peers are in a dense cluster (one of whom is also a bridge and one is in separate cluster, and may eventually learn which).

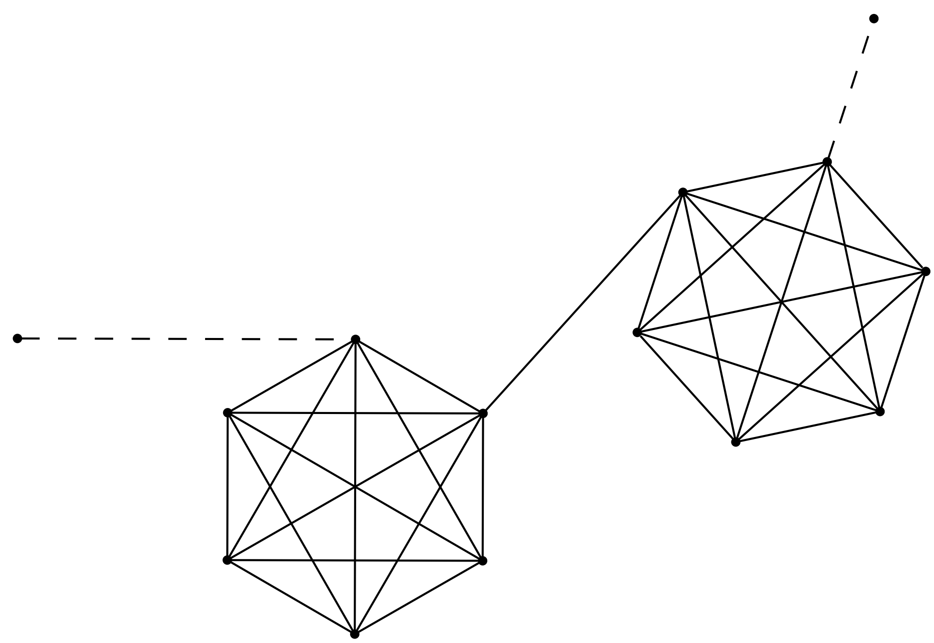

Figure 1: A highly structured network

Contrast this with a different network structure, constructed as follows: starting with a single worker, endow this worker with either 5 or 6 peers, choosing 6 with probability $\frac{2}{6}$. For each of this worker's peers, endow them with either 4 or 5 peers, choosing 5 with probability $\frac{2}{6}$, *ad infinitum*. This network may be visualized as a random *tree*, that continues forever, as in Figure 2. This network shares the same degree distribution as the former network, but workers cannot infer anything from their own type - the types of their peers are independent of each other and uncorrelated. This network is *complex* - the type of network we consider.

To implement assumption A.1, we consider random, complex networks in which agents believe the social network they inhabit was generated by a so-called "configuration" model, a very general process introduced by Bender and Canfield (1978). We follow the formulation of Chung and Lu (2002) to randomly assign links.[5] This procedure can generate large, complex networks with arbitrary degree distributions.

It works as follows: for a fixed number N of nodes, fix a desired degree distribution $\Omega(z)$. Assign

[5]For further details, see Vega-Redondo (2007)

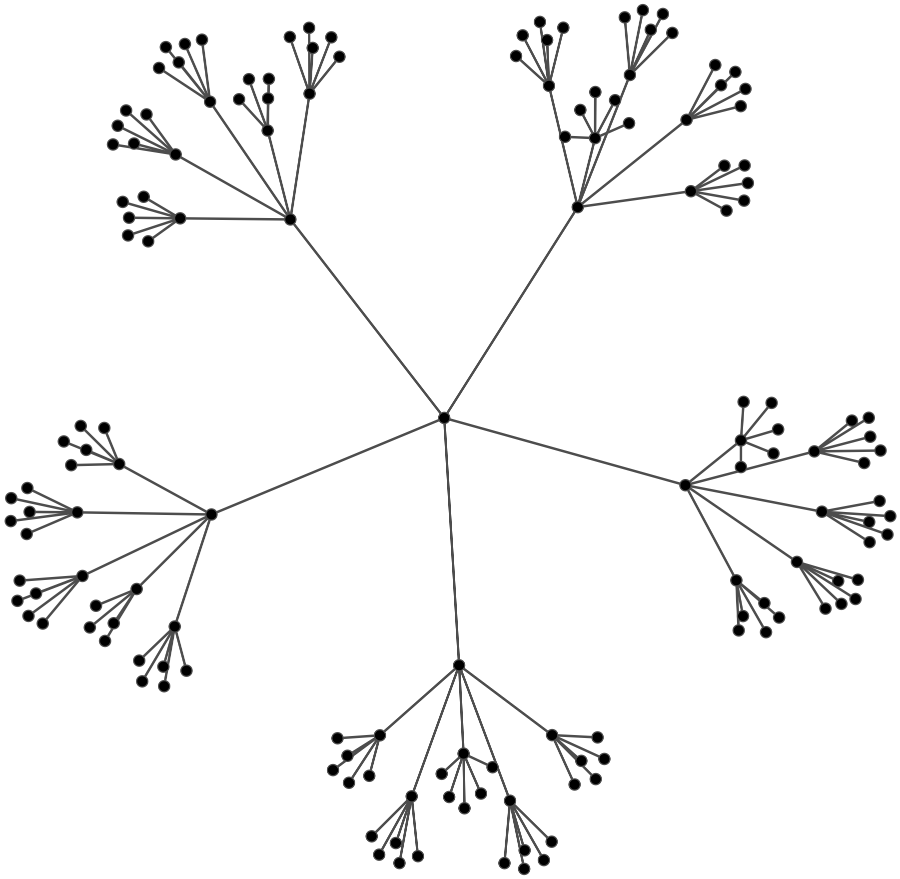

Figure 2: A random tree network

to each node i weight $\Omega(i)$. For each pair of nodes ij, form a link between them at random, with probability proportional to $\Omega(i)\Omega(j)$. This procedure, as N rises, will generate a large, complex network with the desired degree distribution. Networks generated by this process are almost all trees, in that there is no local structure. To put this concretely, the probability a worker's peers are themselves peers vanishes and there are no correlations between node properties. In other words, a node of degree j is not more likely to be connected to nodes of one type than another. Hence, the degree distribution of a neighbor is the same for every neighbor, for every node.

In some sense, our network structure assumption A.1 also limits the information in the economy because with a locally tree-like structure, a worker's own degree does not convey any information about the expected degree of her peers. This actually imposes some numerical restrictions on the degree distribution, in particular, that it's clustering coefficient goes to zero and the number of links is small enough to be in the "critical region," as described in Dorogovtsev et al. (2008). In our implementation, we assume power-law distributions and so a single parameter governs both these features. To ground ideas, networks that satisfy this restriction include those derived from scale free, i.e. power-law, degree distributions when properly parameterized. In our numerical exercises we will use this family also because of its empirical relevance (Vega-Redondo (2007), Newman

(2004)).

Information Structure Assumptions

Next, we rely on two informational assumptions on what the worker knows of the network. To set up notation, let a worker's information set be $\mathcal{I}$, which is equivalent to her state, (i, w, k, χ), in the full information case. Assumption A.2 is that a worker cannot observe their peer's state: neither their position in the network, their employment status, nor their wage.

Assumption (A.2, Limited Observability)**.** *For any worker with information set, $\mathcal{I}$, $\chi \notin \mathcal{I}$ but $||\chi|| \in \mathcal{I}$.*

As we will show, the effect of this assumption is to "anneal" the network: similar to redrawing the network each period, workers will need to take expectations over the state of their connections. When they form these expectations, however, we will limit them only to use the structure of the economy. Our next assumption A.3 limits their memory and ensures that this inference can be based only on global averages.

Assumption (A.3, Limited Memory)**.** *For any worker whose information set is $\mathcal{I}$, $k \notin \mathcal{I}$.*

By assuming workers cannot recall the sequence of job offers from their peers, they can only use the structure of the economy, including the degree distribution and expected distributions of labor market variables to form a belief about the value of a job passed from a peer. The worker's state, and information set, is now (i, w, z).

With these assumptions, we can introduce an important result that makes the rest of the model considerably more tractable. Assumptions A.1-A.3 allow us to use the "degree-based mean field approach," replacing idiosyncratic local effects with "expected" average effects. The important implication is that rather than χ characterizing a worker's position in the network, her number of connections z is sufficient. Workers of type (i, w, z) are treated as an equivalence class.

Proposition 1. *In a network described by A.1 and given information assumptions A.2 and A.3, the degree z is sufficient to characterize an agent's network position, χ.*

The proof, which also uses model features introduced below, is given more completely in Appendix A, showing how the full information problem reduces to use only the worker's state and information set (i, w, z). We sketch its intuition here. Our Assumption A.1 applies the "degree-based mean field approach" (see Pastor-Satorras et al. (2015)) to the network structure. This means that workers with the same number of connections are in equivalent positions in the network graph in expectation. However, for a given realization, two workers with equal numbers of peers will differ both because the number of connections their connections have may differ and because the labor market status of these connections may differ. Assumption A.2 establishes that a worker cannot directly observe these features of her peers and therefore must form expectations when she solves her optimization problem. Imposing that workers take expectations is equivalent to "annealing" the network, as if the connections were redrawn every period. Assumption A.3 prevents a worker from inferring additional details of their position based on the history of referred offers. Now, the expectations formed by any worker with z connections will be the same, which takes workers whose network location is equivalent in expectation to be actually equivalent. Therefore, workers with the same number of connections form an equivalence class because there is no usable information about a worker's network. And, hence, a worker's number of connections z is sufficient to characterize her position in the network.

3 A Model of Search and Networks in Labor Markets

3.1 Network Search Technology

In this section we will present job and offer flows with a search technology in which network and direct search co-exist. Because we use the degree-based mean field approach, we have a relatively parsimonious network structure which allows us to characterize worker types simply by their number of links and therefore to present most of the model objects using this scalar quantity as the dimension on which workers are *ex ante* heterogeneous.

Workers learn about job opportunities when unemployed and when employed. Job information is acquired through either direct search or a worker's social network. Job information arriving directly from employers depend on a worker's current employment state. Let γ^i denote the arrival rate of direct offers while a worker has employment status $i \in \{0, 1\}$. These offers are drawn from a firm offer distribution $F(w)$.

Job offer information can also be acquired through a worker's social network. The peers distribution of types $\Psi(s)$ is distinct from the overall distribution of types $\Omega(s)$ because those with more peers are more likely to connected to the agent whose problem we solve. This is sometimes called the *paradox of friendship*: a very general phenomenon of networks identified by Feld (1991), implying that one's friends are likely to have more friends than oneself. This is because agents with many peers, and a large type s, are disproportionately likely to be one's peers. The distribution over peer types is given by $\Psi(s) = s\Omega(s)/\langle z \rangle$, where $\langle z \rangle = \int_{z=1}^{\infty} z\Omega(z)dz$ is the *average degree* in the network.

We will define the distribution of types conditional on that one sending a referral as $\tilde{\Psi}(s)$, which will also generically differ from $\Psi(s)$ because not all peers are equally likely to have sent a referral. For a given worker with employment status i, the joint probability of a worker of type s transmitting the offer and being a peer is $\gamma^1 n(s)\frac{\nu}{s}\Psi(s)$, while the total probability of such an event integrates over types s. Thus, canceling terms in the numerator and denominator the conditional distribution of types is

$$\tilde{\Psi}(s) = \frac{n(s)\frac{1}{s}\Psi(s)}{\int n(z)\frac{1}{z}\Psi(z)dz} = \frac{n(s)\Omega(s)}{\int n(z)\Omega(z)dz}$$

where the second equality results from substituting in the definition $\Psi(s) = \frac{s\Omega(s)}{\langle z \rangle}$. Notice that this distribution puts more weight on the types who are more likely to be employed than the unconditional distribution $\Omega(z)$. Because better-connected peers receive offers out of unemployment more quickly, this will tend to put more weight on higher z workers.

The probability a worker of type z in state i receives an offer via a peer in her social network is, in discrete time $\left(1 - \left[1 - \nu\gamma^1 \int \frac{n(s)}{s}\Psi(s)ds\right]^z\right)$. Taking the limit as the number of arrivals per period goes to infinity, the continuous time arrival rate is

$$\rho(z) = \left(1 - \exp\left(-z\nu\gamma^1 \int \frac{n(s)}{s}\Psi(s)ds\right)\right) \tag{1}$$

where ν is the probability at which job information is passed along.[6] The probability ν can be interpreted as a "socializing" parameter, representing how strongly tied to one another peers in this network are and, hence, how likely they pass along a potential job.[7] The arrival rate γ^1 appears because jobs to be passed are acquired in the same way as on-the-job search where the peer is searching for her peers rather than herself. Given that ν is unbounded above, this assumption is without loss of generality.

From equation (1) we can see already that heterogeneity in z implies heterogeneity in job finding rates. The density of the network affects the finding rate in two ways. A job offer to be passed arrives at rate $\gamma^1\nu$ but is multiplied by $\frac{1}{s}$ because every one of the peer's connections entails more competition with other workers connected to the same individual. On the other hand, if the network is denser and searching workers have more connections z, they have more chances to get an offer and hence faster finding rates. Note also how the mean field approach appears: we integrate over the distribution of peers' types $\Psi(s)$, rather than knowing the actual number of connections any of a searcher's peers have. Furthermore, we take as separate, the probability an offer arrives and the wage associated with that offer because assumption A.2 prevents worker from observing the origin of the referral (otherwise that would tell the worker her connection's wage).

The distribution of wages acquired through referral will be different from the offer distribution $F(w)$. Rather, it is a function of the equilibrium distribution of earnings by type and the employment rate by type. Let $G(w, z)$ be the earnings distribution among agents with z links, or in other words, the proportion of employed workers of type z earning a wage no greater than w. The wage

[6] Notice we are assuming that jobs are passed to peers at the same rate, regardless of employment status. We could weaken this assumption, and allow $\nu^0 > \nu^1$ but this would complicate some proofs and require constraints on $0 \leq \nu^0 - \nu^1 < const$, constraints we believe are loose in the empirically relevant range. We will discuss the implications of relaxing this assumption in Proposition 4.

[7] This approach follows Calvo-Armengol and Jackson (2004). In words, the probability one of your peers passes you job information depends on the job arrival rate, the socializing parameter ν, the average employment rate n, and the average degree in the network $\langle z \rangle$. This captures both the benefit of there being highly connected peers and the cost of more competition for this information. More-connected peers are likely to be employed and thus pass information but they also have more potential recipients.

distribution in the population is then

$$G(w) \doteq \int_z \Omega(z)G(w,z)dz \tag{2}$$

However, the type distribution, conditional on receiving a wage offer is not the same as the population's distribution. The expected distribution of referral wages is then

$$\tilde{G}(w) = \int_s \tilde{\Psi}(s)G(w,s)ds.$$

Hence, a job offer learned via a worker's network is, therefore, a random draw from the endogenous earnings distribution $\tilde{G}(w)$ of a random peer weighted by the probability that a peer is actually employed.

