

Directed Search On the Job and the Wage Ladder*

Alain Delacroix
Department of Economics
Purdue University
(delacroixa@mgmt.purdue.edu)

Shouyong Shi
Department of Economics
University of Toronto
(shouyong@economics.utoronto.ca)

2003

Abstract

In this paper we characterize the equilibrium in a labor market where employed workers search on the job and firms direct the search by announcing wages and employment probabilities for the applicants. All workers/jobs are homogeneous and free entry of firms determines the number of jobs. The equilibrium features a wage ladder, with a finite number of rungs. Workers on each particular rung of the ladder choose (optimally) to apply to only the jobs at one level above their current wage, despite that they observe all higher wage offers. Workers choose not to leap several rungs at a time on the wage ladder because the jobs at one level above their current wage provide a significantly higher employment probability, and hence a higher expected surplus, than the jobs at two or more levels above. The wage ladder has the following properties: (i) The gap between two adjacent rungs on the ladder becomes smaller and smaller as wage increases; (ii) A worker's quit rate decreases with wage; (iii) A worker's wage, on average, increases with the employment duration; (iv) The average length of time an unemployed worker will take to return to his previous wage increases with that wage; (v) The density of offer wages decreases with wage; (vi) The density of employed wages can be decreasing, increasing, or hump-shaped. The directed search framework replicates empirical regularities on the wage path of workers and the distribution of offer and employed wages that undirected search cannot.

Keywords: Directed Search, On-the-Job Search, Wage Ladder.

* Corresponding author: Shouyong Shi, Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, Canada, M5S 3G7; (email: shouyong@economics.utoronto.ca); telephone: 416-978-4978; fax: 416-978-6713. We would like to thank Dale Mortensen and Rob Shimer for comments. We also received useful comments from the participants of Minnesota Summer Workshop on Macroeconomics (Minneapolis, 2002), National Bureau of Economic Research Summer Workshop (Cambridge, 2002), Canadian Theory Meeting (Toronto, 2002), Midwest Theory Meeting (Notre Dame, 2002), and the workshops at Northwestern University (2002) and Purdue University (2002). Shi would like to acknowledge the Social Sciences and Humanities Research Council of Canada for financial support. All errors are ours alone.

1. Introduction

We study a large labor market where employed workers search on the job and firms direct the search process intentionally. All workers are identical. The number of workers is large and fixed, while the number of jobs (also identical) is determined by free entry. All firms simultaneously announce a wage level and an employment probability for an applicant. The applicants, including unemployed ones, first observe all offers and then decide which job to apply to. After receiving the applicants, a firm selects one and pays the announced wage. Search is *directed*, in the sense that firms explicitly take into account how their offers will affect workers' application strategies. The game repeats without memory. Recruiting generates endogenous transitions of workers between jobs, while exogenous separation sends workers into unemployment. We characterize the stationary equilibrium in this market and study its properties.

Table 1. Search models

type of search	whether employed workers are allowed to search on the job	
	no	yes
undirected	Diamond (1982) Mortensen (1982) Pissarides (1990)	Burdett and Judd (1983) Burdett and Mortensen (1998) Pissarides (1994)
directed	Peters (1991) Montgomery (1991) Acemoglu and Shimer (1999) Burdett et al. (2001)	This paper

Directed search on the job is a realistic feature of the labor market.¹ However, the search literature (see Table 1 for a rough guide) offers very little knowledge about the equilibrium with directed search on the job. The popular search models assume that search is *undirected*, e.g., Diamond (1982), Mortensen (1982) and Pissarides (1990). That is, firms ignore how their offers affect the number of matches they will receive and workers do not choose which job to apply to ex ante, although workers do choose whether to accept a job ex post. These popular models also rule out on-the-job search by assuming that unemployed workers are the only workers who search. The subsequent research has relaxed these two assumptions separately but not simultaneously. For example, Burdett and Mortensen (1998) and Pissarides (1994) examine undirected search on the job, while Peters (1991), Montgomery (1991), Acemoglu and Shimer (1999) and Burdett et al. (2001) examine directed search without on-the-job search.² Our focus in this paper will

¹Reviewing the evidence by Blanchard and Diamond (1989), Pissarides (1994) concludes that about 20% of the monthly hires in the US are direct job-to-job movements. He argues that this fraction is about 40% in UK.

²Burdett and Judd (1983) analyzed sequential asearch in the goods market and did not frame their model as on-the-job search. However, the key element in their model, that some agents each have two or more price quotes while others have only one, is borrowed by Burdett and Mortensen (1998) to generate wage dispersion with on-the-job search. Similarly, Peters (1991) analyzed directed search in the goods market, but his analysis can be

be on the characterization of the equilibrium with the combined feature of directed search and on-the-job search.

Our study is also motivated by the fact that existing search models fail to capture some important aspects of the wage distribution among homogeneous workers with homogeneous jobs. First, with undirected search on the job, the density of the wage distribution is a strictly increasing function, but the empirical density is hump-shaped with the hump occurring at a low wage level (see Kiefer and Neumann (1993)).³ Second, workers seem to climb up the wage ladder over time in reality, which existing search models fail to produce. In particular, wage dispersion in models with undirected search on the job does not constitute a wage ladder, because all workers can ascend immediately to the end of the wage spectrum with positive probability.

Of course, one can introduce heterogeneity among workers or jobs into existing models to make their predictions more realistic. To generate the hump-shaped wage distribution, for example, one may argue that workers differ in ability or that jobs differ in productivity. To generate gradual wage increases, one may argue that workers' abilities are gradually observed by their employers (e.g., Jovanovic (1979)), or there is match-specific productivity, or there is learning-by-doing on the job. These *ex ante* and *ex post* sources of heterogeneity are realistic, and our work does not diminish their importance. However, because there is sizable dispersion of wages in the data among similar workers with similar jobs, it is useful to explore theoretical models that generate wage dispersion among homogeneous workers, as we do here.

We show that the equilibrium with directed search on the job is a wage ladder, which comprises of a finite number of wage levels. Firms are indifferent between posting these equilibrium wages. A firm offering a higher wage is compensated by a higher probability of hiring a worker successfully. The expected surplus of a recruiting firm (i.e., the product of the hiring probability and the *ex post* surplus) is the same at all equilibrium wages. This expected surplus exactly covers the vacancy cost, and so all recruiting firms earn zero net expected profit. In contrast, workers are *not* indifferent between the different wages in equilibrium. An applicant chooses to apply to only such jobs that offer him the maximum expected surplus (i.e., the employment probability times the *ex post* gain from the job). These jobs lie one rung above the applicant's current wage on the ladder. That is, workers choose to climb up the wage ladder over time, one rung at a time, rather than leap on the ladder.

The wage ladder has strong implications on equilibrium wages and worker flows. First, the gap between two adjacent rungs on the ladder becomes smaller and smaller as wage increases, and so the marginal gain from climbing up on the ladder diminishes. Second, a worker's quit rate readily adapted to the labor market. Other examples of directed search models are Cao and Shi (2000), Julien et al (2000), Peters (2000), Shi (2001, 2002a,b), and Shimer (2001).

³In directed search models without on-the-job search, a wage differential can arise among homogeneous workers who work for homogeneous jobs, but some additional elements must be introduced. In Shi (2002a), firms are different in size (i.e., the number of employees), where large firms pay higher wages than small firms do. In Shi (2002b) and Shimer (2001), the presence of high-skill workers induces partial sorting and generates a wage differential among low-skill workers.

decreases with wage, because the probability of getting a higher wage falls as the worker climbs up the ladder. Third, a worker's wage, on average, increases with his employment duration, because a longer duration is likely to indicate that the worker has climbed more rungs on the wage ladder. Fourth, the average length of time an unemployed worker will take to return to his previous wage increases with that wage. Fifth, the density of offer wages is a strictly decreasing function; i.e., there are more firms recruiting at a low wage than at a high wage. This occurs despite the fact that a high wage makes hiring more successful than a low wage. Finally, the density of employed wages can decrease with wage, but not always so. Although there are more vacancies at a low wage than at a high wage, which generates a large flow of workers into a low wage, there are also more workers quitting low-wage jobs. Depending on which of these two flows dominates, the density of employed wages can be decreasing, increasing, or hump-shaped. In particular, when a low-wage job has a sufficiently higher inflow of workers than a high-wage job, the density of employed wages is a decreasing function.⁴

The wage ladder and its properties occur here without any of the familiar assumptions that induce gradual wage increases. In particular, (i) there is no gradual revelation of workers' productivity, learning-by-doing or match-specific productivity, as all workers have the same productivity which is observable before match; (ii) there is no differential information among employed workers regarding job openings, as all employed applicants observe all job openings before they apply; and (iii) firms do not discriminate employed workers according to their current wages, as each firm selects all applicants it receives with the same probability in equilibrium. Rather, a worker chooses to apply to only the jobs one level about his current wage because such jobs offer a significantly higher employment probability, and hence higher expected surplus for the applicant, than other jobs do.

To understand this result better, imagine two applicants who are currently employed at, respectively, a low wage and a high wage. These two applicants differ in the trade-off between the wage and the employment probability, because their current wages are different. A high-wage applicant cares more about the wage level and less about that employment probability than a low-wage applicant does. More precisely, since the same wage offer yields a lower ex post gain to a high-wage applicant than to a low-wage applicant, the same amount of increase in the wage offer represents a larger proportional increase in the expected surplus to a high-wage applicant than to a low-wage applicant. Exploring this difference in the trade-off, firms can separate the groups of applicants by offering high-wage applicants a high wage with a low employment probability and low-wage applicants a low wage with a high employment probability. In a stationary equilibrium, the separation produces the wage ladder. Clearly, directed search and on-the-job search are both important for supporting the wage ladder as an equilibrium.

⁴With the wage ladder, workers employed at low wages do not apply to very high wages. This contributes in part to the decreasing density at high wage levels. As a comparison, the undirected search model of Burdett and Mortensen (1998) produces the result that workers at all wage levels apply to the highest wage with positive probability, which generates a much higher density of workers employed at the highest wage than at lower wages.

Before specifying the model, it may be useful to preview the analytical difficulty of combining on-the-job search with directed search. The source of the difficulty is that on-the-job search makes a worker’s application strategy in the future depend on his current application strategy (i.e. where to apply now influences where to apply next once that new wage is obtained). A recruiting firm anticipates and explores this dependence, which makes the equilibrium a complicated fixed-point problem. This complexity does not arise when on-the-job search is excluded, because then all separations are exogenous. Because of the complexity, a large part of this paper will be devoted to characterizing the equilibrium in a tractable way.

2. A Model of Directed Search On the Job

2.1. The Labor Market and Job Search

A labor market is populated by a large (exogenous) number, L , of risk neutral and infinitely-lived workers. All workers are identical. When employed, a worker supplies one unit of labor and produces y units of output per unit of time. When unemployed, a worker receives an unemployment benefit, b . The unemployment rate u is endogenous. For convenience, we refer to a worker’s wage as the worker’s type and call a worker at wage w a w -worker. Also, we refer to b as an unemployed worker’s “wage” and write $w_0 = b$. There are also a large number of firms, determined endogenously by free entry, each of which has one job to offer. All jobs are the same, and the cost of a vacancy per unit of time is $C > 0$. Time is continuous. Workers and firms discount future at a rate of time preference $r > 0$.

Employed workers search on the job. A worker employed at wage w receives an opportunity of job application at rate $\lambda(w) > 0$, according to the Poisson process. The opportunity allows the worker to observe the job descriptions offered by all firms. We set $\lambda(w) = \lambda$ for all employed workers in order to ensure that the difference in λ between workers is not the cause of wage dispersion in this paper. A job description consists of a wage offer and the firm’s selection rule (described later). All recruiting firms announce the job descriptions simultaneously and commit to them. After observing the job descriptions, applicants choose their application strategies simultaneously. Each applicant can apply to only one job and must incur a small fixed cost $S > 0$ for the application.⁵ The application strategy can be mixed over the job openings. Because the applicants observe firms’ offers before the application, a firm can choose the offer intentionally to attract particular applicants, i.e., to direct workers’ search.