3.2 Households

We will consider separately workers who are employed and unemployed, so the workers' value function $V(i,w,z)$ are split into $V^0(z)$ and $V^1(z,w)$. The state of the unemployed worker is her number of connections $z \in [1,\infty)$ and the state of an employed worker is $(z,w) \in [1,\infty) \times \mathbb{W}$. Given the distribution of offers obtained through direct contact, F, and the distribution of earnings that may be referred, G, the value function of an unemployed worker of type z is

$$\begin{aligned} rV^0(z) = b + \gamma^0 &\left\{ \int_{R(z)}^{\bar{w}} \left[V^1(z,x) - V^0(z)\right] dF(x) \right\} \\ &+ (1-\gamma^0)\rho(z) \int_{R(z)}^{\bar{w}} \left[\int_s \tilde{\Psi}(s) \frac{\partial G(x,s)}{\partial x} \left(V^1(z,x) - V^0(z)\right) ds \right] dx, \end{aligned} \tag{3}$$

where r is the economy's interest rate. In words, the value of an unemployed worker with z peers is the unemployment benefit, the expected value of hearing about a job directly, plus the expected value of hearing about a job from a peer.

Again taking F and G as given, the value function of an employed worker with z connections and wage w is

$$\begin{aligned} rV^1(z,w) = w + \delta\left[V^0(z) - V^1(z,w)\right] + \gamma^1 &\left\{ \int_w^{\bar{w}} \left[V^1(z,x) - V^1(z,w)\right] dF(x) \right\} \\ &+ (1-\gamma^1)\rho(z) \int_w^{\bar{w}} \left[\int_s \tilde{\Psi}(s) \frac{\partial G(x,s)}{\partial x} \left(V^1(z,x) - V^1(z,w)\right) ds \right] dx. \end{aligned} \tag{4}$$

That is, the expected discounted lifetime income of a worker currently employed, with z peers in her network, is her wage w plus the expected value of become unemployed plus the expected value of hearing of a better job either directly from firms or through her social network. Her job-search policy is simple, so we plugged into the value function in equation (4): she takes any job offering wages greater than her current wage w.

Given an offer w, an unemployed worker will accept it if it is higher than her reservation $R(z)$, which may endogenously depend on peers, z. If unemployed, the worker receives unemployment benefits b. At the reservation wage $R(z)$, we have that $V^1(z, R(z)) = V^0(z)$. Plugging $w = R(z)$ into equation (4) and using equation (3), we obtain

$$\begin{aligned}
R(z) - b &= \left(\gamma^0 - \gamma^1\right) \left\{ \int_{R(z)}^{\bar{w}} \left[V^1(z,x) - V^0(z)\right] dF(x) \right\} \\
&\quad + \left[(1-\gamma^0)\rho(z) - (1-\gamma^1)\rho(z)\right] \left\{ \int_{R(z)}^{\bar{w}} \left[\int_s \tilde{\Psi}(s) \frac{\partial G(x,s)}{\partial x} \left(V^1(z,x) - V^1(w,z)\right) ds \right] dx \right\} \\
&= \left(\gamma^0 - \gamma^1\right) \left\{ \int_{R(z)}^{\bar{w}} V_x^1(z,x) \left[1 - F(x)\right] dx \right\} \\
&\quad - \left[(\gamma^0 - \gamma^1)\rho(z)\right] \left\{ \int_{R(z)}^{\bar{w}} V_x^1(z,x)(1 - \tilde{G}(x)) dx \right\} \qquad (5)
\end{aligned}$$

where the derivative $V_x^1(z, x)$, using equation (4), is given by:

$$V_x^1(z,x) = \left[r + \delta + \gamma^1(1 - F(x)) + (1-\gamma^1)\rho^1(z)(1 - \tilde{G}(x)) \right]^{-1}. \qquad (6)$$

3.3 Firms

On the labor demand side, we will denote the average size of a firm whose employees have t links and get paid a wage w as $l(w, t)$. To construct this, we will look at flows in and out of a firm by type.

Separation Rate : First looking at outflows, a firm offering wage w will lose workers to unemployment at the exogenous separation rate δ. Workers at this firm will be directly contacted by other firms at rate γ^1, and the probability the firm they are contacted by offers a wage above w is $1 - F(w)$. Workers at this firm are contacted by their peers at rate $(1 - \gamma^1)$, and these peers will communicate a job offer at a firm with a wage above w with probability $\rho(z)(1 - \tilde{G}(w))$, which gives the probability an acceptable job offer is passed to that worker. Hence, the separation rate $\beta(w, t)$ is the sum of these terms:

$$\beta(w,t) = \delta + \gamma^1 \left[1 - F(w)\right] + (1-\gamma^1)\rho(t)(1 - \tilde{G}(w)) \qquad (7)$$

and the mass of workers of type t leaving the firm is thus $l(w, t)\beta(w, t)$.

Compared to Burdett and Mortensen (1998), where separations would be given by $\beta(w) = \delta + \gamma^1(1 - F(w))$, there is an additional source of "poaching" of workers that depends on the number of peers a firm's worker have. This increases the outflow of higher type workers relatively more than lower type workers.

Hiring Flows : Consider now recruiting to this firm. First, there is a flow of recruits via direct contact. Unemployed and employed workers are contacted directly by firms at rate γ^0 and γ^1, respectively. These job contacts are spread among the $\frac{1}{M}$ firms, and they will result in jobs for the unemployed (employed) if the offered wage is above their reservation wage (current wage).

In our economy, there are $\Omega(z)(1-n(z))\mathbb{I}_{R(z)\leq w}$ workers of type z that may be recruited from unemployment at wage w where $\mathbb{I}_{R(z)\leq w} = 1$ if $R(z) \leq w$ and zero otherwise. There are also $\Omega(z)n(z)G(w,z)$ employed workers of type z earning less than w who may be recruited by this employer. Direct recruiting is therefore given by

$$\frac{\Omega(z)}{M}\left\{[1-n(z)]\,\gamma^0\mathbb{I}_{R(z)\leq w} + n(z)\gamma^1 G(w,z)\right\} \ . \tag{8}$$

Recruiting also occurs indirectly through the social network of a firm's employees. Consider an employed worker of type t at this firm, which employs $l(w,t)$ such workers. An employed worker may contact all t of her peers about job opportunities at her own firm. The total amount of possible referrals made by these workers is $\ell(w,t)t$, i.e., the more connections employees have, the more referrals they will make.

Consider a given worker who is contacted by her employed peer. With probability $\tilde{\Psi}(z)$ she is of type z. With probability $1-n(z)$ she is unemployed. With probability ν the job offer is transmitted, and it will be accepted if the wage is above her reservation wage $R(z)$. On the other hand, with probability $n(z)$, the worker contacted is employed. The job offer is then transmitted with probability ν, and it will be accepted if the wage is above her current wage, which happens with probability $G(w,z)$. Thus, the indirect recruiting of workers of type z by employees of type t of a firm paying w is given by

$$\gamma^1\ell(w,t)t\tilde{\Psi}(z)\left\{(1-n(z))\mathbb{I}_{R(z)<w}\nu + n(z)G(w,z)\nu\right\} \ .$$

Integrating this over t, we get the total volume of (indirect) recruiting of workers of type z from all the firm's current employees, i.e.,

$$\int_t \gamma^1\ell(w,t)t\tilde{\Psi}(z)\left\{[1-n(z)]\,\nu\mathbb{I}_{R(z)\leq w} + n(z)\nu G(w,z)\right\}dt \ . \tag{9}$$

Note that unlike direct recruiting, referrals are not spread evenly among the M active firms. They are in a sense "targeted" because a firm with better connected workers will receive more referred workers. The term $\ell(w,t)\times t$ reflects how the firms' current employees are not only valuable because of their productivity (which is constant at 1), they are also valuable in how they can recruit others. The labor force is, in this sense, the firms' search capital. From the perspective of the type z worker who will get the referral, this employed worker is her connection. This means that connections are distributed by $\tilde{\Psi}(\cdot)$, so the firms worker passes referrals to more connected workers. Combining Equations (8) and (9), the total hiring $h(w,z)$ of workers of type z by a firm paying w is thus

$$\begin{aligned} h(w,z) \quad &\equiv \quad \frac{\Omega(z)}{M}\left\{[1-n(z)]\,\gamma^0\mathbb{I}_{R(z)\le w}+n(z)\gamma^1 G(w,z)\right\} \\ &\quad + \int_{t=1}^{\infty}\gamma^1\times\ell(w,t)t\tilde{\Psi}(z)\left\{(1-n(z))\mathbb{I}_{R(z)\le w}\nu+n(z)G(w,z)\nu\right\}dt. \end{aligned} \tag{10}$$

Compared to Burdett and Mortensen (1998), where recruiting would be $h(w) = ((1-n)\gamma^0 + n\gamma^1 G(w))/M$, in our model recruiting also occurs indirectly through the social network of a firm's employees—current employees are the firms' search capital. In this way, both hiring and separation flows depend on the economy-wide wage distribution and the type distribution within a firm. The additional terms imply that, with network search, high wage firms will have even faster hiring than low wage firms because of their workers because, recall, well-connected workers climb the ladder faster and are therefore more concentrated at high wage firms. Thus, network search will imply a steeper size distribution and even larger high wage firms compared to Burdett and Mortensen (1998).

Putting these together, the change in labor of type z is

$$\dot{\ell}(w,z) = h(w,z) - \beta(w,z)\ell(w,z)$$

Because there is a range of workers with different numbers of employed peers $z \in [1,\infty)$ who will earn a wage w, the total labor input per firm at this particular wage is given by

$$L(w) = \int_z \ell(w,z)dz. \tag{11}$$

An employer's flow of revenue generated per employed worker is normalized to 1. The flow of profits earned by a firm offering wage w is therefore

$$\pi = (1-w)L(w).$$

4 Steady State Equilibrium

We now discuss the steady-state equilibrium, what we call a *Sufficient Recursive Equilibrium*. To build to this, we solve for the steady state wage, employment and firm-size distributions.

4.1 Employment and Wage Distributions

An equilibrium solution is a household solution, $V^0(z)$, $V^1(z,w)$ and $R(z)$, in which we set $r = 0$. Optimizing firm behavior implies a wage offer distribution $F(w)$, such that every wage in the support of $F(w)$ is associated with the same maximizing profit π. Finally, both sides of the market are consistent with the definitions of $G(w,z)$ and $n(z)$, the distributions of wages and employment which we now define.

The law of motion of the employment rate of a worker with z peers, $n(z)$, is

$$n_{t+1}(z) = (1-\delta)n_t(z) + [1-n_t(z)]\left\{\gamma^0\left[1-F\left(R(z)\right)\right] + (1-\gamma^0)\rho(z)\left(1-\tilde{G}(R(z))\right)\right\}. \tag{12}$$

To explain these terms, workers separate into non-employment at a rate of δ. The rate of meeting from unemployment is $\gamma^0 + (1-\gamma^0)\rho(z)$ and the chance one of these is above the reservation is either $F(R(z))$ or $\tilde{G}(R(z))$, depending on where the offer originates.

In steady state, $n_{t+1}(z) = n_t(z) \equiv n(z)$, using equation (12). Hence, the steady state employment rate of such workers is:

$$n(z) = \frac{\gamma^0\left[1-F\left(R(z)\right)\right] + (1-\gamma^0)\rho(z)\left(1-\tilde{G}(R(z))\right)}{\delta + \gamma^0\left[1-F\left(R(z)\right)\right] + (1-\gamma^0)\rho(z)\left(1-\tilde{G}(R(z))\right)}, \tag{13}$$

and the economy's average employment rate is given by

$$N = \int_z n(z)\Omega(z)dz\ . \tag{14}$$

From the steady-state allocation of matches for each type z, the number of employed workers receiving a wage no greater than w, $G(w,z)n(z)$, can be calculated. Because this is the steady state, its change must be equal to 0, which we write as follows:

$$\begin{aligned} 0 &= [1-n(z)]\left[\gamma^0\left\{F(w) - F(R(z))\right\}\right] \\ &+ [1-n(z)]\left[(1-\gamma^0)\rho(z)\left\{\tilde{G}(w) - \tilde{G}(R(z))\right\}\right] \\ &- n(z)G(w,z)\left\{\delta + \gamma^1\left[1-F(w)\right] + (1-\gamma^1)\rho(z)(1-\tilde{G}(w))\right\}\ . \end{aligned} \tag{15}$$

This holds for any wage w, such that $w \geq R(z)$ for some z. Then we rearrange terms in equation (15) to define $G(w,z)$, obtaining

$$G(w,z) = \frac{[1-n(z)]\left\{\gamma^0\left[F(w) - F(R(z))\right] + (1-\gamma^0)\rho(z)\left\{\tilde{G}(w) - \tilde{G}(R(z))\right\}\right\}}{n(z)\left[\delta + \gamma^1\left[1-F(w)\right] + (1-\gamma^1)\rho(z)(1-\tilde{G}(w)\right]} \tag{16}$$

From equations (13) and (16), the steady-state distribution of wages earned by employed workers of type z is

$$G(w,z) = \frac{\delta\frac{\gamma^0[F(w)-F(R(z))]+(1-\gamma^0)\rho(z)\left[\tilde{G}(w)-\tilde{G}(R(z))\right]}{\gamma^0[1-F(R(z))]+(1-\gamma^0)\rho(z)\left[1-\tilde{G}(R(z))\right]}}{\delta + \gamma^1\left[1-F(w)\right] + (1-\gamma^1)\rho(z)\left[1-\tilde{G}(w)\right]} \tag{17}$$

When compared to the classic Burdett and Mortensen (1998) model, where $G(w) = \frac{\delta(F(w)-F(R))/(1-F(R))}{\delta+\gamma^1(1-F(w))}$, we have an additional force that leads to hiring into higher wages, $\frac{(\tilde{G}(w)-\tilde{G}(R(z))}{(1-\tilde{G}(w))}$. As long as $\tilde{G} < F$, as shown below, $\frac{(\tilde{G}(w)-\tilde{G}(R(z))}{(1-\tilde{G}(w))} < \frac{(F(w)-F(R(z))}{(1-F(w))}$. Thus, even if overall finding rates were the same in the two frameworks, the earnings distribution here mixes with another distribution that first-order stochastically dominates it, and will therefore has more workers at higher wages.

4.2 Steady State Flow of Workers

In the steady state, the flow of workers of type z leaving this firm, equation (7), must equal the flow of workers of type z entering this firm, equation (10), and we must have

$$l(w, z)\beta(w, z) = h(w, z) \tag{18}$$

Notice that the steady-state condition, equation (18), represents the balanced (separation and recruiting) flows of workers, with z peers, for a particular firm offering wage w. If w is below $R(z)$, no workers of this type will be recruited to this firm from unemployment. If there are no workers of type z currently earning less than w, then $G(w, z) = 0$ and no workers of this type will be recruited to this firm from other firms. The steady state level of employment of this type of worker, at this firm, will then be zero—the wage w may not be high enough to attract workers of this type if they have better opportunities. Notice however that this is an equilibrium phenomenon and depends on both $R(w)$ and $G(w, z)$.

The equilibrium measure of workers of type z earning a wage w is either 0, if $w < R(z)$, or

$$l(w, z) = \frac{h(w, z)}{\beta(w, z)} \tag{19}$$

A firm offering wage w will attract workers of many different types. Let $\mathcal{Z}(w)$ be the set of types such that if $z \in \mathcal{Z}$ then $R(z) \leq w$, the workers who will accept a firms wage w. Hence, the firm's total labor force is given by

$$L(w) = \int_{z \in \mathcal{Z}(w)} \frac{h(w, z)}{\beta(w, z)} dz \tag{20}$$

4.3 Equal profits and the domain of wages

As is common in wage posting models with heterogeneous reservation wages, the lowest offered wage is not the lowest reservation wage. A firm offering the lowest wage solves a profit maximization problem trading off profits from a lower wage against the types that might accept, resembling that of Burdett and Mortensen (1998) in their version with heterogeneous reservation wages. It is different however in that our dimension of heterogeneity also affects outflow rates to other firms and the hiring rate from unemployment by this lowest-wage firm. The implication is that profit in the lowest wage firm will not be proportional to the population distribution: the fraction of workers whose reservation wage is greater than R. Then $\underline{w} = \arg\max_w (1-w)\underline{L}(\underline{w})$ where $\underline{L}$ defines the labor force in the lowest-wage firm:

$$\begin{aligned} \underline{L}(\underline{w}) &= \int_{z\in\mathcal{Z}(\underline{w})} \frac{h(\underline{w},z)}{\beta(\underline{w},z)}dz \\ &= \int_{z\in\mathcal{Z}(\underline{w})} \frac{\frac{\Omega(z)}{M}\gamma^0(1-n(z)) + \int_t \gamma^1 \ell(\underline{w},t)tdt\tilde{\Psi}(z)(1-n(z))\nu}{\delta+\gamma^1+(1-\gamma^1)\rho(z)}dz. \end{aligned} \tag{21}$$

$h(\underline{w},z)$ is the flow of recruits to this firm and $\beta(\underline{w},z)$ is a measure of workers of type z leaving the firm (separation rate). We omitted the indicator function $\mathbb{I}_{R(z)\leq \underline{w}}$ because we are only integrating over $z \in \mathcal{Z}$, where $\mathbb{I}_{R(z)\leq \underline{w}} = 1$. Notice this firm does not recruit from other firms, because any employed workers are already earning at least $\underline{w}$. Instead, they hire only from unemployment and among those types whose reservations wage is sufficiently low. They lose to poaching any worker who receives another offer because this offer is higher, almost surely.

In Burdett and Mortensen (1998) with heterogeneous reservation wages, the lowest wage was determined by $\max_w (1-w)\Pr[z < \underline{z} : R(\underline{z}) = w]$ because $L(\underline{w}) \propto \Pr[z < \underline{z} : R(\underline{z}) = w]$, but this is not the case in our model. In that model the labor force depends only on the constant fraction that may recruited directly from unemployment $\int_{z\,:\,R(z)\leq w} \Omega(z)\gamma^0(1-n(z))$, while here there is also network recruiting. Because network recruiting happens at rate $\nu \int \gamma^1 \ell(\underline{w},t)tdt$, even if labor supply of each type t were proportional to the population density of that type, hiring would not be.

At the top of the wage distribution, the separation rate for a firm offering $\bar{w}$ is simply δ, i.e., $\beta(\bar{w},z) = \delta$. Firms paying the highest wage in equilibrium will lose workers only to unemployment, not to other firms. On the hiring side, however, this firm recruits directly or indirectly all types of workers from all other firms as well as from the unemployment pool. Since $F(\bar{w}) = 1$ and $G(\bar{w},z) = 1$, the total hiring $h(\bar{w},z)$ of workers of type z by a firm paying $\bar{w}$ is

$$h(\bar{w},z) = \frac{\Omega(z)}{M}\left\{[1-n(z)]\,\gamma^0 + n(z)\gamma^1\right\} + \nu\int_t \gamma^1 \ell(\bar{w},t)t\tilde{\Psi}(z)\nu dt \tag{22}$$

Note that the employment rate of potential recruits $n(z)$ only controls the contact rate, i.e., the rate at which employed and unemployed workers are contacted about job opportunities.

The equilibrium measure of workers of type z earning wage $\bar{w}$ is $l(\bar{w},z) = h(\bar{w},z)/\delta$ and the total labor force of a firm offering the highest wage $\bar{w}$ in equilibrium is given by

$$L(\bar{w}) = \int_z \frac{h(\bar{w},z)}{\delta}dz\ , \tag{23}$$

which does not depend on $\bar{w}$, except that it must be the highest offered wage.

With free entry in equilibrium, every wage offered, regardless of worker type, must yield the same steady-state profit. In other words, the steady-state profit flow of a firm offering the lowest wage in the market ($\underline{w}$) must be equal to the profit of an employer paying the highest wage in equilibrium ($\bar{w}$):

$$\pi^* = [1 - \underline{w}]\, L(\underline{w}) = (1 - \bar{w})L(\bar{w}) \tag{24}$$

4.4 Equilibrium Definition

We are now ready to present our definition of the *Sufficient Recursive Equilibrium*. We call this equilibrium concept "sufficient" because it is defined using z, the number of peers to whom a worker is connected in a social network, as a sufficient statistic for the worker's position in the network. The more general equilibrium, in which the worker's network state is given by a much more complicated object (χ) is described in Appendix A. There, we prove that our framework exploits the mean-field approach to reduce the meaningful heterogeneity to z.

Definition 1. The state of the unemployed worker is her number of connections $z \in [1, \infty)$ and the state of an employed worker is $(z, w) \in [1, \infty) \times \mathbb{W}$, where $\mathbb{W} \subset \mathbb{R}_+$ is the firms posted wage offers set. A *Sufficient Recursive Equilibrium* is the unemployed searcher's value function $V^0 : [1, \infty) \to \mathbb{R}$, the worker's value function $V^1 : [1, \infty) \times \mathbb{W} \to \mathbb{R}$ and the reservation wage $R : [1, \infty) \to \mathbb{R}$, a level of profit at each posted wage $\pi : \mathbb{W} \to \mathbb{R}_+$, a distribution of posted wage offers $F : \mathbb{W} \to [0, 1]$, the employment distribution $n : [1, \infty) \to [0, 1]$ and workers' earnings distribution $G : \mathbb{W} \times [1, \infty) \to [0, 1]$ and, such that

1. the value functions $V^0(z)$, $V^1(z, w)$ and the reservation wage $R(z)$ satisfy equations (3), (4) and (5), respectively,

2. every wage drawn from the firm's offer distribution $F(w)$ is associated with the same steady state profit, i.e., $\pi = (1 - w)L(w) = \bar{\pi}$, $\forall\ w \in \mathbb{W}$. The bounds of $\mathbb{W}$ are determined by equations (21), (23) combined with equation (24), and

3. the employment distribution $n(z)$ and the workers' earnings distribution $G(w, z)$ are consistent with equations (13) and (16), respectively.

The principal task of solving for an equilibrium is now to find the distribution, F, for posted wages that is consistent with equal profits. This distribution must be consistent with separation flows in Equation 7 and the equal profit condition.

4.5 Equilibrium wage dispersion

We have thus far focused on equilibria with wage dispersion. For most of the parameter space, this is the unique equilibrium because there is always a profitable deviation from a degenerate wage distribution. Our argument takes two parts, first imagining all of the mass at 1 and then taking the mass is at $\bar{w} < 1$. Essentially, this is the same argument for the existence of a unique equilibrium with wage dispersion in a model similar to Burdett and Mortensen (1998).

If $\bar{w} = 1$ is the maximum and minimum wage, then $\pi = 0$. Consider a deviation with $w = 1 - \epsilon$. The labor at this firm is given by Equation (21) which is strictly positive for any value of $\rho(z)$

so long as $\exists z: \ R(z) \leq 1-\epsilon$. Then $\pi(1-\epsilon) = (1-1+\epsilon)L(1-\epsilon) > 0$ and the deviation was profitable, hence the distribution cannot be degenerate at the competitive wage.

Suppose instead that $\bar{w} < 1$ is the only wage in equilibrium. Consider then a deviation of $w = \bar{w} + \epsilon$, which will be able to poach any worker from firms offering $\bar{w}$. This means that steady state $L(\bar{w}+\epsilon)$ will be given by Equation (23), while $L(\bar{w})$ is as in Equation (21). Because $1-w$ is continuous, $\exists \epsilon$ such that $(1-\bar{w})L(\bar{w}) < (1-\bar{w}-\epsilon)L(\bar{w}+\epsilon)$: there exists a sufficiently small deviation that will still attract a much larger work force.

As in Mortensen and Vishwanath (1994), if a large enough fraction of the job finding comes through referrals then there may be a degenerate wage distribution with mass at the competitive wage. There exists a large enough ν, such that $\rho(z)$ is uniformly high enough so that π both falls to 0. In our environment, the conditions for this situation are not easily summarized, but can be numerically verified.

5 Properties of the Equilibrium

Our principal result relates the wage distributions passed via network search to those obtained via direct search. It states that the distribution of offers from a peer stochastically dominates the offer distribution obtained directly from firms. To establish this stochastic dominance result, we first show that the reservation wage of workers is decreasing in their type. It follows from this result that the employment rate of a worker is increasing in her type, i.e., her number of peers and, because the expected distribution of referral wages weights these worker more, this distribution dominates the wage distribution in the population. We are then ready to show that the referral wages distribution also first-order stochastically dominates the offer distribution - our main result, Proposition 3. All proofs can be found in the proofs appendix, Appendix A.

Lemma 1. *The reservation wage $R(z)$ is decreasing in z: $\frac{\partial R(z)}{\partial z} \leq 0$.*

In principle, this need not be so; the faster job finding rate of higher type workers might lead them to increase their selectivity, and so have a higher unemployment rate than workers who are less selective due to their lower network job finding. In equilibrium, however, their job finding rate via networks is fast enough that they are more willing to accept lower wage jobs and forgo the faster job finding that comes from off-the-job rather than on-the-job search than workers of lower type. From this follows the fact that the employment rate $n(z)$ is increasing in z.

Lemma 2. *The employment rate $n(z)$ is increasing in z: $\frac{\partial n(z)}{\partial z} \geq 0$.*

Workers of higher type are in equilibrium more likely to be employed, due to both fast job finding and their lower reservation wage. In equilibrium, high type workers climb the job ladder so quickly, and the firms offering such low wages are so small, that workers of higher type still have higher earnings than those of lower type. Thus, the employment rate is increasing in z, and because $\tilde{G}$ places greater weight on types with a higher employment rate, this leads to the stochastic dominance of G by $\tilde{G}$.

Proposition 2. *$\tilde{G}(w)$ first-order stochastically dominates $G(w)$.*

This result is driven by the network structure of the labor market. Workers of higher type are more likely to be your peer than workers of lower type. They are also more likely to earn higher wages of workers than lower type. Thus, the wages earned by peers are even higher than the wages earned by a typical worker. Thus the "friendship paradox" - that the average peer has more friends than the average worker - is an additional source of the benefit of network search over direct search, beyond simply sampling $G(w)$ rather than $F(w)$. We thus have that the distribution of jobs found via networks first-order stochastically dominates the distribution of direct job offers from firms.

Proposition 3. *$\tilde{G}(w)$ first-order stochastically dominates $F(w)$.*

Because referred jobs are from workers already on the wage ladder, G will naturally first-order stochastically dominate F, a feature that can be found in the original Burdett-Mortensen model. Proposition 2 goes further to say that not only are the wages of employed workers better than wages found via direct search, but jobs referred by peers are better than jobs sampled randomly from employed workers.

This additional offer distribution $\tilde{G}$, qualitatively different from the direct contact offer distribution $F(w)$, is unique to our model with $\nu > 0$. In Proposition 4, we look at comparative statics in ν to understand how this change affects other equilibrium objects in the model. We evaluate these starting at $\nu = 0$ because that is how we nest the standard Burdett-Mortensen model in ours. This makes $\nu = 0$ useful because we understand analytically the solution to the model and allows for compares between our model to the well-known benchmark.

Proposition 4. *Evaluated at $\nu = 0$, for sufficiently high γ^1,*

1. *the interval of wages gets wider, $\frac{\partial \underline{w}}{\partial \nu} \leq 0$, $\frac{\partial \bar{w}}{\partial \nu} \geq 0$*
2. *firms at the lowest wages get smaller and at the highest wages they get larger, $\frac{dL(\underline{w})}{d\nu} \leq 0$ and $\frac{dL(\bar{w})}{d\nu} \geq 0$*
3. *profit declines $\frac{d\pi}{d\nu} \leq 0$.*

Proposition 4 shows that introducing network search gives workers additional bargaining power because large high-wage firms increase disproportionately. Increasing ν, the probability at which job information is passed along, makes the lowest wage firms smaller because they lose workers to higher wage firms that are better at recruiting through networks. [8] This lowest wage firm also

[8]We define a sufficient but not necessary condition on γ^1 relative to α, γ^0 that ensures this.

choose a lower optimal wage, because the reservation wage policy is decreasing in ν, just as it was in z. We use an envelop-like condition to show that though $\underline{w}$ declines, $L(\underline{w})$ shrinks enough that equilibrium profit is lower for all firms. The top end moves consistently: if low wage firms are smaller because they are losing workers to poaching, high wage firms get larger by recruiting though their network. Because these firms are getting larger, but profit is getting smaller they also must be offering higher wages. Proposition 4 summarizes these arguments made through a series of Lemmas in Appendix A.6.

6 Numerical Results

We now describe the results from a calibrated version of our model. With this numerical solution, we will explore the quantitative implications of network search for wage and employment dynamics, focusing in particular on how heterogeneity in the number of contacts affects the model's endogenous outcomes such as the wage offer distributions and wage growth paths. We first describe a unique dataset on network search from the New York Federal Reserve's Survey of Consumer Expectations (SCE). Then, we describe the features of the model itself and numerically compare the solution to a similar model without network search, i.e., a Burdett-Mortensen benchmark. Next, we compare the results of a model where workers all have the same job finding rates to one in which there is finding rate heterogeneity. And finally, we use the model to explore and discuss some empirical results about network search and job finding through labor market networks.

6.1 Survey of Consumer Expectations and the network finding rate

In this survey, the Federal Reserve Bank of New York asks households a battery of questions about their expectations regarding their own future and the economy. There is also a once a year labor supplement with detailed questions about job search and, particularly useful for our purposes, the methods of search they are using and have used (see Armantier et al. (2017) for a general description of the data and Faberman et al. (2017) for the labor market supplement).