Once employed, a worker produces and is paid the posted wage until separation, either exogenously or endogenously. Exogenous separation occurs at rate $\sigma > 0$ according to the Poisson process, after which the worker returns to the unemployment pool. Endogenous separation occurs when a worker quits the current job to accept another job. As in many models of on-the-job search, e.g., Burdett and Mortensen (1998), we assume that a worker’s current employer does

⁵The small cost S is needed here to help the existence of an equilibrium. See subsection 4.2.

not match outside offers. This assumption may not be realistic, but it serves two purposes here. First, it keeps our model comparable to other on-the-job search models. Second, it makes the model tractable and serves as a benchmark.

Unemployed workers' search is directed in the same way as an employed worker's search. Let λ_0 be the arrival rate of a job application opportunity to an unemployed worker. In general, $\lambda_0 \neq \lambda$, but our analysis applies equally well to the special case $\lambda_0 = \lambda$. Allowing for $\lambda_0 \neq \lambda$, we can nest directed search models without on-the-job search as a special case, as discussed immediately below.

The combination of on-the-job search and directed search is a unique feature. It distinguishes our model from directed search models in the literature, such as Acemoglu and Shimer (1999) and Burdett, et al. (2001), which allow only unemployed workers to search. Those models correspond to the special case $\lambda = 0 < \lambda_0$. Another dimension of comparison is with on-the-job search models that assume undirected search, e.g., Burdett and Mortensen (1998). The most important feature of directed search is that each applicant observes many offers before application. The induced ex ante competition between recruiting firms is central to our results.⁶

Nevertheless, our model has the following frictions that are familiar in previous search models. First, job application opportunities are not abundant, in the sense that λ and λ_0 are finite. This is a proxy for the cost of gathering information about jobs. Second, each applicant can apply to only a small number of jobs at a time, which is set to be one in our paper. This is a proxy for the constraint that an applicant can attend only one interview at a time. Third, agents cannot coordinate their decisions, which creates the possibility of unmatched agents.

Let wage levels lie in the set of real numbers, \mathfrak{R} . Let $N(\cdot)$ be the cumulative distribution function of workers over wages, with a density function $n(\cdot)$. Because we classify the unemployment benefit as a "wage" level, the mass of the distribution N at $w_0 = b$ is equal to u . The distributional density of (employed) workers over wages in $\mathfrak{R} \setminus \{b\}$ is $n(\cdot)/(1 - u)$, which is also called the *employed wage density*. Let $V(\cdot)$ be the cumulative distribution function of vacancies over wages. The corresponding density function, denoted $v(\cdot)$, is called the *offer wage density*. Let K be the total number of vacancies, which is endogenous, and denote $k = K/L$. The number of vacancies at wage w is $v(w)K$.

We are interested in the equilibrium in a large market, i.e., one in which $L \rightarrow \infty$. However, to detail agents' strategies, we will first analyze a market where L is large and finite, and then take the limit $L \rightarrow \infty$.⁷ The expected number of w -applicants is $\lambda n(w)L$ (or $\lambda_0 uL$ if $w = w_0$). Assume that this number is an integer, without loss of generality.

⁶The assumption that each applicant observes all firms' offers before applying to one simplifies the analysis, but it is not necessary for the analytical results. For ex ante competition to occur, it is sufficient to assume that each applicant observes two offers that are randomly drawn from all recruiting firms' offers.

⁷We view continuous wage distributions as the limits of discrete distributions. This procedure avoids possible measurability problems associated with analyzing individuals' strategies directly in an environment with a continuous wage distribution.

2.2. Strategies, Payoffs and Decisions

Call a w -worker who has just received a job application opportunity a w -applicant. If $w = w_0$, the applicant is unemployed. To unify the notation for employed and unemployed workers, let $\lambda(w)$ be the arrival rate of a job application to a w -worker, where $\lambda(w) = \lambda$ for all $w \neq w_0$ and $\lambda(w_0) = \lambda_0$.

A w -applicant's strategy is an application probability function $p(\cdot, w): \mathfrak{R} \rightarrow [0, 1]$, where $p(w', w)$ is the probability with which the applicant applies to *each* of the job openings at wage w' . An implicit assumption is that the applicant must assign equal probability to apply to all identical job openings. Also, all applicants of the same type use the same strategy. Denote $T(w) = \{w' : p(w', w) > 0\}$, which is a w -applicant's target set of wages.

The probability $p(\cdot, w)$ is not a convenient object to use in a large market. Because there will be many firms offering each wage, an applicant will apply to each of these firms with a very small probability. That is, $p(w', w) \rightarrow 0$ for all w' . To find a convenient object to describe an applicant's strategy, notice that the expected number of w -applicants whom a firm offering w' will receive is:

$$a(w', w) = p(w', w)\lambda(w)n(w)L. \quad (2.1)$$

Recall that $\lambda(w)n(w)L$ is the expected number of w -applicants. Term $a(w', w)$ the *queue length* of w -applicants for a firm that posts wage w' . Because $a(w', w) > 0$ if and only if $p(w', w) > 0$, the queue length represents an applicant's strategy well. Moreover, if $p(w', w) > 0$ in a finite market, $a(w', w)$ remains strictly positive even when the market becomes large. For these reasons, we describe a w -applicant's strategy as a function $a(\cdot, w): \mathfrak{R} \rightarrow \mathfrak{R}_+$. Denote $A(w) = \{a(w', w)\}_{w' \in \mathfrak{R}}$. Then, a w -applicant's target set is $T(w) = \{w' : a(w', w) > 0\}$. Because the application probabilities $p(w', w)$ must sum to one over w' , the following constraint must hold:

$$\sum_{w'} a(w', w)v(w') = \lambda(w)n(w)/k.$$

A recruiting firm announces a job description, which consists of a wage level $w \in \mathfrak{R}$ and a criterion that determines which applicant will get the job. This ex post selection criterion depends on the distribution of applicants whom the firm will receive or attract. Because the applicants use mixed strategies to apply to identical jobs, the distribution of applicants whom the firm will receive is stochastic. For each realization of such random variables, the selection criterion must be specified. This is cumbersome but, fortunately, it is not necessary. For the decision problems, it suffices for the firm to specify an "ex ante" employment probability, denoted $q(w, w')$, for each applicant of type w' (see the supplementary Appendix G for a proof). The probability $q(w, w')$ is the probability with which a w' -applicant gets a job w when he applies to the job, and it is calculated by aggregating ex post employment probabilities over all realizations of the distribution of applicants whom the firm will receive. We refer to q simply as the *employment probability*.

It is optimal for a firm to give the same employment probability to all applicants. That is, $q(w, w') = q(w)$ for all w' (see the supplementary Appendix G for a proof). This is because the

firm's probability of successfully hiring a worker is a concave function of the expected number of applicants whom the firm will attract (see (2.3) later). By giving equal employment probability to all applicants, the firm maximizes the hiring probability and, because all workers have the same productivity, the strategy maximizes the expected surplus from recruiting.⁸ Therefore, a recruiting firm's strategy is to announce a wage level $w \in \mathfrak{R}$ and an employment probability $q(w) \in [0, 1]$ for every applicant. Let $Q = \{q(w)\}_{w \in \mathfrak{R}}$ be a profile of employment probabilities across offer wages.

To describe the agents' payoffs, let $J_e(w)$ be the value function of a worker who is currently employed at wage w , and $J_u = J_e(w_0)$ the value function of an unemployed worker. For firms, let J_v be the value function of a vacancy and $J_f(w)$ of a firm that currently employs a worker at w . We will calculate these value functions later.

A w' -applicant's payoff of applying for a job w is the expected surplus, $q(w)[J_e(w) - J_e(w')]$. This applicant chooses $a(w, w') > 0$ if and only if the expected surplus is greater than or equal to the maximum of expected surplus that the applicant can get elsewhere. Call this maximum expected surplus the applicant's *market surplus* and denote it $E(w')$. In a large market, individual agents take such market surpluses as given.⁹ Then, a w' -applicant's optimal choice of $a(w, w')$ satisfies the following condition:

$$a(w, w') = \begin{cases} = \infty, & \text{if } q(w)[J_e(w) - J_e(w')] > E(w') \\ \in [0, \infty), & \text{if } q(w)[J_e(w) - J_e(w')] = E(w') \\ = 0, & \text{otherwise.} \end{cases} \quad (2.2)$$

In the first case, the queue length of w' -applicants for a firm offering w is infinite because every w' -applicant applies to the firm with probability one and there are a large number of such applicants. In the last case, the queue length of w' -applicants for a firm is zero because such applicants have no incentive to apply to the firm. These two cases will not arise when w is an equilibrium wage.¹⁰ Thus, if w is an equilibrium wage, the expected surplus of applying to a firm posting w must be equal to the applicant's market surplus.

The payoff of a firm posting $(w, q(w))$ is the expected surplus, $h[J_f(w) - J_v]$, where h denotes the firm's probability of successfully hiring a worker. This hiring probability is a function of both w and $q(w)$, but we shorten it to $h(w)$. Because a firm fails to recruit a worker only when all potential applicants have applied to other firms, the hiring probability is:

$$h(w) = 1 - \prod_{w'} [1 - p(w, w')]^{\lambda(w')n(w')L}.$$

⁸When workers are heterogeneous in productivity, recruiting firms may give different priorities to different workers, see Shi (2002b) and Shimer (2001).

⁹Peters (2000), Cao and Shi (2000), and Burdett, et al. (2001) analyze directed search without imposing this assumption. They show that the effect of each individual agent's actions on the worker's market surplus becomes negligible when the size of the market approaches infinity.

¹⁰If $a(w, w') = \infty$, then each w' -applicant who applies to a firm offering w is chosen with probability 0. The expected surplus is 0 in this case, and hence cannot be higher than the applicant's market surplus, contradicting the condition required for $a(w, w') = \infty$. On the other hand, if $a(w, w') = 0$, then the firm offering w does not attract any applicant, and so w is not an equilibrium wage.

Using the queue length defined in (2.1) and the fact that $(1-p)^{1/p} \rightarrow e^{-1}$ when $p \rightarrow 0$, we have:

$$h(w) = 1 - \exp \left[- \sum_{w'} a(w, w') \right]. \quad (2.3)$$

Thus, a firm's hiring probability increases with the queue length of applicants whom the firm will attract. Moreover, the hiring probability is a concave function of the queue length.

To find the link between the hiring probability and the employment probability, notice that the hiring probability can also be interpreted as the expected number of workers that the firm hires, because the firm hires only one worker for a job. This interpretation implies the relationship $h(w) = q(w) \sum_{w'} a(w, w')$. Combining this with the above formula, we have the following negative relationship between the employment probability and the queue length:

$$q(w) = \left[1 - \exp \left(- \sum_{w'} a(w, w') \right) \right] / \sum_{w'} a(w, w'). \quad (2.4)$$

Hence, q and h have the following relationship:

$$q(w) = - \frac{h(w)}{\ln [1 - h(w)]} \equiv \Psi(h(w)). \quad (2.5)$$

Let Ψ^{-1} be the inverse of Ψ , so that $h(w) = \Psi^{-1}(q(w))$. The function Ψ is decreasing because a firm is able to provide a high employment probability to applicants only when the firm will receive a small number of applicants, in which case the firm's hiring probability must be low.

We can formulate a recruiting firm's decisions $(w, q(w))$ as follows:

$$(\mathcal{P}) \quad \max \Psi^{-1}(q(w)) [J_f(w) - J_v]$$

subject to

$$q(w) [J_e(w) - J_e(w')] \geq E(w'), \text{ for all } w' \text{ such that } T(w') \ni w. \quad (2.6)$$

The firm takes other firms' decisions and the applicants' market surpluses as given.

The above formulation captures the key feature of directed search – the trade-off between wage and the probability of forming a match. For a recruiting firm, a higher wage offer is likely to yield a lower value J_f , but it is also likely to attract more applicants and increase the hiring probability. The firm chooses such an offer $(w, q(w))$ that maximizes its expected surplus. For an applicant, a higher wage offer is likely to yield a higher value of employment J_e , but it is also likely to come with a lower employment probability. The applicant applies only to those firms whose offer $(w, q(w))$ maximizes the applicant's expected surplus, as the dual to (\mathcal{P}) suggests.