A useful feature of the SCE is that it includes the prior search method of currently employed workers: Employees are asked how they found their current job. We categorize each worker by whether they found their job through a job referral, and whether that came from a professional connection.[9] In the sample 23% of jobs were found through networks and the rest through a variety of methods. Beyond details of the search process, we also have data on the earnings and characteristics of this job, earnings in the job they left and what caused their separation. More details of our sample selection and summary statistics are available in Appendix B.

For our calibration section, we will use a key feature of this data: The probability a worker finds her next job through network referral is increasing along the wage ladder. This is to say,

[9]Specifically, the SCE categorizes referrals as those from business associates or from "a friend or relative." We take the former as a referral. This is more consistent with our model which is really about professional networks and more in line with prior work emphasizing the importance of weak-ties.

conditional on changing jobs, it is increasingly likely that job was found via referral. Between the 5th and 95th percentile of the earnings distribution, a worker at the bottom of the earnings distribution, workers are about 6 percentage points less likely than average to find their next job via their network. Whereas, at the top of the earnings distribution, they are 5 percentage points more likely. Both of those figures control for observable characteristics of the recruiting firm. We return to this result in Figure 3 with its model counterpart.

To quantify the elasticity of network finding to prior wage, in Table I we estimate the probability of finding a job via network search as follows

$$\Pr[\text{Network find}] = f(\beta_1 salary_{ik} + \beta_2 small_j + \beta_3 J2J_{ij} + \beta_4 PT_{ij}).$$

Here, k indexes the prior employer, while j indexes the employer to which the worker goes. The effects are quite strong: If a worker changes jobs, her likelihood of utilizing a network would increase by about 6 pp if her salary were one standard deviation, 0.9 log points, above the mean.

We will discuss in greater detail why this is the case, but heuristically, there are two reasons for this increase. First, at higher wages workers are more likely to be well connected and hence sample more frequently from the network distribution. Second, because the network-search offer distribution first-order stochastically dominates direct-search offer distribution it becomes an increasingly share of acceptable poaching offers at higher wages.

	(1)	(2)	(3)	(4)
Log prior employer salary	0.0643***	0.0584***	0.0570***	0.0645***
	(3.08)	(2.85)	(3.28)	(3.07)
<500 employees		0.0123		
		(0.29)		
Voluntary Job-to-Job			0.0803**	
			(2.39)	
Part-time				0.0068
				(0.08)
Firm Size Dummies	X		X	
Industry Dummies	X	X	X	
Observations	544	544	544	544

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table I: Estimated marginal effects, replacing prior earnings in equation 34

6.2 Calibration

To first normalize some parameter values, we choose $b = 0$ and the mass of firms $M = 1$. The rest of the parameters are jointly calibrated using flow rates from the CPS and data on network finding from the SCE. Particularly, we need values for $\nu, \gamma^0, \gamma^1, \delta$ and the distributional parameter α. These will be exactly identified by five moments of the data.[10]

Three of our moments are standard labor market search moments: the unemployment rate, the rate at which unemployed workers find jobs and the rate at which employed workers find jobs with new employers. We introduce parameters ν, α, however, that are specific to our environment with network search. Roughly speaking, ν governs how common are referrals, and we target 24%, the fraction of hires by a network contact in our SCE data. In terms of our model elements this moment is defined as

$$0.24 = \frac{\int_z \int_w l(w,z) \int_{t=1}^{\infty} \gamma^1 \ell(w,t) t \Psi(z) \left\{(1-n(z))\mathbb{I}_{R(z)<w}\nu + n(z)G(w,z)\nu\right\} dt dw dz}{\int_z \int_w h(w,z) dw dz}.$$

Though our theory accommodate an arbitrary degree distribution, we use a power-law distribution for $\Omega(z)$:

$$\Omega(z) = (\alpha - 1)z^{-\alpha},$$

where the power-law exponent α determines how heavy the tail of the distribution is, i.e., how common are nodes with much higher than the mean number of peers.[11][12] For $\alpha < 2$, the mean degree in the network diverges (the tail is "too heavy"), while for $\alpha > 3.48$, the network is so sparsely connected that the "giant component" will not exist; in this case, workers will not all be path-connected to each other in the network.[13]

To identify this parameter, we use the increase over the job ladder in the probability of finding a job via network. The intuition being that inequality in the network structure changes the difference in composition between high wage and low wage matches while also changing the degree to which $\tilde{G}$ dominates F, both of which contribute to the increase in network finding rate at higher wage jobs.[14]

Table II presents our baseline calibration parameters, above the line those are those jointly calibrated and below are the ones set without solving the model. We verify via 448 random restarts of the calibration routine that these are the globally optimal, unique parameters.

In Figure 3, we show how those who move to a new job are, at higher wages, increasingly likely to do so via their network. This figure plots $\frac{(1-\gamma^1)\rho(z)(1-\tilde{G}(w))}{\gamma^1(1-F(w))+(1-\gamma^1)\rho(z)(1-\tilde{G}(w))}$, which is increasing

[10]As mentioned in Section 3.1, we constrain the arrival rate on and off the job to be the same for network and direct contact offers, i.e., $\nu^1 = \nu^0$. We could relax this and introduce more moments, which would be computationally trivial but significantly complicate the proofs. For consistency, there will be only ν.

[11]A power law with exponent α is thus equivalent to a Pareto distribution with shape $\alpha - 1$ and scale 1.,

[12]Power-law networks have been identified and studied in a variety of contexts, including empirical social networks. For example, the patterns on friendship on Facebook were found to be well-described by a power-law in Ugander et al. (2011).

[13]For a discussion of so-called "phase transitions" in complex networks, see Vega-Redondo (2007), Hofstad (2016).

[14]A reasonable alternative strategy would be to use hazard profiles, similar to Menzio et al. (2016). However, there are many reasons for heterogeneity in finding rates and we worry about overstating the degree of network heterogeneity by matching our distribution to all of this dispersion.

Parameter	Value	Moment	Model	Data
γ^0	0.24	Average finding rate out of unemployment	0.24	0.25
γ^1	0.10	Average finding rate from employment	0.02	0.02
ν	0.04	Fraction of hires through the network	0.13	0.23
α	2.34	Slope of network finding rate	0.26	0.25
δ	0.13	Average separation rate into unemployment		

Table II: Baseline calibration parameters and targets

in w for any given z. As we have discussed, this result is a direct consequence of $\tilde{G}$ first-order dominating F, so that as we get higher in the distribution, there is more mass left in $1-\tilde{G}(w)$, the probability of a dominating network offer, than in $1-F$. Both $\tilde{G}$ and F distributions are induced in equilibrium, but $\tilde{G}$ is most sensitive to α - it becomes steeper as the network becomes more connected, which is controlled by α.

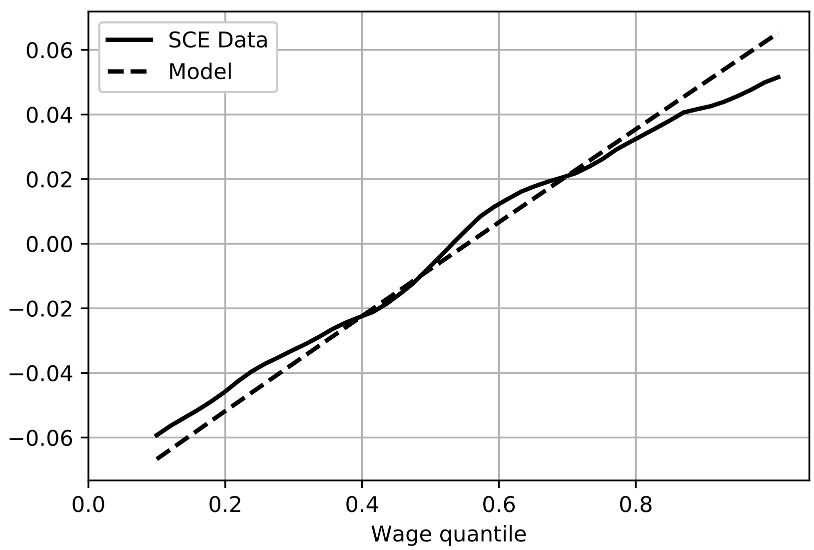

Figure 3: The increase in the probability that a worker's next job will be found through network-search, model and data.

6.3 Characteristics of the solution

The most important features of our model are the two ways labor market networks affects the job search of an individual worker. First, better connected workers find jobs more quickly. The probability a worker of type z, receives an offer via a peer in her social network is an increasing, concave function defined in equation (1). Given our calibration of ν and α the arrival rate via network for the average worker is 8.3% and one with double the average number of links gets offers at 15.9%.

The second effect is that workers draw from a different offer distribution when they search for a job. Offers through the network have higher wages, a result shown in our theoretical result, Proposition 3. The distribution of offers found through a worker's network, $\tilde{G}(w)$, first-order stochastic dominates the distribution, $F(w)$, from which the direct searcher finds her job. This was because the earnings distribution G dominates the offer distribution F, and also because the distribution of wages among connections $\tilde{G}$ dominates G. Because workers with more connections draw more often from this dominant distribution, their unconditional distribution dominates those with fewer connections. This is visualized in Figure 4.

Figure 4 plots distributions for both methods, direct contact and network search, against the distribution of wages. Notice the direct contact line is above the 45° because $F > G$. The network search line is slightly below 45° though it is visually difficult to see. The black lines in Figure 4 are weighted averages of two distributions F and $\tilde{G}$. The better-connected worker draws more often from $\tilde{G}$, so her average puts more weight on the dominating distribution $\tilde{G}$ compared to F. Not only does a better-connected worker draw more offers, she also draws from a better offer distribution.

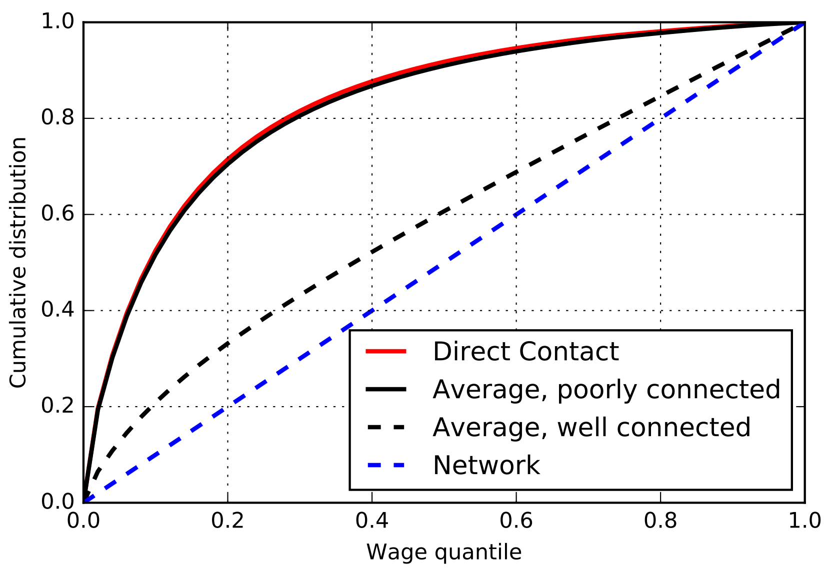

Figure 4: Average distribution of wage offers by contact method conditional on number of peers

To step inside the mechanism through which $\tilde{G}$ dominates F, recall that $\tilde{G}$ weights conditional earnings distributions by the relative employment share, rather than their population weights. Better-connected workers, who have higher wages and higher employment therefore dominate the distribution. Figure 5 illustrates this point by plotting the employment rate and average wage by a worker's number of peers z. Both are increasing, so $\tilde{G}$ puts increasing weight on populations whose wages are higher than the average.

A key result of our network structure—some workers have better access to $\tilde{G}$, which dominates F—is that some workers climb the ladder more quickly. To demonstrate this idea, we will consider

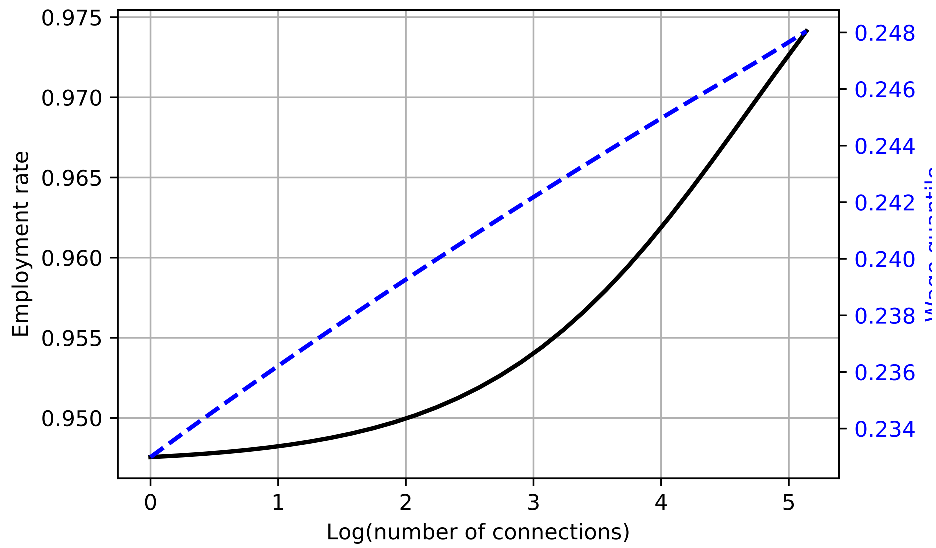

Figure 5: Employment rate (Left Axis) and average wage within overall distribution (Right Axis) both increase with network connectedness

the half-life before an employed worker's wages are expected to converge on the upper-most wage, $\bar{w}$. As in any model with a job ladder, as a worker's wage increases, she is decreasingly likely to get an offer that dominates and so the speed at which she climbs the ladder is also decreasing. Hence, a good measure of the instantaneous rate a worker is climbing the job ladder is the half-life, which measures half of the expected time to converge to wage $\bar{w}$ with wages growing at their instantaneous rate given wage w and connections z. To be specific, the half-life is $\frac{\log 2}{\lambda(w,z)}$, where the instantaneous convergence rate, $\lambda(w, z)$, is

$$\lambda(w, z) = -\log\left(\frac{\bar{w} - E[w'(w, z)]}{\bar{w} - w}\right)$$

Figure 6 illustrates how quickly the wage will approach the maximum wage $\bar{w}$. Each line conditions on the starting wage and traces how the expected time to reach the maximum wage decreases as the number of connections increases. This figure highlights that the effect of a better network is stronger at high wages, i.e. the slope is steeper. At these high rungs, the chance of a worker getting a dominating wage offer through the direct offer distribution is quite low compared with the chance of getting one through network search. This result is due to the fact that $1-F(w)$ gets considerably smaller than $1 - \tilde{G}(w)$ when w is high. This feature is also demonstrated in Figure 3, where we saw the worker's next accepted job offer is increasingly likely to come from the network as a she climbs the wage ladder.