Now we specify the value functions and the endogenous separation probability. Because an employee quits only when he gets a job from another firm, the endogenous separation probability of a worker at wage w is:

$$\rho(w) = \lambda(w) \sum_{w'} q(w') p(w', w) v(w') K = \frac{k}{n(w)} \sum_{w'} q(w') a(w', w) v(w'). \quad (2.7)$$

The value functions are given as follows:

$$rJ_v = -C + h(w)[J_f(w) - J_v], \quad (2.8)$$

$$rJ_f(w) = (y - w) - [\sigma + \rho(w)][J_f(w) - J_v], \quad (2.9)$$

$$rJ_e(w) = \left[\begin{array}{l} w - \sigma [J_e(w) - J_u] - \chi(\sum_{w'} a(w', w)) \lambda S \\ + \frac{k}{n(w)} \sum_{w'} q(w') [J_e(w') - J_e(w)] a(w', w) v(w'), \end{array} \right] \quad (2.10)$$

$$rJ_u = b - \lambda_0 S + \frac{k}{u} \sum_{w'} q(w') [J_e(w') - J_u] a(w', w) v(w'). \quad (2.11)$$

The function $\chi(\cdot)$ in (2.10) is an indicator function, with $\chi(\Sigma a) = 1$ if $\Sigma a > 0$ and $\chi(\Sigma a) = 0$ if $\Sigma a = 0$.

We explain (2.10) for example. Eq. (2.10) equates the permanent income of a worker employed at wage w , $rJ_e(w)$, to the expected “cash flow” in such employment. The cash flow consists of the current wage, the loss in value in the event of exogenous separation, and the expected gain from searching on the job. The gain from searching on the job is the difference between the last two terms in (2.10). If this difference is non-positive, the worker will choose $\Sigma a = 0$, in which case the last two terms in the equation are zero.

2.3. Definition of Equilibrium

As formulated above, each firm can choose to post any number $w \in \mathfrak{R}$ as the wage level. However, the set of equilibrium wage offers is $\Omega \equiv \{w \in \mathfrak{R} : v(w) > 0\}$. Define $w_1 = \inf(\Omega)$ and $w_M = \sup(\Omega)$. Let $\Omega_0 = \Omega \cup \{w_0\}$ and call Ω_0 the extended support of equilibrium wages. Clearly, the accumulative distribution of wages over Ω_0 is $N(\cdot)$; the density of employed workers over Ω is $n(\cdot)/(1 - u)$, and the density of vacancies (or offer wages) over Ω is $v(\cdot)$.

Definition 2.1. *A (symmetric) equilibrium in the labor market consists of the aggregate characteristics $(\Omega, N(\cdot), V(\cdot), k)$, the profile of employment probabilities $Q = (q(w))_w$, and the applicants’ strategies $(A(w))_w$, where $A(w) = (a(w', w))_{w'}$, such that the following requirements are met: (i) Given the aggregate characteristics and other firms’ strategies, each firm’s decision $(w, q(w))$ is optimal, and $w \in \Omega$; (ii) Given the firms’ decisions and the aggregate characteristics, each applicant’s application decision is optimal; (iii) The agents’ strategies are symmetric within each type, in the sense described earlier; (iv) There is free entry of firms: $J_v = 0$ for every recruiting firm; (v) The aggregate characteristics are stationary.*

In the symmetric equilibrium, $p(w', w) = 1/[v(w')K]$ if $w' \in T(w)$ and 0 otherwise. So, $a(w', w) = \lambda(w)n(w)/[v(w')k]$ if $w' \in T(w)$ and 0 otherwise. Also, as required by the equilibrium, we will set $J_v = 0$ in the remainder of this paper.

3. Configuration of the Equilibrium

3.1. Separation of Applicants by Their Current Wages

We establish the following lemma that will be central to our analysis.

Lemma 3.1. (Singleton) *If there is an equilibrium, then each equilibrium wage attracts at most one type of applicants. Precisely, $a(w^*, w_i)a(w^*, w_j) = 0$ for all $w_i, w_j, w^* \in \Omega_0$ with $w^* > w_j > w_i$, provided that $J_e(\cdot)$ is an increasing function.*

Proof. Because the lemma is trivially true for $w^* = w_1$ ($= \inf \Omega$), we examine a firm posting $w^* > w_1$. The decision problem is (\mathcal{P}) , with w being replaced by w^* , $q(w)$ by $q(w^*)$, etc.. Shorten the notation $q(w^*)$ to q^* . Suppose that (w^*, q^*) is an equilibrium offer. As concluded earlier, the constraint (2.6) must hold as equality for all types of applicants whom the firm attracts. Moreover, the constraint (2.6) must be binding on the firm for at least one type of applicants whom the firm attracts; otherwise, the firm should set $q^* = 0$ to maximize the hiring probability, which contradicts a non-binding constraint.

Suppose that the applicants' constraint binds for $w' = w_i$, with $w_i < w_j$. Consider an alternative offer (\hat{w}, \hat{q}) , where $\hat{w} = w^* + \varepsilon$, $\hat{q}[J_e(\hat{w}) - J_e(w_j)] = E(w_j) = q^*[J_e(w^*) - J_e(w_j)]$, and $\varepsilon > 0$ is an arbitrarily small number. Under the presumption that $J_e(\cdot)$ is an increasing function (verified later), $\hat{q} < q^*$. Then,

$$\begin{aligned} \hat{q}[J_e(\hat{w}) - J_e(w_i)] &= q^*[J_e(w^*) - J_e(w_i)] - (q^* - \hat{q})[J_e(w_j) - J_e(w_i)] \\ &< q^*[J_e(w^*) - J_e(w_i)] = E(w_i). \end{aligned}$$

The inequality follows from the facts that $q > \hat{q}$ and $J_e(w_j) > J_e(w_i)$. Thus, the new offer (\hat{w}, \hat{q}) still attracts w_j -applicants but not w_i -applicants. Because this eliminates a binding constraint on the firm's decision problem with very little change in the offer, the firm's expected surplus increases, contradicting the supposition that (w^*, q^*) is an equilibrium offer.

The proof is similar if the applicant's constraint binds for $w' = w_j$ under the offer (w^*, q^*) . In this case, construct the alternative offer by setting $\hat{w} = w^* - \varepsilon$ and $\hat{q}[J_e(\hat{w}) - J_e(w_i)] = E(w_i)$. This alternative offer attracts w_i -applicants but not w_j -applicants, and it increases the firm's expected surplus. **QED**

The above lemma holds because the indifference curves of any two types of applicants have the following single-crossing property.¹¹ The higher an applicant's type (i.e., his current wage), the lower the ex post surplus he can obtain from a given wage; so, a given amount of increase in the wage offer represents a larger *proportional* increase in the expected surplus to a high-type applicant than to a low-type applicant. Put differently, a high-type applicant cares more about the wage level that an offer provides, and less about the employment probability, than a low-type

¹¹The required condition that $J_e(w)$ increases with w holds in all previous search models and, as verified later, it holds in our model as well.

applicant does. Thus, there is no combination of the employment probability and wage that is most attractive to both high- and low-type applicants and yet maximizes the firm's expected surplus from recruiting. If any symmetric equilibrium exists, each equilibrium wage must attract only one type of applicants.

We depict the single-crossing property in Figure 1. An applicant's indifference curve is derived from the equality form of (2.6):

$$q(w) = \frac{E(w')}{J_e(w) - J_e(w')}, \text{ all } w' \text{ such that } a(w, w') > 0. \quad (3.1)$$

This is a negative relationship between the employment probability and wage, as a low wage must be compensated by a high employment probability in order to induce an applicant to apply. Figure 1 depicts two such indifference curves, one for the applicants of type $w' = w_i$ and the other for $w' = w_j > w_i$. The indifference curve of the high-wage applicant (w_j -applicant) crosses that of the low-wage applicant from above, which reflects the difference discussed above between the two applicants in the willingness to trade-off the employment probability for wage. In Figure 1, we also depict the iso-profit curve in Figure 1, which summarizes different combinations of (q, w) that yield the same level of expected surplus to the recruiting firm.

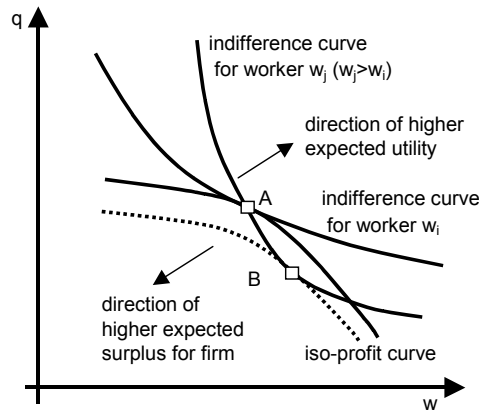


Figure 1

Figure 1 provides the following illustration of Lemma Singleton. If the firm attracts both types of applicants, the firm's offer must be at the intersection of these two indifference curves (point A). For this offer to be optimal for the firm, the firm's iso-profit curve must be tangent to the applicants' indifference curves. This is not possible. If the firm's iso-profit curve is tangent to w_i -applicants' indifference curve at point A , the firm can increase expected surplus by offering the combination at point B that is attractive to only w_j -applicants. Similarly, if the firm's iso-profit curve is tangent to w_j -applicants' indifference curve at point A , the firm can increase expected surplus by increasing q and reducing w along w_i -applicants' indifference curve.

Lemma Singleton provides a sharp contrast between an equilibrium with directed search on the job and an equilibrium with undirected search on the job. In the undirected search model by Burdett and Mortensen (1998), the equilibrium wage support is a continuum. There, because

of undirected search, each firm receives applicants from all wage levels with positive probability, regardless of the firm's offer. All equilibrium wage offers yield the same expected surplus to the firm. Applicants are *not* indifferent between those wages, but they cannot choose ex ante which wage to apply to. This construct is no longer an equilibrium when search is directed, because applicants will apply only to the openings which yield the highest expected surplus to them.

3.2. Equilibrium Configuration

To construct the equilibrium, we assume that the recruiting firm's decision problem (\mathcal{P}) has a unique solution for each type of applicants.¹² Then, the dual of (\mathcal{P}) also has a unique solution. That is, for each type w of applicants, the target set of wages $T(w)$ is singleton. Without confusion, we now refer to $T(w)$ as this single target wage level rather than the set. Then, the equilibrium must be a wage ladder with a finite number of rungs. The argument proceeds as follows.

First, an employed applicant's target wage is strictly higher than his current wage; i.e., $T(w) > w$ for all $w \neq w_0$. This is because workers employed at all wages receive a job application opportunity at the same rate λ and have the same access to all job openings. By applying to a job that offers the same wage as his current wage, an applicant incurs the fixed application cost without any gain. By applying to a wage lower than his current wage, an applicant loses even more because of the wage reduction.

The same argument applies to an unemployed applicant if $\lambda_0 \geq \lambda$. However, if $\lambda_0 < \lambda$, then it is possible that $T(w_0) < w_0$. In this case, an unemployed applicant may accept a wage lower than the unemployment benefit in an attempt to gain a better access to higher wages.

Second, the path of future equilibrium wages starting from any equilibrium wage w (including w_0) contains only a finite number of wage levels. This path is the sequence $(T^i(w))_{i=1}^j$, where $T^i(w) = T^{i-1}(T(w))$ for all i . The number j is finite because there are costs for firms to maintain a vacancy and for workers to apply for jobs. The difference between any two adjacent wage levels, $T^i(w) - T^{i-1}(w)$, must be bounded below by a strictly positive number in order to cover the fixed application cost. In a finite number of steps, the ascending wage sequence will reach a level at which recruiting yields an expected surplus below the vacancy cost.

Third, every employed wage in the equilibrium can be reached in a finite number of steps from w_0 . That is, for every $w \in \Omega$, there exists a non-negative integer j such that $w = T^j(w_0)$. To see this, suppose that an equilibrium wage $w \in \Omega$ cannot be reached from w_0 . Then we can trace backward to find the source of this wage, using the sequence $\{T^{-i}(w)\}_{i \geq 0}$, where T^{-1} is the inverse function of T . Given the features of T discussed above, T^{-1} is a strictly decreasing function and the difference $[T^{-(i-1)}(w) - T^{-i}(w)]$ is bounded below by a strictly positive amount for any $i \geq 1$. Thus, the descending sequence $\{T^{-i}(w)\}_{i \geq 0}$ reaches a minimum in a finite number of steps, say m . Because w cannot be reached from w_0 in a finite number of steps by the

¹²We cannot verify such uniqueness analytically, but will verify it with numerical examples later.

supposition, $T^{-m}(w) \neq w_0$ and $T^{-m}(w)$ cannot be reached from w_0 in a finite number of steps. In fact, $T^{-m}(w)$ cannot be reached from w_0 at all because, if it can ever be reached from w_0 , the number of steps needed is finite by the previous result. At the wage $T^{-m}(w)$, there is an outflow of workers because of endogenous and exogenous separation, but there is no inflow of workers. The measure of workers employed at this wage must be zero in the stationary equilibrium.

Therefore, the equilibrium is a wage ladder that has a finite number of rungs. Denote $w_i = T^i(w_0)$. The set of employed wages in equilibrium is $\Omega = (w_i)_{i=1}^M$, where M is a finite integer and $w_i > w_{i-1}$ for all $i = 2, 3, \dots, M$.¹³ If $\lambda_0 \geq \lambda$, then $w_1 > w_0$; if $\lambda_0 < \lambda$, then $w_1 < w_0$ is possible.

The wage ladder may contain only one wage level or no wage at all. This can happen if, for example, the unemployment benefit (w_0) is sufficiently close to productivity (y). However, when y is sufficiently higher than w_0 , adding even a small amount of on-the-job search (i.e. λ positive, but very close to 0) generates a non-degenerate wage distribution Ω (i.e. $M > 1$). Since the case $\lambda = 0$ corresponds to previous directed search models, those models' prediction that the wage distribution is degenerate is not robust to the introduction of on-the-job search.

We depict the wage ladder in Figure 2a. Each wage w_i attracts only the w_{i-1} -applicants, as pictured by the arrows. Exogenous separation, which takes the separated workers to unemployment, is not depicted here.

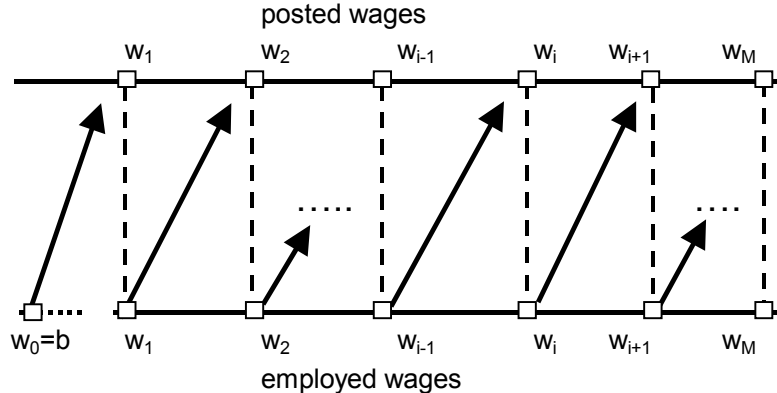


Figure 2a

With the wage ladder as the equilibrium, an applicant applies only to the wage one rung above his current wage, and so the quit rate of a worker employed at w is $\rho(w) = \lambda(w)q(T(w))$. The applicant does not apply to wages two or more rungs above because it is not optimal to do so, not because he does not have access to those higher wages. The lower employment probability accompanying these higher wages make these wages an inferior tradeoff between wage and the

¹³Each equilibrium wage on the ladder is necessarily a mass point of the wage distribution. In contrast, a mass point cannot exist in on-the-job search models where search is undirected, e.g., Burdett and Mortensen (1998). There, the argument is that a firm can increase the wage offer slightly above the mass point, which increases the number of applicants the firm gets by a discrete amount (as a result of undirected search) and hence increases expected profit. This argument becomes invalid when search is directed, because it is not optimal for the applicants to respond to a marginal increase in the wage offer by a discrete increase in the application probability – doing so by a large number of applicants would reduce their employment probability by a discrete amount.

employment probability.

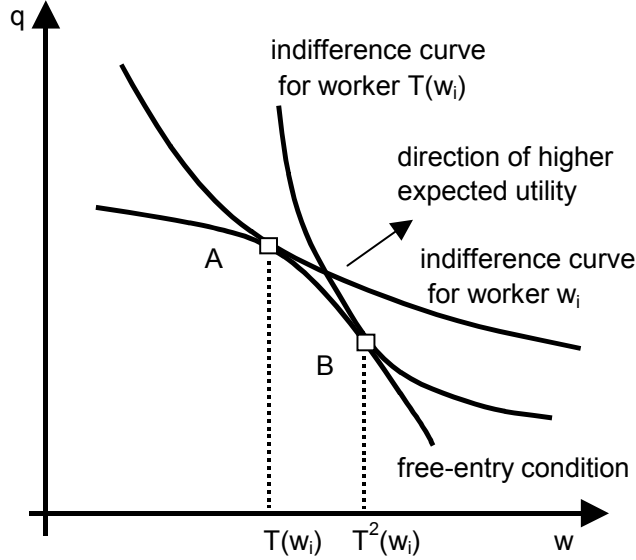


Figure 2b

We can illustrate this tradeoff in equilibrium by a diagram similar to Figure 1. However, the iso-profit curve must be replaced by firms' free-entry condition, $h(w)[J_f(w) - J_v] = C$. Substituting J_f from (2.9), $\rho(w) = \lambda(w)q(T(w))$ and $J_v = 0$, we can write this condition as:

$$q(w) = \Psi \left(\frac{r + \sigma + \lambda(w)q(T(w))}{y - w} C \right), \quad (3.2)$$

where Ψ is a decreasing function defined in (2.5). Figure 2b depicts a possible case of this free-entry condition, in which the condition gives q as a decreasing and concave function of w . Every equilibrium wage level is a tangency point between the indifference curve of a particular type of applicants and the free-entry condition. Figure 2b shows two such wages, $T(w_i)$ for the applicants at wage w_i and $T^2(w_i)$ for the applicants at wage $T(w_i)$. For a w_i -applicant, a job opening at wage $T^2(w_i)$ is inferior to an opening at $T(w_i)$ because the former offers a much lower employment probability.

The tradeoff between wage and the employment probability is common in all directed search models, with or without on-the-job search. The difference is that on-the-job search makes an agent's payoff dependent on future strategies, which significantly complicates the task of finding an equilibrium. In particular, the firm's value function $J_f(w)$ depends on the employee's future application decision, $T(w)$, which in turn makes the free-entry condition for firms offering w dependent on this future application decision. Similarly, an applicant's indifference curve depends on the future application decision through the worker's value function $J_e(w)$. Such dependence makes it difficult to verify the properties of the curves presumed in Figure 2b. A general translation of the diagram into a mathematical formulation yields very little about the analytical properties of the equilibrium (see the supplementary Appendix H).

In the few sections to follow, we will explore the recursive structure of the wage ladder to find the analytical properties of the equilibrium. By focusing on the recursive approach, we are able to construct the equilibrium and in the process to provide some intuition. To do so, we need to show that the configuration in Figure 2a leaves no incentive for firms and applicants to deviate. Because one firm's deviation to a wage outside the equilibrium set sends its potential employee off the equilibrium path for the application for higher wages in the future, we need to specify what firms would do to the applicants whose wages happen to lie outside Ω .¹⁴ Because it is optimal for a firm to offer equal probability to all applicants, as discussed before, we impose throughout the paper the following restriction on beliefs off the equilibrium path:

Restriction (Off-eqm): The employment probabilities satisfy $q(w, w') = q(w)$ for all applicants' types $w' \in \mathfrak{R}$ and for all wage offers w , including $w' \notin \Omega_0$ and $w \notin \Omega$.

Before proceeding further, we simplify the notation. Denote $n_i = n(w_i)$ and $v_i = v(w_i)$ for each $w_i \in \Omega$, with $n_0 = u$, and $a_i = a(w_i, w_{i-1})$. Writing q and h as functions of a , rather than of w , we transform (2.3) and (2.4) as follows:

$$h_i = h(a_i) \equiv 1 - e^{-a_i}, \quad q_i = q(a_i) \equiv (1 - e^{-a_i})/a_i. \quad (3.3)$$

Clearly, $h'(a) > 0$ and $q'(a) < 0$. Also, let ρ_i be the rate at which a w_i -worker endogenously separates from the job. Then, (2.1) and (2.7) become:

$$\begin{aligned} \text{for } i \geq 2: \quad & a_i = \lambda n_{i-1}/(v_i k), \quad \rho_{i-1} = \lambda q_i, \\ \text{for } i = 1: \quad & a_1 = \lambda_0 u/(v_1 k), \quad \rho_0 = \lambda_0 q_1. \end{aligned} \quad (3.4)$$

4. Characterizing and Computing the Equilibrium

There are many potential types of deviations from the ladder depicted in Figure 2a. In this section, we characterize the equilibrium by temporarily imposing the following restriction:

Temporary Restriction (One-rung): For all $i \in \{0, \dots, M-1\}$ and all deviations $w^d \in (w_i, w_{i+1})$, either (i) the firm posting w^d attracts only applicants at w_{i-1} and after receiving the job, such a worker's next application is to w_{i+1} , or (ii) the firm posting w^d attracts only applicants at w_i and after receiving the job, such a worker's next application is to w_{i+2} .

We will eliminate this restriction in section 6 but, for now, the restriction simplifies the characterization of the equilibrium. This restriction asks that a deviation should disturb the wage ladder by only one rung, in either the type of applicants whom the deviation attracts or the direction in which the deviator's employee will apply to other jobs in the future.

¹⁴If the belief off the equilibrium path is unrestricted, an arbitrary set of wages may be supported as an equilibrium. For example, consider an arbitrary set of wages Ω and suppose that for each wage w_i in this set, the firms recruiting at w_i give positive employment probability only to w_{i-1} -applicants. Then, even a slight deviation from w_{i-1} will reduce the recruit's future employment probability for higher wages to zero. Knowing this, workers may not apply to the deviating firm at all, and this successfully supports Ω as an equilibrium.

4.1. Wages Lower Than the Highest Level

Examine wage $w_i \in \Omega$ for $1 \leq i \leq M - 1$. Consider a deviation $w^d \in (w_{i-1}, w_{i+1})$. Under the restriction One-rung, there is no loss of generality to suppose that the deviating firm intends to attract only the applicants at w_{i-1} , in which case w^d can be regarded as a deviation from w_i because a w_{i-1} -applicant who gets the job w^d will apply to w_{i+1} next.¹⁵ Let a^d be the queue length of w_{i-1} applicants whom w^d attracts. Then, the deviating firm's hiring probability is $h(a^d)$ and each applicant's employment probability is $q(a^d)$, given by (3.3) with a^d replacing a . After an applicant gets the job, the probability of future quit rate is λq_{i+1} (note that we invoked the restriction Off-eqm here). If the deviating firm successfully hires a worker, the firm's and the employee's value functions are as follows, which are adapted from (2.9) and (2.10):

$$J_f(w^d) = \frac{y - w^d}{r + \sigma + \lambda q_{i+1}}, \quad (4.1)$$

$$J_e(w^d) = \frac{w^d + \sigma J_u - \lambda S + \lambda q_{i+1} J_e(w_{i+1})}{r + \sigma + \lambda q_{i+1}}. \quad (4.2)$$

It is more convenient to formulate the deviator's decisions as (w^d, a^d) , rather than (w^d, q^d) . Taking $(q_{i+1}, J_e(w_{i+1}))$ and other firms' strategies as given, the decisions (w^d, a^d) solve the following problem similar to (\mathcal{P}) :

$$\begin{aligned} (\mathcal{P}d) \quad & \max h(a^d) J_f(w^d) \\ \text{s.t.} \quad & q(a^d) [J_e(w^d) - J_e(w_{i-1})] = E(w_{i-1}). \end{aligned}$$

For w_i to be an equilibrium wage, the solution to $(\mathcal{P}d)$ must be $(w^d, a^d) = (w_i, a_i)$. The first-order conditions and the constraint of $(\mathcal{P}d)$ yield:

$$\begin{aligned} J_e(w_i) - J_e(w_{i-1}) &= \frac{a_i}{e^{a_i} - 1 - a_i} J_f(w_i), \\ q_i [J_e(w_i) - J_e(w_{i-1})] &= E(w_{i-1}), \end{aligned}$$

where $J_f(w_i)$ and $J_e(w_i)$ obey (4.1) and (4.2), respectively, with w_i replacing w^d . The first equation states that the applicant's surplus after getting the job, $J_e(w_i) - J_e(w_{i-1})$, is a share $a_i/(e^{a_i} - 1)$ of the total surplus. Notice that this share decreases endogenously with the queue length a_i .