From another perspective, well-connected workers' fast rise implies that high wage firms have better connected workers than those at the median. On the other hand, the lowest wage firms also

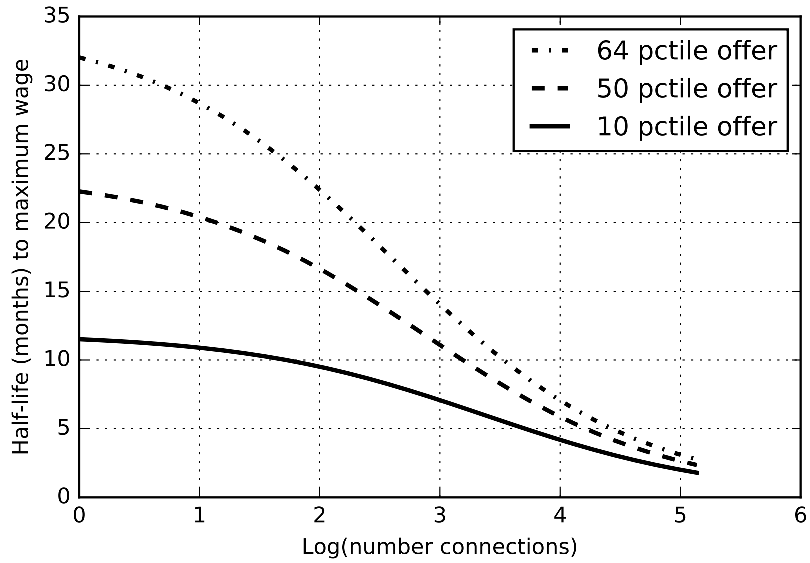

Figure 6: Half-life of wage growth paths to maximum wage: different starting wages, labeled by their quantile in the overall offer distribution, and different network connections z.

have mostly highly connected workers because they have a lower reservation wage. This leads to an asymmetric U-shape in types by firm posted wage. Despite having many well-connected workers, the very low wage firms still hire mostly via direct contact rather than referrals because γ^0 is large relative to either γ^1 or $\rho(z)$. Thus a firm hiring mostly from unemployment is not hiring much via referral. But, as wages rise and firms' workforce becomes better matched, this additional search capital implies they can poach workers from other firms much more effectively.

Figure 7 plots the fraction of a firms' hiring that comes through referrals, which rises monotonically with the posted wage. The highest wage firm has workers with almost 50% more connections than the median wage firm, and these are very effective at recruiting connections. This is also caused by a selection effect, that well-connected workers get more offers and are therefore more likely to rise to the highest wage firms.

6.4 Comparison to heterogeneous finding rates

To isolate how network heterogeneity affects the wage draw distribution from the frequency of draws, we consider a version with only heterogeneous finding rates and homogeneous offer distributions. We keep the distribution of finding rates on and off the job implied by $(\gamma^i, \rho(z))$ but we integrate over the implied offer distributions to take F and the average sampling from $\tilde{G}$. This creates a new offer distribution,

$$F^i_{HS}(w) = \int_z \gamma^i F(w) + (1-\gamma^i)\rho(z)\tilde{G}(w)\Omega(z)dz.$$

To highlight the differences between our model and one with only heterogeneous finding rates,

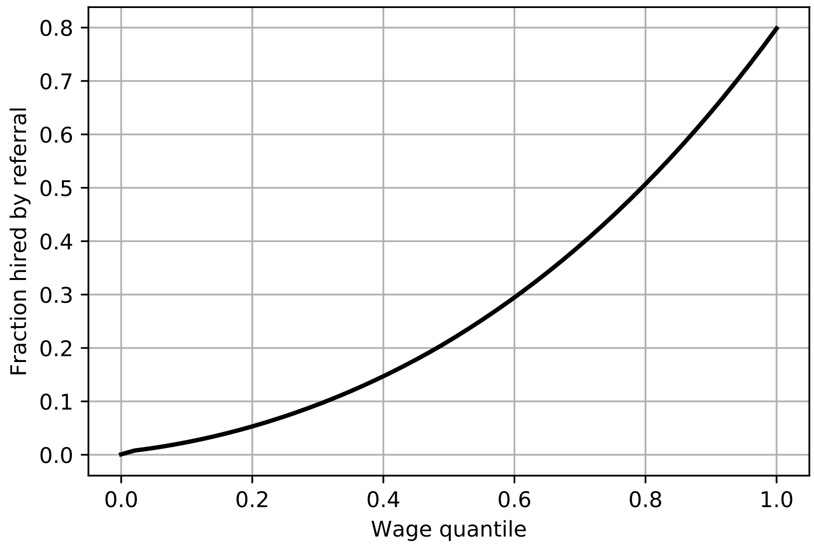

Figure 7: The fraction of hires coming from referrals at each wage.

Figure 8 plots half-lives until reaching $\bar{w}$, as before. The blue lines are the half-lives to the top wage if only finding rates differ and the black lines are the half-lives presented in Figure 6 for the network search model. With network search, the half-life decreases more quickly in z than if workers were only different in the arrival rate of offers. Those with higher numbers of peers climb the ladder faster because of faster offer arrival rates and also because the distribution from which they draw stochastically dominates both the distribution of offers for those with fewer peers. The crossing point between the blue and black lines is the point at which a type-z offer distribution begins to stochastically dominate the average offer distribution F^1_{HS}.

7 Understanding some empirical findings

Empirical work often find that networks are important to job-seekers' outcomes. Here, we discuss how these findings can be generated by different offer distributions and differences in the composition of workers who find jobs through their network. We restrict our attention to three main results in the literature and analyze them through the prism of our model. A body of empirical studies present evidence that a worker whose job is found through her network has a higher wage, a longer employment duration and a shorter unemployment spell. These regularities are often attributed to the information that network connections better convey to potential employers (e.g. Simon and Warner (1992), Topa (2011), Cappellari and Tatsiramos (2015), Marmaros and Sacerdote (2002), and Dustmann et al. (2015)), but we show and illustrate numerically that such results can also be generated in a job-ladder model with heterogeneity in network peers, as presented here.

Our model highlights two channels through which network-found matches are different than

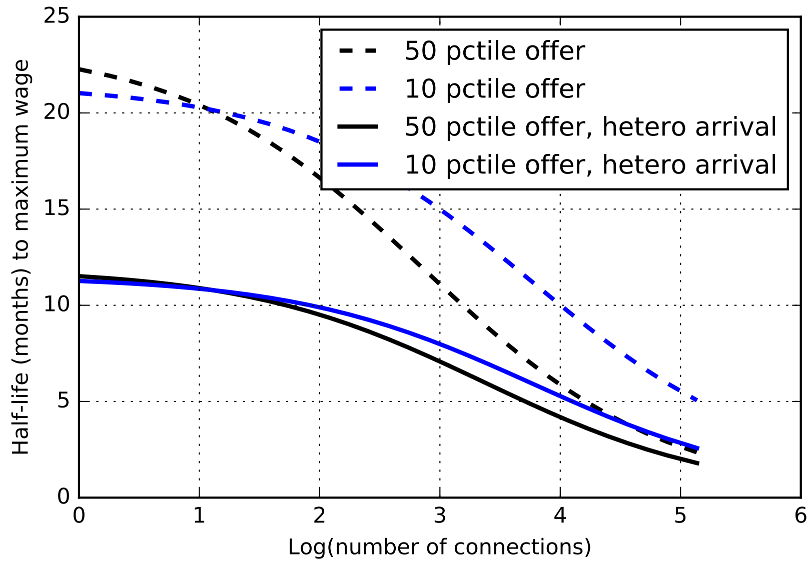

Figure 8: Half-life until the maximum wage comparing heterogeneous search rates and the network search model. Network search has shorter half-lives because better connected workers also draw from a better distribution.

those found by direct contact. First, when a worker finds a job through her own network it implies that she drew from a different (better) wage distribution and as we have shown, $\tilde{G}$ first-order stochastic dominates the distribution F from which the direct searcher finds her job.

The other, more subtle *ex post* reason why network-found matches are different is because they tend to be found by different types of people. A worker who found a job through network search had a higher arrival rate, lower reservation wage and better draw distribution because those who had found jobs through networks have, on average, more peers. The intuition is that one with more network connections gets proportionally more of her draws from the network distribution than a worker with fewer links. Then when we see the pool of searchers who found a job through peers, more of those are better-connected. Figure 9 demonstrates this idea in the calibrated model by plotting the log density of the number of peers conditional on whether a worker finds a job (either job-to-job or through unemployment) through network or direct contact search. The compositional differences will drive the observational differences across finding method.

Table III illustrates the expected differences between workers finding jobs through network or direct contact search. Above the line, the table lists statistics for workers finding jobs through unemployment and then job-to-job transitions are below the line. Workers who find jobs via their network of contacts have on average more peers, spend less time searching and should expect a higher initial wage than those workers who become employed through direct contact search. We see this effect very clearly in the two rows giving the average number of connections, z of those who make a transition from unemployment or job-to-job relative to the average at risk population. A worker who finds a job from unemployment via networks has more than 4 times as many

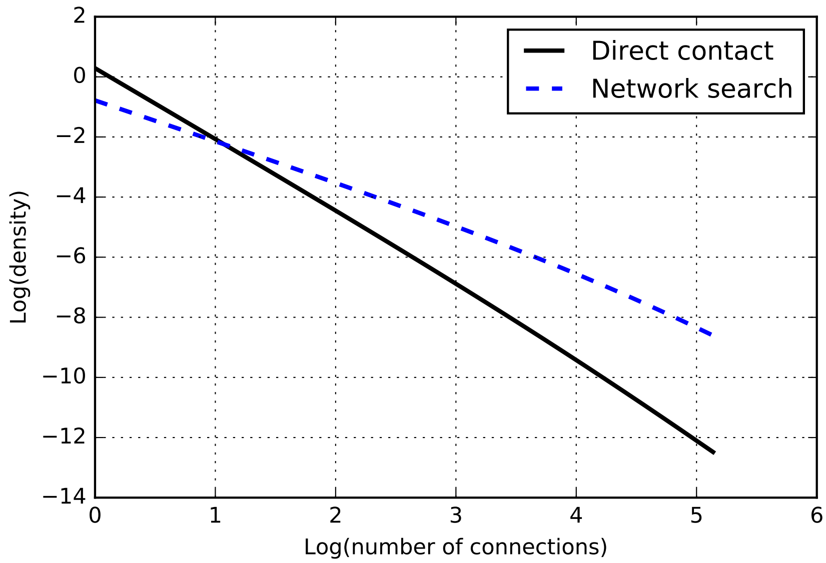

Figure 9: Log density of the number of peers conditional on having found a job by direct contact search or network search.

peers as the average unemployed person, while a worker making a job-to-job transition through network search has about 2.5 times as many as the average employed person. The difference comes from two sources. First, the stock of unemployed people disproportionately have low numbers of peers. Second the direct and referral offers distributions are both basically completely acceptable to unemployed workers, thus a worker who finds a job through network must have a very high $\rho(z)$, whereas as a worker climbs the job ladder the network search distribution, $\tilde{G}$ is becoming increasingly likely to be the main source of dominating offers, regardless of $\rho(z)$.

Both network search effects, the composition and draw distribution, imply that network-found jobs have higher wages. These forces are at odds however as they affect expected duration. The higher matched wage itself, within the job-ladder structure, implies these jobs found through the network are expected to last longer. On the other hand, a worker who finds a job through her network probably has more peers and therefore, conditional on , finds another job more quickly and has a shorter tenure at the first job. Quantitatively, the job-ladder effect dominates, and so jobs found via network search last longer. Longer match duration is consistent with empirical evidence from Hellerstein et al. (2015).

Figure 10 shows the expected duration of a job match conditional on wage for direct contact search and network search. Table III suggests the expected duration of job matches are longer for workers who find jobs through their network, i.e., have longer tenures, but this is purely because they match at a higher rung on the job ladder. Conditional on wage, a match formed through a network contact has only 77% of the average expected duration because these workers are better-connected and therefore have higher offer arrival rates. Unconditionally, however, these matches last more than twice as long because they are at higher wages and hence have a lower probability of a dominating offer.

Composition alone makes unemployment duration shorter for workers who find jobs through network search. This is a direct corollary of Lemmas 1 and 2, where we showed better connected workers had higher finding rates from unemployment. Then, unemployed workers who find jobs through their network are likely to have more peers, and therefore have a higher ex ante finding rate.[15] This feature of our model is consistent qualitatively with evidence from Cingano and Rosolia (2012). However the quantitative effects in our model are small; Table III shows there is only about a 5% decrease in unemployment duration.

	Network Search	Direct Search
Average z relative to unemployed	4.14	0.93
Expected wage qtile; from unemployment	0.251	0.075
Search time relative to average search time	0.951	1.001
Average z relative to employed	2.67	0.80
Expected wage qtile; job-to-job	0.444	0.224
Expected duration of job match	4.87 years	2.70 years

Table III: Expected differences between workers finding jobs through network or directed search. Above the line describe finding from unemployment, below adds features of job-to-job transitions.

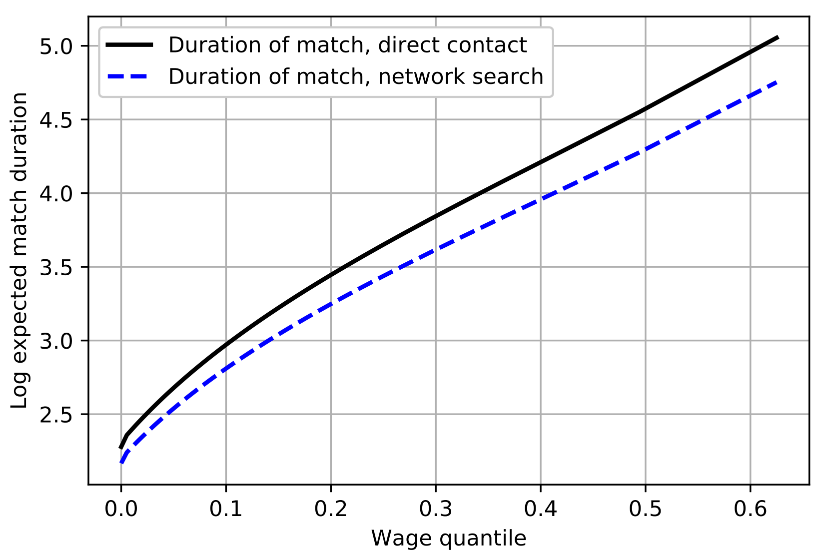

Figure 10: Expected duration of a job match conditional on wage: Direct search and network search.

[15] Qualitatively, this model also implies a downward sloping unemployment exit hazard rate. The composition of the unemployment pool becomes worse connected at higher durations. Quantitatively, the effect is small.