In addition, the value of every vacancy is zero in equilibrium because of free entry. That is,

$$\frac{C}{h_i} = J_f(w_i) = \frac{y - w_i}{r + \sigma + \lambda q_{i+1}}. \quad (4.3)$$

¹⁵There are two other cases, but the same analysis applies with i being renumbered. The first is that the deviating firm intends to attract w_i workers with $w^d \in (w_i, w_{i+1})$. In this case, the restriction One-rung requires the w_i -worker who gets the job w^d to apply next to w_{i+2} , and so w^d can be treated as a downward deviation from w_{i+1} . The second is that the deviating firm intends to attract w_{i-2} workers with $w^d \in (w_{i-1}, w_i)$. In this case, One-rung requires the w_{i-2} -worker who gets the job w^d to apply next to w_i , and so w^d can be treated as an upward deviation from w_{i-1} .

With this condition, we rewrite the first-order conditions of $(\mathcal{P}d)$ as follows:

$$J_e(w_i) - J_e(w_{i-1}) = C/f_i, \quad (4.4)$$

$$E(w_{i-1}) = q_i [J_e(w_i) - J_e(w_{i-1})] = Cq_i/f_i, \quad (4.5)$$

where

$$f_i = f(a_i) \equiv q(a_i) (e^{a_i} - 1 - a_i). \quad (4.6)$$

Finally, for w_{i-1} -applicants to incur the fixed cost of application, the market surplus $E(w_{i-1})$ must be greater than or equal to S . With (4.5), this requirement becomes

$$a_i \leq \bar{a}, \quad \text{where } e^{\bar{a}} - 1 - \bar{a} = C/S. \quad (4.7)$$

Under the restriction One-rung, the conditions (4.3) – (4.7) ensure that (w_i, a_i) are optimal for attracting w_{i-1} -applicants among all wages in (w_{i-1}, w_{i+1}) , for any $i \in \{1, \dots, M-1\}$.

4.2. Highest Wage in Equilibrium

The highest wage w_M must satisfy (4.3) – (4.7), for $i = M$, in order to ensure that it is the optimal wage for attracting w_{M-1} -applicants. In addition, there should not be incentive for a firm to post a wage higher than w_M . A deviation $w^* > w_M$ can be intended for attracting the applicants at w_M or at other wages. However, to rule out a profitable deviation $w^* > w_M$, it is necessary and sufficient to require that posting $w^* > w_M$ to attract w_M -applicants is not profitable.¹⁶ To express this requirement formally, consider a single firm that deviates to $w^* > w_M$ with an intention to attract w_M -applicants, and let a^* be the queue length of w_M -applicants whom the deviation attracts. Then,

$$q(a^*) [J_e(w^*) - J_e(w_M)] = S, \quad (4.8)$$

where $J_e(w) = (w + \sigma J_u)/(r + \sigma)$ for both $w = w^*$ and w_M . The deviator's expected surplus is $h(a^*)J_f(w^*)$, where $J_f(w^*) = (y - w^*)/(r + \sigma)$. This deviation is not profitable if and only if the firm's maximum expected surplus is less than the vacancy cost C . Solving the deviator's maximization problem subject to (4.8), this requirement becomes:

$$w_M > y - (r + \sigma)S e^{\bar{a}}, \quad (4.9)$$

where \bar{a} is defined in (4.7). Clearly, for this condition to be satisfied and yet $w_M \leq y$, we need $S > 0$. This is why we introduced the fixed cost of application.

To explain intuitively why $S > 0$ is needed for an equilibrium, suppose $S = 0$ and $w_M < y$. A firm that deviates to a slightly higher wage $w_M + \varepsilon$ ($\varepsilon > 0$) can always attract w_M -applicants,

¹⁶This condition guarantees that it is not profitable to post $w^* > w_M$ to attract w_{M-j} -applicants, for $j \geq 1$. The proof consists of two stages. First, by construction, w_M yields the highest expected surplus to w_{M-1} -applicants and hence those applicants prefer to apply to w_M to all other wages $w^* > w_M$. Second, if w_{M-1} -applicants prefer w_M to $w^* > w_M$, so do all other applicants at wages w_{M-j} ($j \geq 2$). The proof for the second stage is similar to that of Lemma 6.4 later and hence omitted.

and so it can succeed in hiring almost surely. Relative to posting w_M , the deviation gives the firm a slightly lower ex post surplus but a discrete increase in the hiring probability. Thus, the deviation is profitable. To prevent such profitable deviations, w_M must be equal to or greater than y , which yields negative expected net profit, after the vacancy cost is deducted.

For future use, it is useful to express (4.7) and (4.9) for $i = M$ as requirements on the hiring probability at wage w_M , as follows:¹⁷

$$1 - (1 + \bar{a})e^{-\bar{a}} < h_M \leq 1 - e^{-\bar{a}}. \quad (4.10)$$

4.3. Recursive Characterization

The conditions in previous subsections provide a recursive characterization of the equilibrium. Pick up a number h_M that satisfies (4.10). Then, $q_{M+1} = 0$. Moreover,

$$\begin{aligned} a_M &= -\ln(1 - h_M), & q_M &= h_M/a_M, \\ w_M &= y - (r + \sigma)C/h_M, & J_e(w_M) &= (w_M + \sigma J_u)/(r + \sigma). \end{aligned}$$

The result for w_M comes from setting $i = M$ in (4.3), and the result for $J_e(w_M)$ from setting $i = M$ and $w^d = w_M$ in (4.2). Proving by induction (see Appendix B), we have the following proposition.

Proposition 4.1. (Recursive) *Given h_M , q_{M-j+1} , w_{M-j} and $J_e(w_{M-j})$, the following conditions hold in equilibrium for $j = 0, 1, 2, \dots, M - 2$:*

$$h_{M-j} = \frac{(r + \sigma + \lambda q_{M-j+1})C}{y - w_{M-j}}, \quad (4.11)$$

$$a_{M-j} = -\ln(1 - h_{M-j}), \quad q_{M-j} = h_{M-j}/a_{M-j}, \quad (4.12)$$

$$w_{M-j-1} = w_M + \lambda S - \frac{\lambda q_{M-j}C}{f_{M-j}} - (r + \sigma)C \sum_{t=0}^j \frac{1}{f_{M-t}}, \quad (4.13)$$

$$J_e(w_{M-j-1}) = \frac{\sigma J_u + w_M}{r + \sigma} - C \sum_{t=0}^j \frac{1}{f_{M-t}}. \quad (4.14)$$

In addition, (4.11) and (4.12) hold for $j = M - 1$, and J_u satisfies:

$$J_u = \frac{1}{r} \left[b - \lambda_0 S + \lambda_0 C \frac{q_1}{f_1} \right]. \quad (4.15)$$

The recursive method generates the sequence $(h_i, a_i, q_i, w_i, J_e(w_i))$ for given h_M . For the sequences to be an equilibrium, the value of h_M must be such that the generated value of $J_e(w_1)$ satisfies (4.4) for $i = 1$. We will examine such existence in the next subsection.

¹⁷For $i = M$, (4.7) is equivalent to $h_M \leq h(\bar{a}) = 1 - e^{-\bar{a}}$. To rewrite (4.9), note that $h_M [J_f(w_M) - J_v] = C$ in equilibrium and $J_f(w_M) - J_v = (y - w_M)/(r + \sigma)$.

Once the equilibrium sequence of (h, a, q) is determined, we can obtain the distributions of workers and vacancies. First, because the equilibrium is stationary, the measure of workers who separate from w_i must be equal to the measure of workers newly recruited at wage w_i . That is,

$$\begin{aligned} (\sigma + \lambda q_{i+1})n_i &= \lambda n_{i-1}q_i, \text{ for all } 2 \leq i \leq M, \\ (\sigma + \lambda q_2)n_1 &= \lambda_0 u q_1 = \sigma(1 - u) \text{ and } u = 1 - \sum_{i=1}^M n_i. \end{aligned} \quad (4.16)$$

These equations solve for u and $(n_i)_{i=1}^M$. Second, (3.4) implies

$$v_i = \lambda n_{i-1}/(a_i k) \text{ for } i \geq 2, \quad \text{and } v_1 = \lambda_0 u/(a_1 k). \quad (4.17)$$

Together with $\sum_{i=1}^M v_i = 1$, these equations solve for k and $(v_i)_{i=1}^M$. The density of offer wages is $(v_i)_{i=1}^M$ and of employed wages $(n_i/(1 - u))_{i=1}^M$.

4.4. Hiring Probability at the Highest Wage and the Number of Rungs

The equilibrium values of M and h_M are such that (4.4) is satisfied for $i = 1$. That is, starting with h_M , the computed sequence of J_e satisfies $\Delta = 0$, where $\Delta \equiv J_e(w_1) - J_u - C/f_1$. Setting $j = M - 2$ in (4.14) to obtain $J_e(w_1)$ and substituting (4.15) for J_u , we get:

$$\Delta = \frac{w_M - b + \lambda_0 S}{r + \sigma} - \frac{C \lambda_0 q_1}{(r + \sigma) f_1} - C \sum_{t=0}^{M-1} \frac{1}{f_{M-t}}. \quad (4.18)$$

By Proposition Recursive, the computed a sequence and w_M depend on the chosen values of (M, h_M) , but not on J_u directly. Hence, we write $\Delta = \Delta(M, h_M)$. Equilibrium values of (M, h_M) solve $\Delta(M, h_M) = 0$.

To examine the function $\Delta(M, h_M)$, we maintain the following assumption.

Assumption 1. (Regularity) Define \bar{a} by (4.7). Assume that the following conditions hold:

$$b \leq y + \lambda_0 S - C \left[\frac{(r + \sigma) e^a + \lambda_0}{e^a - 1 - a} \right]_{a=\bar{a}-\ln(1+\bar{a})}, \quad (4.19)$$

$$(r + \sigma)/\lambda > f(\bar{a})/\bar{a}. \quad (4.20)$$

The condition (4.19) ensures that there is at least one wage level that yields a higher present value to the workers than unemployment, while (4.20) is a technical condition necessary for exploring Proposition Recursive. The following lemma is proven in Appendix B.

Lemma 4.2. Fix h_M at any value h^* that satisfies (4.10), and compute the a sequence by Proposition Recursive. Then, there exists an integer $M^* \geq 1$ such that $\Delta(M', h^*) \geq 0$ for all $M' \leq M^*$ and $\Delta(M', h^*) < 0$ for all $M' \geq M^* + 1$.

Now we can find the lowest equilibrium value of h_M as follows. Choose h^* in the above lemma to be the lower bound on h_M given by (4.10) and obtain the corresponding M^* . Then, $\Delta(M^*, h^*) \geq 0$. If $\Delta(M^*, h^*) = 0$, then the chosen value of h^* is the lowest equilibrium value of h_M . Suppose $\Delta(M^*, h^*) > 0$. Then $\Delta(M^* + 1, h^*) < 0$. Reset $M = M^* + 1$ and increase h^* . By Proposition Monotone established later, the a sequence is an increasing function of the chosen value of h_M . So is w_M . Thus, $\Delta(M, h^*)$ is an increasing function of h^* (for given M). For there to be an equilibrium solution for h_M , $\Delta(M, h^*)$ must increase to cross 0 when h^* increases to the upper bound in (4.10). The first crossing gives the lowest equilibrium value of h_M , where $M = M^* + 1$ is the equilibrium value of M .