8 Conclusion

In this paper, we have explored how network search affects wage and employment dynamics. We embedded the well documented evidence jobs are found through networks into an equilibrium wage posting model with on-the-job search. In our model workers hear about job opportunities through their social networks and direct contact . Heterogeneity in the worker's position within the network leads to heterogeneity in wage and employment dynamics: better-connected workers climb the job ladder faster and recover more quickly if they fall. These workers also pass along higher quality referrals, which benefits their connections. In our environment, wage offer distributions are natural consequences of frictions in the job information propagation mechanism through different channels. Differences in job finding rates and reservation wages are a function of a worker's position in the social network because the number of connections governs how quickly a worker can expect to find a referral from a peer. This feature alone provides a micro-foundation for finding-rate heterogeneity; but further, we show that network-based search is important for the type of jobs that are found.

We present network and information assumptions under which it is possible to ensure that the number of peers to whom a worker is connected in a social network is actually sufficient to summarize a worker's state, her position in the network. We show that the wage offer distribution via social networks first-order stochastically dominates the wage offer distribution via direct contact search. Job offers learned from a worker's peers are endogenously better than those a worker would likely find through direct contact search. In our calibration exercise, we show how a workers differ based on whether they found their job through network or direct contact search. Those who find jobs through network-search generally have more peers than a worker who found her job via direct contact search. This different composition then manifests itself in different labor market transitions. Because those who found a job through network search have more peers, they find jobs from unemployment more quickly. Once employed, a worker who found her job through her network, and is therefore better-connected, experiences a higher wage growth for the rest of her career than another who found an equivalent job through direct contact search.

Our analysis has abstracted from many quantitatively important factors, particularly firm-side heterogeneity and other types of worker-side heterogeneity. It is, however, a very flexible starting point to which these other features can be grafted. Our approach, simplifying the state using the mean field approach makes many rich models with network search tractable. We hope that, in the future, it can be used to study how network heterogeneity affects classic questions regarding worker-side heterogeneity such as earnings inequality and unemployment duration dependence. Furthermore, network search can be uniquely useful in understanding measured vacancy yields which across sectors and time.

References

Armantier, O., G. Topa, W. Van der Klaauw, and B. Zafar (2017): "An overview of the Survey of Consumer Expectations," *Economic Policy Review*, 51–72.

Bagger, J., F. Fontaine, F. Postel-Vinay, and J.-M. Robin (2014): "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics," *The American Economic Review*, 104, 1551–1596.

Bayer, P., S. L. Ross, and G. Topa (2008): "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 116, 1150–1196.

Bender, E. A. and E. Canfield (1978): "The asymptotic number of labeled graphs with given degree sequences," *Journal of Combinatorial Theory, Series A*, 24, 296 – 307.

Bentolila, S., C. Michelacci, and J. Suarez (2010): "Social Contacts and Occupational Choice," *Economica*, 77, 20–45.

Bontemps, C., J.-M. Robin, and G. J. van den Berg (2000): "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation," *International Economic Review*, 41, 305–58.

Brown, M., E. Setren, and G. Topa (2013): "Do informal referrals lead to better matches? Evidence from a firm's employee referral system," Staff Reports 568, Federal Reserve Bank of New York.

Burdett, K. and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39, 257–73.

Burks, S., B. Cowgill, M. Hoffman, and M. Housman (2015): "The Value of Hiring through Employee Referrals," *The Quarterly Journal of Economics*, 130, 805–839.

Calvo-Armengol, A. and M. O. Jackson (2004): "The Effects of Social Networks on Employment and Inequality," *American Economic Review*, 94, 426–454.

——— (2007): "Networks in labor markets: Wage and employment dynamics and inequality," *Journal of Economic Theory*, 132, 27–46.

Cappellari, L. and K. Tatsiramos (2015): "With a little help from my friends? Quality of social networks, job finding and job match quality," *European Economic Review*, 78, 55–75.

Chung, F. and L. Lu (2002): "Connected Components in Random Graphs with Given Expected Degree Sequences," *Annals of Combinatorics*, 6, 125–145.

Cingano, F. and A. Rosolia (2012): "People I Know: Job Search and Social Networks," *Journal of Labor Economics*, 30, 291 – 332.

Datcher, L. (1983): "The Impact of Informal Networks of Quit Behavior," *The Review of Economics and Statistics*, 65, 491–95.

Dorogovtsev, S. N., A. V. Goltsev, and J. F. F. Mendes (2008): "Critical phenomena in complex networks," *Reviews of Modern Physics*, 80, 1275–1335.

Dustmann, C., A. Glitz, and U. Schonberg (2015): "Referral-based Job Search Networks," *The Review of Economic Studies*, forthcoming.

Faberman, R. J., A. I. Mueller, A. Sahin, and G. Topa (2017): "Job Search Behavior among the Employed and Non-Employed," NBER Working Papers 23731, National Bureau of Economic Research, Inc.

Feld, S. L. (1991): "Why Your Friends Have More Friends Than You Do," *American Journal of Sociology*, 96, 1464–1477.

Fontaine, F. (2008): "Why are similar workers paid differently? the role of social networks," *Journal of Economic Dynamics and Control*, 32, 3960–3977.

Galenianos, M. (2014): "Hiring through referrals," *Journal of Economic Theory*, 152, 304–323.

Glitz, A. (2013): "Coworker Networks in the Labour Market," IZA Discussion Papers 7392, Institute for the Study of Labor (IZA).

Granovetter, M. (1995): *Getting a Job: A Study of Contacts and Careers*, Chicago, IL, USA: 2nd ed., University of Chicago Press.

Hellerstein, J. K., M. J. Kutzbach, and D. Neumark (2015): "Labor Market Networks and Recovery from Mass Layoffs Before, During, and After the Great Recession," Working Paper 21262, National Bureau of Economic Research.

Hellerstein, J. K., M. McInerney, and D. Neumark (2011): "Neighbors and Coworkers: The Importance of Residential Labor Market Networks," *Journal of Labor Economics*, 29, 659 – 695.

Hofstad, R. v. d. (2016): *Random Graphs and Complex Networks*, vol. 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press.

Holzer, H. (1988): "Search Method Use by Unemployed Youth," *Journal of Labor Economics*, 6, 1–20.

Ioannides, Y. M. and L. Datcher Loury (2004): "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42, 1056–1093.

Ioannides, Y. M. and A. R. Soetevent (2006): "Wages and Employment in a Random Social Network with Arbitrary Degree Distribution," *American Economic Review*, 96, 270–274.

Kugler, A. (2003): "Employee referrals and efficiency wages," *Labour Economics*, 10, 531–556.

Marmaros, D. and B. Sacerdote (2002): "Peer and social networks in job search," *European Economic Review*, 46, 870–879.

Menzio, G., I. Telyukova, and L. Visschers (2016): "Directed Search over the Life Cycle," *Review of Economic Dynamics*, 19, 38–62.

Montgomery, J. D. (1991): "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *American Economic Review*, 81, 1407–18.

Mortensen, D. T. and T. Vishwanath (1994): "Personal contacts and earnings : It is who you know!" *Labour Economics*, 1, 187–201.

Newman, M. E. J. (2004): "Coauthorship networks and patterns of scientific collaboration," *Proceedings of the National Academy of Sciences*, 101, 5200–5205.

Pastor-Satorras, R., C. Castellano, P. Van Mieghem, and A. Vespignani (2015): "Epidemic processes in complex networks," *Reviews of Modern Physics*, 87, 925–979.

Pissarides, C. A. (2000): *Equilibrium Unemployment Theory, 2nd Edition*, vol. 1 of *MIT Press Books*, The MIT Press.

Postel-Vinay, F. and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70, 2295–2350.

Postel-Vinay, F. and H. Turon (2010): "On-The-Job Search, Productivity Shocks, And The Individual Earnings Process," *International Economic Review*, 51, 599–629.

Simon, C. J. and J. T. Warner (1992): "Matchmaker, Matchmaker: The Effect of Old Boy Networks on Job Match Quality, Earnings, and Tenure," *Journal of Labor Economics*, 10, 306–30.

Topa, G. (2011): "Chapter 22 - Labor Markets and Referrals," in *Handbook of Social Economics*, ed. by A. B. Jess Benhabib and M. O. Jackson, North-Holland, vol. 1 of *Handbook of Social Economics*, 1193 – 1221.

Ugander, J., B. Karrer, L. Backstrom, and C. Marlow (2011): "The Anatomy of the Facebook Social Graph," *CoRR*, abs/1111.4503.

Vega-Redondo, F. (2007): *Complex social networks*, 44, Cambridge University Press.

A Proofs Appendix

A.1 Proof of Proposition 1

In this section we will describe the value functions of the worker's problem under the full information structure described in section 2.4, and show that the imposition of Assumptions A.1-A.3 simplify these functions to the model analyzed in section 3. Without imposing Assumptions A.1-A.3, the worker's state consists of their employment status i, wage w, the history of wages offers they have received, k, and recursively, each of these states for each of their peers, i.e. their network position χ. In this setting, a worker's degree z is not sufficient to characterize their network position; two workers of the same degree z may have peers whose own network position κ is very different, and who thus face a different value of job search via networks. The rate of contact from a particular peer, c of type $\kappa(c)$, where $s = ||\kappa(c)||$ is given by

$$i(c)\gamma^1 \frac{\nu}{s}$$

The probability that this peer c is the one that passes a job is a binomial random variable, that depends on the realization of the job contact probabilities from the entire set of peers. For example, it may be that exactly j of the worker's z peers attempt to transfer a wage offer, and with probability $\frac{1}{j}$ this particular peer of type i, w, k, κ is successful.[16] Call this probability $p(c)$.

$$\begin{aligned} rV^0(k,\chi) &= b + \gamma^0 \int_{R(\chi)}^{\bar{w}} V^1(x,k,\chi) - V^0(k,\chi) dF(x) \\ &\quad + (1-\gamma^0) \int_c p(c) \left(V^1(w(c),k,\chi) - V^0(k,\chi)\right) \mathbb{I}_{w(c)\geq R(k,\chi)} dc \qquad (25) \\ rV^1(w,k,\chi) &= w + \delta\left[V^0(k,\chi) - V^1(w,k,\chi)\right] + \gamma^1 \int_w^{\bar{w}} V^1(x,k,\chi) - V^0(k,\chi) dF(x) \\ &\quad + (1-\gamma^1) \int_c p(c) \left(V^1(w(c),k,\chi) - V^1(w,k,\chi)\right) \mathbb{I}_{w(c)\geq w} dc \qquad (26) \end{aligned}$$

Note that the integrations in these value functions are over the known state of each peer of the worker, so our notation c is more explicitly written as $(i(c), w(c), k(c), \kappa(c)) \in \chi$. The only uncertainty is in the problem arises from the friction of the wage transmission procedure, not in the positions or wages of peers.

[16] In this case, the worker with complete information still receives only one job at random in the event that multiple jobs offers are passed to him; she cannot choose, for example, the highest wage offer, even though she knows which peer has the highest wage. We interpret this as a remaining friction in the labor market, even with complete information on the state. One can imagine alternative offer selection procedures, but this would change the wage dynamics drastically from our baseline model, and are not pursued here.

A.1.1 Limited Observability

We now consider the worker's problem when she has limited information on the status of her peers. We will first impose Assumption A.1, that the network was generated according to the configuration model taken to the infinite limit, and Assumption A.2, that the worker does not know the state of any of her peers. Under these assumptions, the worker knows only her own history of wages, and may use this information, along with nature of the configuration model, to infer the network position κ of each of her z peers, and therefore their employment status and wages.

Because she no longer knows the wages of her peers, she cannot distinguish which of them may have passed her particular wage offers in the past. She must therefore view each peer as a draw from a peer type distribution. In the infinite limit of the configuration model of network formation, the network is *locally tree-like*[17]; The types of each of her peers are independent of each other; they are independent draws from some type distribution. She will form a belief of this distribution that depends on her history k; call her belief that a given peer is of type s $\Psi(s|k)$. Importantly the type of peer c is $s(c)$ rather than $s(c), k(c)$ because the peer's own history of offers does not affect the offers she refers.

Each peer of type s is herself the root of a tree whose first generation is of size s, and whose types are drawn from some distribution. The subsequent generations of this tree - the peers of a worker's peers - she will believe are drawn from a different distribution than $\Psi(s|k)$; this is because her local information is less informative the further from her immediate peers the tree progresses. The "offspring" distribution of the branching process that describes her network position is therefore not IID across generations, and each worker's idiosyncratic history k will lead to a different belief about her local network position.

Under this information structure, the probability that at least one of her peers passes her a job is given by

$$\rho(z, k) = \left(1 - \exp\left(-z\left[\int_s n(s)\gamma^1 \frac{\nu}{s}\Psi(s|k)ds\right]\right)\right) \tag{27}$$

She can no longer assign different probabilities of being passed an offer to each peer, as she cannot distinguish one peer from another, but she can come to expect that the probability of being passed any job at all may be higher or lower, as her own history brings her to believe her peers are

[17] See Dorogovtsev et al. (2008) and Hofstad (2016) for a discussion of the "tree ansatz" method for complex networks.

of higher or lower degree s. Under these assumptions, the household value functions are given by

$$\begin{aligned} rV^0(k,z) &= b+\gamma^0\int_{R(k,z)}^{\bar{w}}(V^1(x,k,z)-V^0(k,z))dF(x) \\ &+ (1-\gamma^0)\rho(z,k)\left[\int_s\left[\int_w^{\bar{w}}(V^1(x,k,z))-V^0(k,z))Pr(x|s,k)dx\right]\Psi(s|k)ds\right] \\ rV^1(w,k,z) &= w+\delta\left[V^0(k,z)-V^1(w,k,z)\right]+\gamma^1\int_w^{\bar{w}}V^1(x,k,z)-V^0(k,z)dF(x) \end{aligned}$$

Each worker of type z and history k will form an expectation of the types s of each of their peers, and the wages they might be passed by such a peer. Her belief about the distribution of wages of a random peer is given by $Pr(w|s,k)$, and will depend on that peers own type s and her history k.

A.1.2 Limited Memory

Under Assumption A.3, workers can no longer recall their history of wage offers, and so can no longer update their beliefs about the types of their peers beyond what can be inferred from the structure of the network. Because she can no longer condition on the history k, the worker's value function of being unemployed is given by

$$\begin{aligned} rV^0(z) &= b+\gamma^0\int_{R(z)}^{\bar{w}}(V^1(x,z)-V^0(z))dF(x) \\ &+ (1-\gamma^0)\rho(z)\left[\int_s\left[\int_x(V^1(x,z))-V^0(z))Pr(x|s)dx\right]\Psi(s)ds\right] \\ rV^1(w,z) &= w+\delta\left[V^0(z)-V^1(w,z)\right]+\gamma^1\int_w^{\bar{w}}V^1(x,z)-V^0(z)dF(x) \\ &+ (1-\gamma^1)\rho(z)\left[\int_s\left[\int_w^{\bar{w}}(V^1(x,z))-V^1(w,z))Pr(x|s)dx\right]\Psi(s)ds\right]. \end{aligned}$$

The distribution of peer types s is now simply the peer type distribution $\Psi(s)$; the worker's idiosyncratic history is forgotten, so she does not believe her peers are any different than the average peer in the network; furthermore, she will believe that the distribution of the types of her peers' peers are also no different than that of the average peer. The "offspring distribution" of the branching process that describes her network position is the same for every worker: $\Psi(s)$, independently identically distributed across nodes and across generations. Workers only meaningfully differ in the size of the first generation of this process, z, so each worker of type z has the same belief about her local network position.