Similarly, we can find the highest equilibrium value of h_M . To do so, choose the upper bound of h_M , $1 - e^{-\bar{a}}$, to be the starting value of h^* in the above lemma and compute the corresponding M^* . If $\Delta(M^*, h^*) = 0$, then the starting value of h^* is the highest equilibrium value of h_M . If $\Delta(M^*, h^*) > 0$, then fix $M = M^*$ and reduce h^* . For there to be an equilibrium solution for h_M , $\Delta(M, h^*)$ must decrease to cross 0 when h^* decreases to the lower bound of h_M in (4.10). The first crossing gives the highest equilibrium value of h_M , where $M = M^*$ is the equilibrium value of M . Therefore, the following proposition holds:

Proposition 4.3. (M-exists) *Set $h^* = 1 - (1 + \bar{a})e^{-\bar{a}}$ and compute M^* as in Lemma 4.2. There exists an equilibrium value of h_M if and only if*

$$\Delta(M^*, 1 - e^{-\bar{a}}) \geq 0. \quad (4.21)$$

Under this condition, there exist h_L and h_H , which possibly coincide with each other, such that all equilibrium values of h_M lie in $[h_L, h_H]$. The equilibrium value of M is either M^ or $M^* + 1$.*

It is analytically difficult to verify (4.21) or to check whether the solution for h_M is unique. In section 5.2 we will provide some numerical examples.

5. Properties of the Equilibrium

We delay to section 6 the proof that, even without the restriction One-rung, the wage ladder constructed in the previous section can be an equilibrium. In the current section, we investigate the properties of the equilibrium. Whenever possible, we will suppress the index $i = M - j$, use the subscript $\pm t$ to stand for $M - j \pm t$, where $t \geq 1$.

5.1. Analytical Properties

To study the properties of the equilibrium, we examine the properties of the sequence computed with any given value of h_M that satisfies (4.10). We start by deriving some useful equations from Proposition Recursive. Subtracting (4.13) for j and $j + 1$, we get:

$$w - w_{-1} = \frac{(r + \sigma + \lambda q)C}{f} - \frac{\lambda C q_{+1}}{f_{+1}}. \quad (5.1)$$

This condition holds for all $i = M - j$, where $j \in \{1, 2, \dots, M - 2\}$. For $j = 0$, replace the last term in the equation by $\lambda Cq(\bar{a})/f(\bar{a}) = \lambda S$. Also, for all $j \in \{1, \dots, M - 2\}$, we have:¹⁸

$$h_{-1} = (r + \sigma + \lambda q) \left/ \left(\frac{r + \sigma + \lambda q_{+1}}{h} + \frac{r + \sigma + \lambda q}{f} - \frac{\lambda q_{+1}}{f_{+1}} \right) \right. . \quad (5.2)$$

A similar equation holds for $j = 0$, with the first q_{+1} being replaced by 0 and the term $\lambda q_{+1}/f_{+1}$ by $\lambda q(\bar{a})/f(\bar{a}) = \lambda S/C$.

We establish the following propositions, in Appendices C and D, respectively.

Proposition 5.1. (Monotone) *For any given h_M that satisfies (4.10), the sequence constructed in Proposition Recursive has the following monotonicity properties for all $0 \leq j \leq M - 2$ (where the subscript $M - j$ is suppressed):*

$$a_{-1} < a \leq \bar{a}, \quad h_{-1} < h, \quad q_{-1} > q. \quad (5.3)$$

$$(r + \sigma)/\lambda > f(a_{-1})/a_{-1}. \quad (5.4)$$

$$a_{-1} > a - \ln(1 + a), \quad (5.5)$$

$$da/dh_M > 0, \quad dw/dh_M > 0. \quad (5.6)$$

Proposition 5.2. (W-property) *For any given h_M that satisfies (4.10), the computed sequence satisfies: (i) $w > w_{-1}$ and (ii) $E(w_{-1}) > E(w) \geq S$. In addition, if the necessary condition for the equilibrium described later in Lemma No-leap holds, then (iii) $w - w_{-1} > w_{+1} - w$.*

These propositions reveal interesting properties of the equilibrium, as listed below:

- A firm is more likely to succeed in hiring at a higher wage than at a lower wage, while an applicant is more successful getting a low-wage job than a high-wage job. This result arises from the fact that the queue length of applicants increases endogenously with wage.
- A worker at a high wage is less likely to quit than a worker at a low wage. The quit rate of a worker at wage w is $\rho(w) = \lambda q(a_{+1})$. Because the employment probability for the next wage ($q(a_{+1})$) decreases as wage increases, the quit rate falls.
- The longer a worker has been employed since his last unemployment, the higher his current wage is on average. Moreover, the higher wage a worker had just before becoming unemployed, the longer he will take on average to return to this previous wage level.

¹⁸This equation comes from the following relationships:

$$\begin{aligned} \frac{r + \sigma + \lambda q}{h_{-1}} &= \frac{1}{C} (y - w_{-1}) = \frac{1}{C} (y - w) + \frac{1}{C} (w - w_{-1}) \\ &= \frac{r + \sigma + \lambda q_{+1}}{h} + \frac{1}{C} (w - w_{-1}) = \frac{r + \sigma + \lambda q_{+1}}{h} + \frac{r + \sigma + \lambda q}{f} - \frac{\lambda q_{+1}}{f_{+1}}. \end{aligned}$$

The first equality comes from using (4.11) for $j + 1$, the second equality from rewriting, the third equality from using (4.11) for j , and the last equality from substituting (5.1).

- A worker's ex post value of employment increases with wage. That is, $J_e(w) > J_e(w_{-1})$, which we used in Lemma Singleton. Also, $E(w_i) \geq S$ for all $1 \leq i \leq M - 1$. So, all applicants employed below the highest wage are indeed willing to incur the fixed cost S to apply to higher wages.
- The gap between two adjacent rungs on the ladder shrinks as a worker climbs up the ladder. So, the higher the wage, the smaller the wage gain in the next job change.
- An applicant's expected surplus diminishes as the applicant moves up the wage ladder, i.e., $E(w_{-1}) > E(w)$, despite the fact that the ex post value of employment increases with wage. This implies that an applicant's employment probability must decrease more rapidly than the increase in wage along the wage ladder.

All these properties are realistic features of the labor market. What is special about these properties is that they all originate from the search frictions, not from exogenous differences between workers. The last property is a necessary (although not a sufficient) condition for the wage ladder to be an equilibrium in a directed search environment. Since an applicant observes all offers, to induce him to apply only to the next wage level, the expected surplus must be higher from applying for such jobs than for the jobs at higher wages.

The next proposition describes some features of the wage density in the equilibrium (see the middle part of Appendix D for a proof):

Proposition 5.3. (W-density) *The density (frequency) of offer wages decreases with wage. A sufficient condition for the density of employed wages to be decreasing at the upper end of the wage support is:*

$$\frac{\sigma}{\lambda} > \frac{1 - (1 + \bar{a})e^{-\bar{a}}}{\bar{a} - \ln(1 + \bar{a})}. \quad (5.7)$$

When r is sufficiently close to 0, a sufficient condition for the above inequality is $C/S > 2.373$. A sufficient condition for the density of employed wages to be increasing at the upper end of the wage support is $\sigma/\lambda < q(\bar{a})$.

The density of offer wages is a decreasing function. This result is opposite to what Burdett and Mortensen (1998) established in a on-the-job search model with undirected search. Our result is easy to understand. In the stationary equilibrium, the flow of workers into every equilibrium wage w must be equal to the outflow. Because the outflow consists of exogenous separation and quits, the inflow must exceed the number of quits. The inflow of workers into w is the number of new hires at w , i.e., hv . The number of quits from w is equal to the number of new hires at the next wage level, $h_{+1}v_{+1}$, because the applicants at w are the sole source of hiring at the next wage. Thus, $hv > h_{+1}v_{+1}$ in the stationary equilibrium. This necessarily implies $v > v_{+1}$, because the hiring probability at w is lower than that at w_{+1} . Therefore, the density of offer wages necessarily decreases with wage.

The density of employed wages can also be decreasing, but it is not always so. This is because the density of employed wages depends on both the inflow and the outflow of workers. Although there is a larger flow of workers into a low wage than into a high wage, the quit rate is also higher for a low wage than for a high wage. There are more workers employed at a low wage than at a high wage if and only if the difference between the inflows into the two wages is larger than the difference between the outflows. This is satisfied at the upper end of the wage distribution if the hiring cost is large relative to the application cost. In general, however, the density of employed wages may not be monotonic with respect to wages and if it is monotonic, it may not be increasing.

5.2. Numerical Examples

We provide numerical examples to illustrate the equilibrium. Consider the following values of the parameters: $y = 1000$, $b = 0$, $C = 60$, $S = 1$, $r = 0.02$, $\lambda = 0.025$, $\lambda_0 = 0.2$, and $\sigma = 0.125$. These parameter values satisfy all requirements for an equilibrium (see Proposition 6.7). Under these parameter values, there is a unique equilibrium and the wage distribution has four rungs. The unemployment rate is $u = 40.9\%$ and the overall vacancy-worker ratio is $k = 0.41$.¹⁹ Other characteristics of this equilibrium are summarized in Table 2.

Table 2. Equilibrium in a numerical example

i	w_i	v_i (%)	$\frac{n_i}{1-u}$ (%)	a_i	q_i (%)	h_i (%)
1	947.6	95.7	87.8	0.21	90.4	18.6
2	982.6	4.0	11.2	0.78	69.6	53.9
3	988.9	.2	1.0	1.77	46.9	83.0
4	990.8	<.1	.1	2.94	32.2	94.7

These numerical results confirm the analytical properties of (w, a, q, h, v) established in Propositions Monotone and W-property. A notable feature is that the small difference between two adjacent wages induces large differences in the employment probability, the hiring probability and the density of offer wages. For example, when the wage increases from 947.6 to 982.6, the employment probability falls sharply from 90.4% to 69.6%, the hiring probability increases from 18.6% to 53.9% and the density of offer wages falls from 95.7% to 4.0%. A predominant fraction of firms recruit at the lowest wage.

The density of employed wages is also a decreasing function of wages in this equilibrium. A large fraction of workers are employed at the lowest wage, although the distribution is less skewed toward low wages than the offer wage distribution.

In Tables 3 we illustrate the importance of on-the-job search for the wage ladder, by changing λ while fixing other parameters (including λ_0). When $\lambda = 0$, on-the-job search is shut down and, as previous models of directed search predict, the wage distribution is degenerate, with $w = 953.2$,

¹⁹It is possible to find parameter values which allow for the possibility of multiple equilibria (with different number of rungs).

$a = .21$, $u = 40.9\%$ and $k = .40$. However, the wage ladder becomes non-degenerate as soon as workers can search while employed. Even for $\lambda = 10^{-4}$, the support of equilibrium wages is significantly dispersed. This result illustrates that the degeneracy of the wage distribution in previous directed search models is not robust. As the opportunity of on-the-job search increases further, most of the characteristics of the equilibrium change very little. The support of the distribution Ω , the employment probabilities $\{q_i\}$, the hiring rates $\{h_i\}$, the unemployment rate u and market tightness k remain very similar. Even though the employment probabilities upon applying are unaffected, workers get to apply more often as λ increases, and as a result, the distributions of offer and employed wages become less skewed.

Table 3. Various degrees of on-the-job search

	$\lambda = 0$	$\lambda = 10^{-4}$	$\lambda = .01$	$\lambda = .025$		$\lambda = 0$	$\lambda = 10^{-4}$	$\lambda = .01$	$\lambda = .025$
w_1	953.2	953.2	951.0	947.6	a_1	.21	.21	.21	.21
w_2		984.0	983.4	982.6	a_2		.79	.78	.78
w_3		989.6	989.4	988.9	a_3		1.82	1.80	1.77
w_4		991.0	990.9	990.8	a_4		3.28	3.13	2.94
%	$\lambda = 0$	$\lambda = 10^{-4}$	$\lambda = .01$	$\lambda = .025$	%	$\lambda = 0$	$\lambda = 10^{-4}$	$\lambda = .01$	$\lambda = .025$
v_1	100	99.9	98.2	95.7	$\frac{n_1}{1-u}$	100	99.9	94.7	87.8
v_2		<.1	1.8	4.0	$\frac{n_2}{1-u}$		<.1	5.1	11.2
v_3		<.1	<.1	.2	$\frac{n_3}{1-u}$		<.1	.2	1.0
v_4		<.1	<.1	<.1	$\frac{n_4}{1-u}$		<.1	<.1	.1
			$\lambda = 0$	$\lambda = 10^{-4}$	$\lambda = .01$	$\lambda = .025$			
u			40.9%	40.9%	40.9%	40.9%			
k			.40	.40	.40	.41			

6. Wage Ladder Is an Equilibrium

In this section, we eliminate the restriction One-rung imposed in section 4 and show that the wage ladder is an equilibrium under certain conditions. We divide wages that violate the restriction One-rung into two categories, those that belong to the equilibrium wage support and those that lie outside. The two subsections below examine the two categories in turn. The restriction Off-eqm is maintained throughout. As in the previous section, we suppress the index $i = M - j$ whenever possible and use the subscript $\pm t$ to stand for $M - j \pm t$.

6.1. Workers Apply Only to the Next Wage Level

In this subsection, we confine deviating wages to the equilibrium set of wages, Ω . A wage $w \in \Omega$ is a deviation that violates the restriction One-rung if the wage w is intended to attract the applicants at w_{i-1} but $w \neq w_i$. For such deviations to be not profitable, the expected surplus that a w_{i-1} -applicant obtains from applying to wage w_i must be larger than or equal to that from applying to any other wage in Ω . That is,

$$q[J_e(w) - J_e(w_{-1})] \geq q_{+t}[J_e(w_{+t}) - J_e(w_{-1})]. \quad (6.1)$$

This requirement hold for all $t \in \{1, 2, \dots, j\}$ and all $j \in \{1, 2, \dots, M - 1\}$. Notice that we have invoked Restriction Off-eqm to compute the payoff of a w_{-1} -applicant who applies to wages other than w , because such application is off the equilibrium path.

Lemma 6.1. *For all j and all $t \geq 2$, if the w_{-1} -applicants prefer applying to a job at wage w to a job at w_{+1} , then so do the w_{-t} -applicants.*

Proof. Suppose that the w_{-1} -applicants prefer to apply to w relative to w_{+1} , i.e.,

$$q [J_e(w) - J_e(w_{-1})] \geq q_{+1} [J_e(w_{+1}) - J_e(w_{-1})].$$

For all $t \geq 2$, we have

$$\begin{aligned} & q [J_e(w) - J_e(w_{-t})] - q_{+1} [J_e(w_{+1}) - J_e(w_{-t})] \\ = & \{q [J_e(w) - J_e(w_{-1})] - q_{+1} [J_e(w_{+1}) - J_e(w_{-1})]\} + (q - q_{+1}) [J_e(w_{-1}) - J_e(w_{-t})]. \end{aligned}$$

The difference in $\{.\}$ is non-negative by the supposition. The last term on the right-hand side is also positive, because $q > q_{+1}$ and $J_e(w_{-1}) > J_e(w_{-t})$ for all $t \geq 2$. Thus, the above deference is positive, implying that applying to w yields a higher expected surplus for a w_{-t} -applicant than applying to w_{+1} . **QED**

The intuition for the above lemma is the same as that for Lemma Singleton. An applicant at wage w_{-1} is more willing to sacrifice the employment probability for the wage level than an applicant at lower wages w_{-t} ($t \geq 2$). If the high employment probability offered with wage w is more attractive to an applicant at w_{-1} than the higher wage w_{+1} , it must be so as well to an applicant at lower wages w_{-t} .

The above lemma significantly reduces the number of inequalities we need to verify for (6.1). For each $j \in \{1, 2, \dots, M - 1\}$, it suffices to verify (6.1) for only $t = 1$. However, there are still $M - 1$ such inequalities. To reduce the number further, we use (4.4) to rewrite (6.1) for $t = 1$ as:

$$\frac{q(a)}{q(a_{+1})} - 1 - \frac{f(a)}{f(a_{+1})} \geq 0. \quad (6.2)$$

For given a_{+1} , define $\phi(a_{+1})$ as the solution for a to the equality form of (6.2). Because the left-hand side of (6.2) is a decreasing function of a , the inequality is equivalent to $a \leq \phi(a_{+1})$.

We prove the following lemma in Appendix E:

Lemma 6.2. *The function $\phi(.)$ exists, is unique for each a_{+1} , and has the following properties: (i) $\phi' > 0$; (ii) $a_{+1} > \phi(a_{+1}) > a_{+1} - \ln(1 + a_{+1})$; and (iii) if $a \leq \phi(a_{+1})$, then $a_{-1} < \phi(a)$.*

Property (iii) in Lemma 6.2 says that an applicant at wage w_{M-2} has the strongest incentive among all applicants to leap on the wage ladder. To verify (6.2) for all $j \in \{1, 2, \dots, M - 1\}$, it is sufficient to verify it for $j = 1$. For $j = 1$, (6.2) becomes $a_{M-1} \leq \phi(a_M)$, which is equivalent to

$h_{M-1} \leq h(\phi(a_M))$. Express (5.2) for $j = 0$ by replacing q_{+1} by 0 and $\lambda q_{+1}/f_{+1}$ by $\lambda S/C$. Using this equation to substitute for h_{M-1} , we rewrite $h_{M-1} \leq h(\phi(a_M))$ as follows:

$$0 \leq \frac{r + \sigma}{\lambda} \left(\frac{1}{h_M} + \frac{1}{f_M} - \frac{1}{h(\phi(a_M))} \right) - q_M \left(\frac{1}{h(\phi(a_M))} - \frac{1}{f_M} \right) - \frac{S}{C}.$$

Rewriting this condition further, we have:²⁰

Lemma 6.3. (No-leap) *Given the wage levels computed in Proposition Recursive, it is optimal for applicants to apply only to the next wage level, rather than higher ones, if and only if*

$$\frac{r + \sigma}{\lambda} \geq \left(\frac{q_M}{h(\phi(a_M))} - \frac{q_M}{f_M} + \frac{S}{C} \right) / \left(\frac{1}{h_M} + \frac{1}{f_M} - \frac{1}{h(\phi(a_M))} \right). \quad (6.3)$$

6.2. Deviations That Lie Outside the Equilibrium Wage Support

Now we examine deviations that violate the restriction One-rung and lie outside Ω . Let $w^d \in (w_{-1}, w)$ be such a deviation. This deviation can violate One-rung either in the type of applicants it attracts, or in the direction in which the prospect employee will apply to other jobs, or in both. The following lemma narrows down the types of such deviations we need to consider.

Lemma 6.4. *The following statements are true regarding any deviation $w^d \in (w_{-1}, w)$. (i) If an applicant gets the w^d -job, then his future application is to either w or w_{+1} . (ii) If w_{-2} -applicants do not have incentive to apply to w^d , then neither do w_{-t} -applicants, where $t \geq 3$.*

Proof. For (i), we show that a w^d -applicant does not have incentive to apply to w_{+t} , for all $t \geq 2$. Suppose that a worker gets the job w^d . Let $J_e^d(w^d)$ be the value function of such a worker employed at w^d . Because this worker is not restricted to applying to w_{+1} next, $J_e^d(w^d)$ may not obey (4.2). However, whatever job opportunities a worker at w^d will have in the future, a worker employed at wage w will have as well with the same probability (under Restriction Off-eqm). Thus, $J_e^d(w^d) < J_e(w)$. For the worker employed at w^d , applying to w_{+1} next yields a higher expected surplus than to any higher wage w_{+t} ($t \geq 2$), as shown below:

$$\begin{aligned} & q_{+t} \left[J_e(w_{+t}) - J_e^d(w^d) \right] - q_{+1} \left[J_e(w_{+1}) - J_e^d(w^d) \right] \\ &= q_{+t} \left[J_e(w_{+t}) - J_e(w) \right] + q_{+t} \left[J_e(w) - J_e^d(w^d) \right] - q_{+1} \left[J_e(w_{+1}) - J_e^d(w^d) \right] \\ &< q_{+1} \left[J_e(w_{+1}) - J_e(w) \right] + q_{+t} \left[J_e(w) - J_e^d(w^d) \right] - q_{+1} \left[J_e(w_{+1}) - J_e^d(w^d) \right] \\ &< q_{+1} \left[J_e(w_{+1}) - J_e(w) \right] + q_{+1} \left[J_e(w) - J_e(w^d) \right] - q_{+1} \left[J_e(w_{+1}) - J_e(w^d) \right] \\ &= 0. \end{aligned}$$

²⁰To rewrite the condition as a lower bound on $(r + \sigma)/\lambda$, note that $\frac{1}{h} + \frac{1}{f} = [1 - (1 + a)e^{-a}]^{-1}$. So, the fact $\phi(a) > a - \ln(1 + a)$ implies $\frac{1}{h(a_M)} + \frac{1}{f(a_M)} > \frac{1}{h(\phi(a_M))}$. Moreover, if the right-hand side of the condition is an increasing function of a_M (which seems true from numerical examples), then we can replace (6.3) by the following sufficient condition (using the facts that $a_M \leq \bar{a}$ and $C/S = (e^{\bar{a}} - 1 - \bar{a})$):

$$\frac{r + \sigma}{\lambda} \geq \frac{q(\bar{a})}{h(\phi(\bar{a})) \left[\frac{1}{h(\bar{a})} + \frac{1}{f(\bar{a})} \right] - 1}.$$

The first inequality follows from our previous result that a w -applicant prefers to apply to w_{+1} relative to w_{+t} for all $t \geq 2$. The second inequality comes from the facts that $q_{+t} < q_{+1}$ for all $t \geq 2$ and $J_e(w) > J_e^d(w^d)$.

For (ii), suppose that w_{-2} -applicants do not have incentive to apply to w^d . Then,

$$q^d [J_e^d(w^d) - J_e(w_{-2})] \leq q_{-1} [J_e(w_{-1}) - J_e(w_{-2})],$$

where q^d is the probability with which an applicant to w^d gets the job. Using an argument similar to that established $J_e(w) > J_e^d(w^d)$, we have $J_e^d(w^d) > J_e(w_{-1})$. Then, the above inequality implies $q^d < q_{-1}$. Now, for all $t \geq 3$, the following relationships hold:

$$\begin{aligned} & q^d [J_e^d(w^d) - J_e(w_{-t})] - q_{-1} [J_e(w_{-1}) - J_e(w_{-t})] \\ = & q^d [J_e^d(w^d) - J_e(w_{-2})] + q^d [J_e(w_{-2}) - J_e(w_{-t})] - q_{-1} [J_e(w_{-1}) - J_e(w_{-t})] \\ \leq & q_{-1} [J_e(w_{-1}) - J_e(w_{-2})] + q^d [J_e(w_{-2}) - J_e(w_{-t})] - q_{-1} [J_e(w_{-1}) - J_e(w_{-t})] \\ < & q_{-1} [J_e(w_{-1}) - J_e(w_{-2})] + q_{-1} [J_e(w_{-2}) - J_e(w_{-t})] - q_{-1} [J_e(w_{-1}) - J_e(w_{-t})] \\ = & 0. \end{aligned}$$

The first inequality follows from the supposition about w_{-2} -applicants, and the second inequality from $q^d < q_{-1}$. Therefore, a w_{-t} -applicant gets a higher expected surplus from applying to w_{-1} than to w^d . **QED**

The intuition for the above lemma is similar to that for Lemma Singleton. This similarity is obvious for part (ii), because part (ii) extends Lemma 6.1 to wages outside the equilibrium support and Lemma 6.1 relies on the intuition for Lemma Singleton. To see the link between part (i) and Lemma Singleton, notice that $w > w^d$. According to the intuition for Lemma Singleton, an applicant at w is more willing to sacrifice the employment probability for the wage level than an applicant at w^d . Because the high employment probability with wage w_{+1} is more attractive to a w -applicant than higher wages, it must be even more attractive to an applicant at the lower wage w^d .