The distribution of wages conditional on s, $Pr(x,s)$, is now $\frac{\partial G(x,s)}{\partial x}$, so that the value functions coincide exactly with section 3. Taken together, these assumptions amount to the worker behaving as though the types of her peers are drawn anew every period, as if the network were randomly rewired every period. In this "annealed" network (as opposed to the realized, or "quenched" network), the degree-based mean-field dynamics correctly characterize the network, and each worker's state is fully characterized by (i, w, z).

A.2 Proof of Lemma 1

Lemma 3. *The reservation wage $R(z)$ is decreasing in z: $\frac{\partial R(z)}{\partial z} \leq 0$.*

Proof. The reservation wage $R(z)$ of worker of type z is defined by

$$R(z) = b + (\gamma^0 - \gamma^1) \int_{R(z)}^{\bar{w}} V_x(z,x)(1 - F(x))dx - \Big((\gamma^0 - \gamma^1)\rho(z)\Big) \int_{R(z)}^{\bar{w}} V_x(z,x)(1 - \tilde{G}(x))dx$$

and it's derivative is therefore given by

$$\begin{aligned} R'(z) =& (\gamma^0 - \gamma^1) \left(\int_{R(z)}^{\bar{w}} \frac{\partial V_x(z,x)}{\partial z} (1 - F(x))\, dx + V^0(z) \left(1 - F(R(z))\right) R'(z) \right) \\ & -(\gamma^0 - \gamma^1)\rho(z) \left(\int_{R(z)}^{\bar{w}} \frac{\partial V_x(z,x)}{\partial z} (1 - \tilde{G}(x))dx + V^0(z)(1 - \tilde{G}(R(z)))R'(z) \right) \\ & -(\gamma^0 - \gamma^1)\rho'(z) \int_{R(z)}^{\bar{w}} V_x(z,x)(1 - \tilde{G}(x))dx \end{aligned}$$

Solving for $R'(z)$, we have

$$R'(z) = \frac{Q}{1 + (\gamma^0 - \gamma^1)V^0(z)(1 - F(R(z))) - (\gamma^0 - \gamma^1)\rho(z)V^0(z)(1 - \tilde{G}(R(z)))} \tag{28}$$

where

$$\begin{aligned} Q =& (\gamma^0 - \gamma^1) \left(\int_{R(z)}^{\bar{w}} \frac{\partial V_x(z,x)}{\partial z} (1 - F(x))\, dx \right) \\ & -(\gamma^0 - \gamma^1)\rho(z) \left(\int_{R(z)}^{\bar{w}} \frac{\partial V_x(z,x)}{\partial z} (1 - \tilde{G}(x))dx \right) \\ & -(\gamma^0 - \gamma^1)\rho'(z) \int_{R(z)}^{\bar{w}} V_x(z,x)(1 - \tilde{G}(x))dx \end{aligned} \tag{29}$$

The denominator of Equation (28) is positive. It is in fact the value of the integrand that defines $R(z)$ evaluated at its lower bound, and so it is positive. Hence, to show that $R'(z) \leq 0$, it suffices to show that $Q \leq 0$. Since $V_x(z, w)$ is given by

$$V_x(z,x) = (r + \delta + \gamma^1(1 - F(x)) + (1 - \gamma^1)\rho(z)(1 - \tilde{G}(x)))^{-1},$$

we have

$$\frac{\partial V_x(z,x)}{\partial z} = (-1)(1 - \gamma^1)\rho'(z)(1 - \tilde{G}(x))(V_x(z,x))^2.$$

We can write the integral in the first, second and third terms of Q (Equation 29) as

$$\int_{R(z)}^{\bar{w}} (\gamma^0 - \gamma^1) \left((-1)(1-\gamma^1)\rho'(z)(1-\tilde{G}(x))(V_x(z,x))^2 \right) (1-F(x))dx$$

$$\int_{R(z)}^{\bar{w}} -(\gamma^0 - \gamma^1)\rho(z) \left((-1)(1-\gamma^1)\rho'(z)(1-\tilde{G}(x))(V_x(z,x))^2 \right) (1-\tilde{G}(x))dx$$

$$\int_{R(z)}^{\bar{w}} -(\gamma^0 - \gamma^1)\rho'(z)V_x(z,x)(1-\tilde{G}(x))dx$$

Since these integrals are all over the same domain, we can combine them, so that the sign of $R'(z)$ depends on the sign of the integrand in the definition of Q, which is

$$\begin{aligned} &(\gamma^0 - \gamma^1) \left((-1)(1-\gamma^1)\rho'(z)(1-\tilde{G}(x))(V_x(z,x))^2 \right) (1-F(x)) \\ &-(\gamma^0 - \gamma^1)\rho(z) \left((-1)(1-\gamma^1)\rho'(z)(1-\tilde{G}(x))(V_x(z,x))^2 \right) (1-\tilde{G}(x))) \\ &-(\gamma^0 - \gamma^1)\rho'(z)V_x(z,x)(1-\tilde{G}(x)) \end{aligned} \tag{30}$$

We can write the terms in expression (30) as, respectively

$$\left((1-\gamma^1)\rho'(z)V_x(z,x)(1-\tilde{G}(x))\right)\left((\gamma^0-\gamma^1)(-1)V_x(z,x)(1-F(x))\right),$$

$$\left((1-\gamma^1)\rho'(z)V_x(z,x)(1-\tilde{G}(x))\right)\left((-1)(\gamma^0-\gamma^1)\rho(z)(-1)V_x(z,x)(1-\tilde{G}(x))\right),$$

$$(-1)(\gamma^0-\gamma^1)\rho'(z)V_x(z,x)(1-\tilde{G}(x))$$

Collecting terms, we can write this as

$$\left((1-\gamma^1)\rho'(z)V_x(z,x)(1-\tilde{G}(x))\right)(-1)\left((\gamma^0-\gamma^1)V_x(z,x)(1-F(x))-(\gamma^0-\gamma^1)\rho(z)V_x(z,x)(1-\tilde{G}(x))+1\right)$$

Since the expression $(1-\gamma^1)\rho'(z)V_x(z,x)(1-\tilde{G}(x)) \geq 0$, the sign depends on

$$(-1)\left((\gamma^0-\gamma^1)V_x(z,x)(1-F(x))-(\gamma^0-\gamma^1)\rho(z)V_x(z,x)(1-\tilde{G}(x))+1\right)$$

Notice that $(\gamma^0-\gamma^1)V_x(z,x)(1-F(x))-(\gamma^0-\gamma^1)\rho(z)V_x(z,x)(1-\tilde{G}(x))$ is in fact the integrand that defines $R(z)$, which must be non-negative from the definition of the value function, that is, that $V_x(z,x)(1-F(x)) = \left(V^1(z,x) - V^0(z)dF(x)\right)$. Thus, due to the leading (-1), this expression is negative, so that $Q \leq 0$, and thus $R'(z) \leq 0$, as was to be shown. □

A.3 Proof of Lemma 2

Lemma 4. *The employment rate $n(z)$ is increasing in z: $\frac{\partial n(z)}{\partial z} \geq 0$.*

Proof. Let $P(z)$ be the job finding rate out of unemployment:

$$P(z) = \gamma^0(1 - F(R(z)) + (1 - \gamma^0)\rho(z)(1 - \tilde{G}(R(z)))$$

Then, by Equation (13), $n(z)$ is given by

$$n(z) = \frac{P(z)}{\delta + P(z)}$$

and so

$$n'(z) = \frac{\delta P'(z)}{(\delta + P(z))^2}$$

Hence $n'(z)$ and $P'(z)$ share the same sign; if P' is positive, the result is proven. We have

$$P'(z) = -\gamma^0 F'(R(z))R'(z) - (1 - \gamma^0)(\rho(z)\tilde{G}'(R(z))R'(z) - (1 - \gamma^0)(1 - \tilde{G}(R(z))\rho'(z)))$$

This is positive if

$$R'(z) \leq \frac{\rho'(z)(1 - \gamma^0)(1 - \tilde{G}(R(z)))}{\gamma^0 F'(R(z)) + (1 - \gamma^0)\rho(z)\tilde{G}'(R(z))}$$

This states that if $R'(z)$ is smaller than a positive cutoff, $n'(z)$ is positive. This will certainly hold if $R'(z) < 0$, as was shown above.

□

A.4 Proof of Proposition 2

Proposition 5. *$\tilde{G}(w)$ first-order stochastically dominates $G(w)$.*

Proof. To show that $\tilde{G}(w)$ first-order stochastically dominates $G(w)$, we will now that $\tilde{\Psi}$ first-order stochastically dominates (FOSD) Ω. Recall that $\Psi = \frac{s\Omega}{<s>}$, so $\frac{\Psi}{\Omega} = \frac{s}{<s>}$, which is increasing, and it is how we show Ψ FOSD Ω. If $\frac{\tilde{\Psi}}{\Omega}$ is increasing, this result is proven.

Note, it is *not* true that $\tilde{\Psi}$ FOSD Ψ. In fact, the opposite is true. So, we have

$$\begin{aligned}
\tilde{\Psi} &= \frac{n(s)\frac{1}{s}\Psi}{<\tilde{s}>} \\
&= \frac{n(s)\frac{1}{s}\frac{s\Omega}{<s>}}{<\tilde{s}>} \\
&= \frac{n(s)\Omega}{<\tilde{s}><s>} \\
\frac{\tilde{\Psi}}{\Omega} &= \frac{n(s)}{<\tilde{s}><s>} \equiv \frac{n(s)}{N}
\end{aligned}$$

Since $n(s)$ is increasing the result is proven. □

A.5 Proof of Proposition 3

Proposition 6. *$\tilde{G}(w)$ first-order stochastically dominates $F(w)$.*

Proof. Finally, we show that $\tilde{G}$ first-order stochastically dominates F. To establish this, it suffices to show that G first-order stochastically dominates F. Recall the job finding rate out of unemployment, $P(z)$

$$P(z) = \gamma^0(1 - F(R(z)) + (1 - \gamma^0)\rho(z)(1 - \tilde{G}(R(z)))$$

We now define two functions

$$\begin{aligned} P^0(w, z) &= \gamma^0 F(w) + (1 - \gamma^0)\rho(z)\tilde{G}(w) \\ P^1(w, z) &= \gamma^1 F(w) + (1 - \gamma^1)\rho(z)\tilde{G}(w) \end{aligned}$$

where $P^0(R(z), z) = P(z)$ is the job finding rate out of unemployment, while $P^1(w, z)$ is the job finding rate when employed at a firm offering wage w.

From the definition of $G(w, z)$ and $n(z)$, we can write

$$\begin{aligned} G(w, z) &= \frac{1 - n(z)}{n(z)} \frac{P^0(w, z) - P^0(R(z), z)}{\delta + (P^1(\bar{w}, z) - P^1(w, z))} \\ \frac{1 - n(z)}{n(z)} &= \frac{\delta}{P^0(\bar{w}, z) - P^0(R(z), z)} \end{aligned}$$

Substituting, we have

$$\begin{aligned} G(w, z) &= \frac{\delta \frac{P^0(w,z) - P^0(R(z),z)}{P^0(\bar{w},z) - P^0(R(z),z)}}{\delta + (P^1(\bar{w}, z) - P^1(w, z)))} \\ &= \frac{\frac{P^0(w,z) - P^0(R(z),z)}{P^0(\bar{w},z) - P^0(R(z),z)}}{1 + \frac{1}{\delta}(P^1(\bar{w}, z) - P^1(w, z)))} \\ &\leq \frac{P^0(w, z) - P^0(R(z), z)}{P^0(\bar{w}, z) - P^0(R(z), z)} \\ &\leq P^0(w, z) - P^0(R(z), z) \\ &\leq P^0(w, z) \end{aligned}$$

where the penultimate inequality follows because $P^0(\bar{w}, z) > P^0(w, z) > P^0(R(z), z)$, and ultimate inequality follows because $P^0(R(z), z) \geq 0$. Thus, we have

$$G(w, z) \leq P^0(w, z),$$

showing the the earnings distribution of workers of type z first order stochastically dominates a “weighted ” offer distribution for these workers. This implies

$$\begin{aligned}
G(w,z) &\leq P^0(w,z)\\
G(w,z) &\leq \gamma^0(F(w)) + (1-\gamma^0)\rho(z)(\tilde{G}(w))
\end{aligned}$$

Integrating both sides with respect to the neighbor type distribution $\tilde{\Psi}(z)$, we have

$$\tilde{G}(w) \leq \gamma^0(F(w)) + (1-\gamma^0)\tilde{G}(w)\tilde{\rho},$$

where $\tilde{\rho} = \int \rho(z)\tilde{\Psi}(z)dz$ is the “average” network contact rate among a worker's peers. This implies

$$\tilde{G}(w) \leq \frac{\gamma^0 F(w)}{1-(1-\gamma^0)\tilde{\rho}} \leq F(w)$$

where the final inequality follows because $\frac{\gamma^0}{1-(1-\gamma^0)\tilde{\rho}} \leq 1$. Thus the result is proven. □

A.6 Proof of Proposition 4

Lemma 5. *The reservation wage is declining in ν, $\frac{\partial R(z;\nu)}{\partial \nu} < 0\Big|_{\nu=0,}$*

Proof. Evaluating the partial derivative:

$$\begin{aligned}
\frac{\partial R(z;\nu)}{\partial \nu} = &-\left(\frac{\partial \rho(z)}{\partial \nu}(\gamma^0-\gamma^1)\right)\int_R^{\bar{w}} V_x(z,x)(1-G(x))dx\\
&-(\gamma^0-\gamma^1)(1-\rho(z))V_x(z,R)
\end{aligned}$$

Because $\frac{\partial \rho}{\partial \nu} = (1-\rho(z))(z\gamma^1 \int \frac{n(s)}{s}\tilde{\Psi}(s)ds) > 0$ and $V_x > 0$ for any value x, both lines are definitely negative and so is the entire partial derivative. □

Lemma 6. *Holding constant equilibrium distributions (n, G, F), the lower-bound, $\underline{w}$, is decreasing in ν from $\nu = 0$: $\frac{\partial \underline{w}}{\partial \nu}\Big|_{\nu=0} \leq 0$*

Proof. Let $\underline{w}^0$ be the initial lowest wage when $\nu = 0$. At $\nu = 0$ we know that $R(z) = R = \underline{w}^0\ \forall z$ because this is simply the BM case. If instead $\nu > 0$ then $R(z) < \underline{w}^0\ \forall z > 0$.

Consider a firm offering $w' \leq \underline{w}^0$, then this firm makes a larger profit than $(1-\underline{w}^0)L(\underline{w}^0)$ if $(1-w')L(w') \geq (1-\underline{w}^0)L(\underline{w}^0)$. Because w' may be arbitrarily close to $\underline{w}^0$, we need to show that at $\underline{w}^0$ the decrease in the firm size $\frac{\partial L(\underline{w}^0)}{\partial w}$ does not have larger magnitude than the increase in profit per worker. Put another way, we need to verify that at $\nu > 0$ that $\frac{\partial \pi(\underline{w}^0)}{\partial w} = (1-\underline{w}^0)\frac{\partial L(\underline{w}^0)}{\partial w} - L(\underline{w}^0) \leq 0$

Rather than setting conditions to establish its magnitude relative to $\frac{L(\underline{w}^0)}{(1-\underline{w}^0)}$, we can approach this piece-wise for the two cases:

1. If $\frac{\partial L(\underline{w})}{\partial w} < \frac{L(\underline{w}^0)}{1-\underline{w}^0}$ then the strict inequality holds, $\frac{\partial \underline{w}}{\partial \nu}\Big|_{\nu=0} < 0$.

2. If instead $\frac{\partial L(w)}{\partial w} \geq \frac{L(\underline{w}^0)}{1-\underline{w}^0}$. Then $w'' > \underline{w}$ cannot be the lowest wage offered because a deviation of $\underline{w}$ would attract the same size firm—recruiting anyone from unemployment—but make higher profit. Hence, in these cases $\left.\frac{\partial \underline{w}}{\partial \nu}\right|_{\nu=0} = 0$.

$\square$

Lemma 7. *For sufficiently high* γ^1, $\left.\frac{d\pi}{d\nu}\right|_{\nu=0} \leq 0$

Proof. To show it is negative, we will directly evaluate the derivative of π with respect to ν:

$$\frac{d\pi}{d\nu} = -\frac{\partial \underline{w}}{\partial \nu} L(\underline{w}) + (1-\underline{w})\left(\frac{\partial L(\underline{w};\nu)}{\partial \nu} + \frac{\partial L(\underline{w})}{\partial w}\frac{\partial \underline{w}}{\partial \nu}\right)$$

The last term reflects the fact that L depends directly on ν and its evaluation point $\underline{w}$ also depends on ν, as shown in Lemma 6.

Recall, the lowest wage for any $\nu \geq 0$, $\underline{w}(\nu)$ must satisfy

$$\frac{\partial \pi}{\partial w} = 0 = (1-\underline{w}(\nu))\frac{\partial L(\underline{w}(\nu);\nu)}{\partial w} - L(\underline{w}(\nu);\nu).$$

Now, we plug in the first-order condition, that at optimal $\underline{w}$, $L(\underline{w}) = (1-\underline{w})\frac{\partial L(\underline{w})}{\partial \underline{w}}$ to get

$$\frac{d\pi}{d\nu} = -\frac{\partial \underline{w}}{\partial \nu}(1-\underline{w})\frac{\partial L(\underline{w})}{\partial \underline{w}} + (1-\underline{w})\left(\frac{\partial L(\underline{w};\nu)}{\partial \nu} + \frac{\partial L(\underline{w})}{\partial w}\frac{\partial \underline{w}}{\partial \nu}\right)$$

Canceling terms, the sign of $\frac{d\pi}{d\nu}$ depends only on the partial $\frac{\partial L(\underline{w};\nu)}{\partial \nu}$ because $(1-\underline{w}) \geq 0$. To address this, we will first take the partial of $\ell(\underline{w}, z)$, and then integrate over these,

$$\begin{aligned}\frac{\partial \ell(\underline{w}, z;\nu)}{\partial \nu} =& \frac{\partial}{\partial \nu}\left(\frac{(1-n(z))\left(\frac{\Omega(z)}{M}\gamma^0 + \nu\tilde{\Psi}(z)\int \ell(\underline{w}, t)tdt\right)}{\delta + \gamma^1 + (1-\gamma^1)\rho(z)}\right)\\ =& \frac{(1-n(z))\left(\tilde{\Psi}(z)\int_t \ell(\underline{w}, t)tdt + \nu\tilde{\Psi}(z)\int \frac{\partial \ell(\underline{w},t)}{\partial \nu}tdt\right)}{\delta + \gamma^1 + (1-\gamma^1)\rho(z)} - \frac{\frac{\partial \rho(z)}{\partial \nu}(1-\gamma^1)}{\delta + \gamma^1 + (1-\gamma^1)\rho(z)}\ell(\underline{w}, z)\end{aligned}$$

evaluating this at $\nu = 0$ and plugging in $\frac{\partial \rho(z)}{\partial \nu} = (1-\rho(z))(z\gamma^1 \int \frac{n(s)}{s}\tilde{\Psi}(s)ds)$ gives us

$$\frac{\partial \ell(\underline{w}, z;\nu)}{\partial \nu} = \frac{(1-n(z))\tilde{\Psi}(z)\int_t \ell(\underline{w}, t)tdt}{\delta + \gamma^1} - \frac{\left(z\gamma^1 \int \frac{n(s)}{s}\tilde{\Psi}(s)ds\right)(1-\gamma^1)}{\delta + \gamma^1}\ell(\underline{w}, z) \tag{31}$$

This is the increase in hiring through referrals minus the increase in outflows due to other firms also hiring more quickly through referrals. Now, taking the partial derivative of the whole firm, plugging in Equation 31 at all z we have

$$\begin{aligned}\frac{\partial L}{\partial \nu} =& \frac{\partial}{\partial \nu}\left(\int_{\hat{z}}^{\infty} \ell(\underline{w}, z)dz\right)\\ =& \int_{\hat{z}}^{\infty} \frac{(1-n(z))\tilde{\Psi}(z)\int_t \ell(\underline{w},t)tdt}{\delta+\gamma^1} - \frac{\left(z\gamma^1 \int \frac{n(s)}{s}\tilde{\Psi}(s)ds\right)(1-\gamma^1)}{\delta+\gamma^1}\ell(\underline{w},z)dz\\ &- \frac{(1-n(\hat{z}))\left(\frac{\Omega(\hat{z})}{M}\gamma^0 + \nu\tilde{\Psi}(\hat{z})\int \ell(\underline{w},t)tdt\right)}{\delta+\gamma^1}\frac{\partial \hat{z}}{\partial \nu} \end{aligned} \tag{32}$$

Then we can rearrange Equation 32 to be

$$\begin{aligned}\frac{\partial L}{\partial \nu} =& \frac{1}{\delta+\gamma^1}\left(\int_{\hat{z}}^{\infty}(1-n(z))\tilde{\Psi}(z)dz\int_{\hat{z}}^{\infty}\ell(\underline{w},t)tdt - \gamma^1(1-\gamma^1)\int_1^{\infty}\frac{n(s)}{s}\tilde{\Psi}(s)ds\int_{\hat{z}}^{\infty}\ell(\underline{w},z)zdz\right)\\ &- \frac{(1-n(\hat{z}))\left(\frac{\Omega(\hat{z})}{M}\gamma^0+\nu\tilde{\Psi}(\hat{z})\int\ell(\underline{w},t)tdt\right)}{\delta+\gamma^1}\frac{\partial\hat{z}}{\partial\nu}\end{aligned} \tag{33}$$

The first term is negative iff $\int(1-n(z))\tilde{\Psi}(z)dz < \gamma^1(1-\gamma^1)\int\frac{n(z)}{z}\tilde{\Psi}(z)dz$, where both of the integrals are evaluated from 1 to ∞, replacing $\hat{z}=1$ because we are evaluating this derivative at $\nu=0$. The second term is always negative because $\frac{\partial\hat{z}}{\partial\nu}>0$. So a sufficient, though not necessary condition is a parameter restrictions that implies. Recall, the employment level and Ψ distributions are simply functions of parameters, and not endogenously determined at $\nu=0$, therefore this can be written as a restriction of γ^1 being large enough relative to a function of δ,γ^0 and α: we need job-to-job finding rates to be high enough relative to the size of the unemployment pool from which $\underline{w}$ firms may recruit. With our power-law distribution, this condition becomes true if $\gamma^1(1-\gamma^1)\frac{\alpha-1}{\alpha+2}\geq\frac{\delta}{\gamma^0+\delta}$. Note, however, that if $\frac{\delta}{\gamma^0+\delta}$ is large, such that the sufficient condition is not satisfied, the second line of Equation 33 becomes more negative. □

Lemma 8. *The direct effect of ν on labor at a top-wage firm is positive,* $\left.\frac{\partial L(\bar{w})}{\partial\nu}\right|_{\nu=0}>0$

Proof. Labor of each type at a top firm is given by

$$\ell(\bar{w},z)\delta = \frac{\Omega(z)}{M}((1-n(z))\gamma^0+n(z)\gamma^1)+\nu\tilde{\Psi}(z)\int_1^{\infty}\ell(\bar{w},t)tdt$$

Integrating over types z, and taking the partial with respect to ν we have

$$\frac{\partial L(\bar{w})}{\partial\nu} = \frac{1}{\delta}\int_1^{\infty}\tilde{\Psi}(z)\int_1^{\infty}\ell(\bar{w},t)tdtdz$$

This is clearly positive because every term over which we are integrating must be positive. □

Lemma 9. *The highest offered wage is increasing in ν from $\nu=0$,* $\left.\frac{\partial\bar{w}}{\partial\nu}\right|_{\nu=0}>0$

Proof. The highest wage $\bar{w}$ must satisfy profit equalization $\pi = (1-\bar{w})L(\bar{w}) = (1-\underline{w})L(\underline{w})$. From Lemma 7 we have that profit is decreasing in ν from $\nu = 0$. If $\frac{\partial \pi}{\partial \nu} < 0$ and $\frac{\partial L(\bar{w})}{\partial \nu} > 0$, then for profit equalization to hold it must be the case that $\bar{w}$ increases.

□

We can now prove the main Proposition:

Proposition. *4:*
Evaluated at $\nu = 0$,

- *The interval of wages gets wider, $\frac{\partial \underline{w}}{\partial \nu} \leq 0$, $\frac{\partial \bar{w}}{\partial \nu} \geq 0$*
- *Firms at the lowest wages get smaller and at the highest wages they get larger, $\frac{dL(\underline{w})}{d\nu} \leq 0$ and $\frac{dL(\bar{w})}{d\nu} \geq 0$*
- *Profit declines $\frac{d\pi}{d\nu} \leq 0$*

Proof. The first bullet point is because of Lemma 6 and Lemma 9. Taking the second bullet point:

$$\frac{dL(\underline{w})}{d\nu} = \frac{\partial L(\underline{w})}{\partial \nu} + \frac{\partial L(\underline{w})}{\partial w}\frac{\partial \underline{w}}{\partial nu}$$

The first term is negative because of Lemma 7, Equation 33, the second term is negative because $\frac{\partial L(\underline{w})}{\partial w} > 0$ and Lemma 6. While

$$\frac{dL(\bar{w})}{d\nu} = \frac{\partial L(\bar{w})}{\partial \nu} + \frac{\partial L(\bar{w})}{\partial w}\frac{\partial \underline{w}}{\partial nu}$$

are all positive terms, the first because of Lemma 8 and the second because of Lemma 9. Finally, profit decline is shown in Lemma 7.

□

B Network search data from the Survey of Consumer Expectations

To select our sample, we require that workers report the search method that led to their current job, as well as the salary associated with it and employment characteristics we use as regressors. We allow either full or part-time employment but restrict to those making at least \$80 per week. We drop those who did not transition from prior employment or unemployment. Because workers may report their current and past salary either as annual, weekly or hourly, we convert all to weekly salaries. To convert hourly wages, we multiply by reported usual hours and divide annual salaries by 52 weeks. Workers do not directly report their usual hours in their prior employment, instead they list the change from their current job, which could, of course, introduce some measurement error. Unfortunately, the public version of the SCE does not include data on worker's demographic characteristics otherwise we could restrict ourselves to a homogeneous sample or statistically control these differences.

In our regressions we control for firm size and industry effects. We use the size bundles available in the SCE and collapse these 2-digit industry codes into approximately 1-digit bins combining industries 1-5, 6-8, 10-14, 16-19 and another with 15 and 9. We categorize a "voluntary job-to-job" transition if there was no interstitial unemployment spell and the separation did not occur because the original job ended. We follow the SCE in defining part-time and full-time work.

	Total	Network	Direct
Wkly Salary	\$1,334.20	\$1,414.73	\$1,309.30
	(\$38.83)	(\$69.93)	(\$46.00)
Usual Hours	41.26	40.55	41.48
	(0.29)	(0.60)	(0.33)
Prior Wkly Salary	\$1,124.42	\$1,309.19	\$1,067.28
	(\$85.96)	(\$234.70)	(\$86.01)
Num Employees	2.83	2.79	2.84
	(0.04)	(0.08)	(0.05)
Voluntary Sep	0.59	0.66	0.57
	(0.01)	(0.03)	(0.02)
Obs	1287	304	983

Table IV: Summary statistics for our SCE sample and split between those who found their current job through a network referral, about $\frac{1}{4}$ the sample, and those who did not. Standard errors in parentheses.

Table IV presents summary statistics for our sample. Unconditionally, workers who found their job via networks have generally higher wages and came from higher wages. The most salient other difference in is that they tend more frequently to separate voluntarily.

We next use a regression to confirm that this data is consistent with evidence in Dustmann et al. (2015), that jobs found through referral have higher salaries, even conditional on other factors. As discussed in Section 5, this is an aspect central to the mechanism in our model, in which network-search gives workers access to job offers from a distribution that first-order stochastic dominates the random-search offer distribution. To find empirical evidence, we estimate a logit, predicting whether the current job was reportedly found through a network referral or direct contact search. Formulating the estimation as such allows us to control for firm-effects that we know also affect wages and the probability of hiring through referral. Specifically, we estimate for worker i with employer j

$$\Pr[\text{Network find}] = f(\beta_1 salary_{ij} + \beta_2 small_j + \beta_3 J2J_{ij} + \beta_4 PT_{ij}) \ . \tag{34}$$

Table V presents the estimation results. The number of observations drops slightly from Table IV because of missing industry codes. In the first column of Table V, we show the relationship between a workers earnings and the probability that they found their job through a network referral. Those who earn more tend to have found their job through their network. In Column (2) we substitute the size dummies for a single dummy representing those with fewer than 500 employees to illustrate that smaller firms tend to hire through referrals, confirming past findings. In Columns (3) and (4) we show that the salary estimates are robust to controlling for whether the transition was directly job-to-job and whether it is part or full time. The estimates of these ancillary coefficients are rather imprecise but the signs are as expected.

	(1)	(2)	(3)	(4)
Log weekly salary	0.0513***	0.0484***	0.0413**	0.0498**
	(2.85)	(2.75)	(2.22)	(2.55)
<500 employees		0.0132		
		(0.42)		
Voluntary Job-to-Job			0.0656**	0.0677**
			(2.32)	(2.40)
Part-time				0.0613
				(1.21)
Firm Size Dummies	X		X	X
Industry Dummies	X	X	X	X
Observations	933	934	933	933

Table V: Estimated marginal effects from equation 34