Under the above lemma, only the following two types of deviations still need be ruled out:

- (I) The deviation $w^d \in (w_{-1}, w)$ attracts w_{-1} -applicants and, after getting the job with w^d , such a worker will apply to w in the future.
- (II) The deviation $w^d \in (w_{-1}, w)$ attracts w_{-2} -applicants and, after getting w^d , such a worker will apply to w_{+1} .

Figure 3 depicts these two deviations, where the dashed arrows indicate the deviating firm's source of applicants and its employee's future application. A type I deviation is profitable only when the support of the wage distribution is too sparse. A type II deviation is profitable only when the support is too dense, and it is meaningful only when the number of rungs on the wage

ladder is three or more. We will investigate these two deviations in turn.²¹

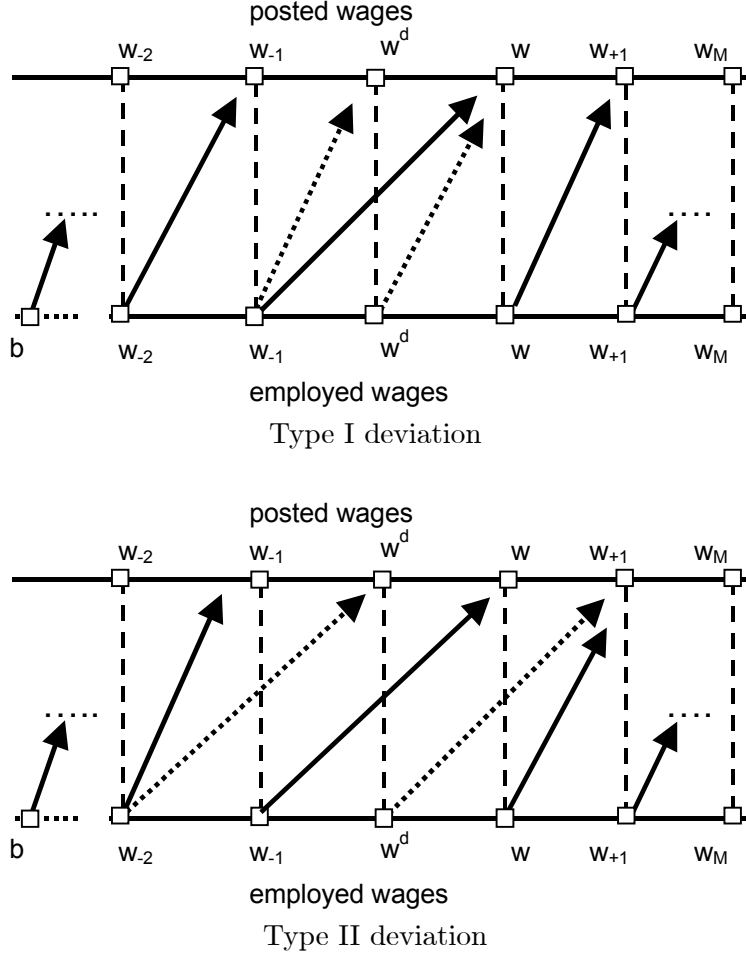


Figure 3

Consider first a type I deviation $w^d \in (w_{-1}, w)$. Let $J_f^d(w^d)$ be the deviating firm's value function after successfully recruiting a worker and $J_e^d(w^d)$ the value function of a worker who gets the job w^d , conditional on that the worker's future application is to w as required in a type I deviation. Then,

$$J_f^d(w^d) = \frac{y - w^d}{r + \sigma + \lambda q},$$

$$J_e^d(w^d) = \frac{w^d + \sigma J_u - \lambda S + \lambda q J_e(w)}{r + \sigma + \lambda q}.$$

²¹Other deviations and, in particular, the following two types of deviations, are not profitable. The first is like the one depicted in the upper panel of Figure 3, but with the employee at w^d applying to w_{+1} in the future. This deviation can be viewed as a downward deviation from w that satisfies One-rung and, by construction, the deviation is dominated by w . The second deviation is like the one depicted in the lower panel of Figure 3, but with the employee at w^d applying to w in the future. This deviation can be viewed as an upward deviation from w_{-1} that satisfies One-rung and, by construction, the deviation is dominated by w_{-1} .

These functions are different from those in (4.1) and (4.2), because the worker's future application direction is different from the one depicted in Figure 2a (again, we invoked Off-eqm).

For the deviation w^d to be profitable, it must satisfy the following conditions:

- (Ia) By applying to w^d , a w_{-1} -applicant's expected surplus is equal to $E(w_{-1})$;
- (Ib) The deviating firm earns an expected surplus greater than C .

These two conditions cannot be both satisfied. To see this, suppose that the deviation satisfies (Ia). Let q^d be the employment probability of an applicant applying to w^d . Then, (Ia) implies

$$q^d [J_e^d(w^d) - J_e(w_{-1})] = E(w_{-1}) = Cq/f. \quad (6.4)$$

Substituting $J_e^d(w^d)$ and $J_e(w_{-1})$ into the above equation, we solve w^d as follows:

$$w^d = w_{-1} + (r + \sigma + \lambda q) Cq / (fq^d). \quad (6.5)$$

The deviating firm's expected surplus is $\pi(a^d) = h^d J_f^d(w^d)$, where h^d is the firm's hiring probability. Substituting $(y - w_{-1})$ from (4.11) and $(w^d - w_{-1})$ from (6.5), $\pi(a^d)$ becomes:

$$\pi(a^d) = h^d \left(\frac{C}{h_{-1}} - \frac{qC}{fq^d} \right) = C \left(\frac{h^d}{h_{-1}} - \frac{a^d q}{f} \right).$$

The surplus $\pi(a^d)$ is maximized at $a^d = a^*$, where $a^* = \ln \left(\frac{f}{qh_{-1}} \right)$ solves $\pi'(a^*) = 0$. The maximum is

$$\pi(a^*) = C \left(\frac{h(a^*)}{h_{-1}} - \frac{a^* q}{f} \right) = \frac{Cq}{f} (e^{a^*} - 1 - a^*) = \frac{e^{a^*} - 1 - a^*}{e^a - 1 - a} C,$$

where the second equality comes from substituting $h_{-1} = e^{-a^*} f/q$, and the third equality from the definition of f . Because $(e^a - 1 - a)$ is an increasing function, a necessary condition for the deviation to be profitable is $a^* > a$. However, a queue length $a^d > a$ will not attract any w_{-1} applicants. To see this, note that a w -worker is employed at a higher wage than a w^d -worker and has all future job opportunities that a w^d -worker will have. Thus, $J_e(w) > J_e^d(w^d)$. If $a^d > a$, then $q^d < q$, which would lead to the following contradiction to (6.4):

$$E(w_{-1}) = q [J_e(w) - J_e(w_{-1})] > q [J_e^d(w^d) - J_e(w_{-1})] > q^d [J_e^d(w^d) - J_e(w_{-1})].$$

Therefore, we have established the following lemma:

Lemma 6.5. (Type-I) *A type I deviation is not profitable.*

The explanation is as follows. A type-I deviation $w^d \in (w_{-1}, w)$ competes against an equilibrium wage w for the same applicants (i.e., w_{-1} -applicants). In comparison with w , the deviation w^d offers a w_{-1} -applicant not only a lower wage but also a lower value for future application. For

the deviation to attract this applicant, it must provide a significantly higher employment probability than a job opening at w does. This implies that the deviating firm's hiring probability must be significantly lower than that of a firm recruiting at w . In this case, the deviator's expected surplus from recruiting will not be high enough to cover the vacancy cost.

Now consider a type II deviation $w^d \in (w_{-1}, w)$, depicted in the lower panel in Figure 3. With this deviation, the deviating firm's ex post value $J_f^d(w^d)$ and the employee's value $J_e^d(w^d)$ are:

$$J_f^d(w^d) = \frac{y - w^d}{r + \sigma + \lambda q_{+1}}, \quad (6.6)$$

$$J_e^d(w^d) = \frac{w^d + \sigma J_u - \lambda S + \lambda q_{+1} J_e(w_{+1})}{r + \sigma + \lambda q_{+1}}. \quad (6.7)$$

In contrast to a type-I deviation, a type-II deviation changes not only the applicant's current payoff but also the future payoff, by changing the potential employee's future application. As a result, it is more difficult to prove that a type-II deviation is not profitable. The following lemma is proven in Appendix F:

Lemma 6.6. (Type-II) *Define β^* by the following equation:*

$$h(\beta^*) = 1 \left/ \left[\frac{1}{h} + \frac{q_{+1}}{(q - q_{+1}) f_{+1}} \right] \right. . \quad (6.8)$$

A type II deviation is not profitable if and only if the following condition is satisfied:

$$\beta^* - h(\beta^*) e^{a-1} + \frac{r + \sigma + \lambda q_{+1}}{r + \sigma + \lambda q} (e^{a-1} - 1 - a_{-1}) \geq 0. \quad (6.9)$$

It is difficult to verify (6.9) analytically, because the condition involves three variables (a_{-1}, a, a_{+1}). However, it is satisfied in the numerical example in section 5. We conclude this section by summarizing the results on existence in the following proposition.

Proposition 6.7. *Maintain Assumption Regularity and Restriction Off-eqm. An equilibrium with the described wage ladder exists under the conditions in Proposition M-exists, Lemma No-leap and (if $M \geq 3$) Lemma Type-II.*

7. Conclusion

In this paper we have studied the equilibrium in a large labor market where employed workers search on the job and firms direct workers' search intentionally using wage offers and employment probabilities. All applicants observe all posted job openings before the application. There is wage dispersion among workers, despite the fact that all workers (and all jobs) are homogeneous. Moreover, equilibrium wages form a ladder and workers choose to climb up the ladder over time, one rung at a time. This occurs without the familiar elements that prevent workers from jumping

over the ladder, such as a gradual increase in productivity, differential information among workers, and intentional discrimination by firms according to workers' current wages.

Other properties of the wage ladder are as follows: (*i*) a low-wage job is easier to be obtained than a high-wage job, and so a low-wage job experiences a higher quit rate than a high-wage job; (*ii*) The distance between two adjacent rungs on the wage ladder becomes smaller as wage increases; (*iii*) The longer a worker has been employed since his last unemployment, the higher his current wage is on average and, the higher wage a worker had just before becoming unemployed, the longer time he will take to return to the previous wage level; (*iv*) the density of offer wages is a decreasing function; and (*v*) the wage density is decreasing at high wages when the hiring cost is larger relative to the application cost. In particular, properties (*ii*), (*iv*)-(*v*) are not true in the undirected search framework with on-the-job search. On the other hand, directed search models without on-the-job search generate a degenerate wage distribution among homogeneous workers. In fact, our numerical results show that this degeneracy is not robust, because allowing for even a very small probability of on-the-job search generates a non-degenerate wage ladder.

We have illustrated how useful search frictions can be in explaining labor market phenomena. When search is directed, wages can direct workers' applications ex-ante and as a result, both the meeting technology and the division of the surplus are endogenized. When on-the-job search is allowed, the tradeoff between wages and employment (hiring) probability, which is the underlying principle of directed search, leads to a wage ladder equilibrium. We could use that same framework to look at various extensions, such as learning by doing on the job or match specific productivity.²²

²²One way to introduce learning by doing is to assume that workers' productivity depends on their positions on the wage ladder, specifically that each time they move up the ladder, their productivity increases by some increment, and that their skills fully depreciate when they lose their job. This would allow the worker's labor market history to remain tractable, yet carry the flavor of learning by doing. Of course, some of the properties of equilibrium (such as the fact that firms offer the same employment probability to all workers) may not hold anymore, as firms may not be indifferent as to the type of workers it would like to attract.

References

- [1] D. Acemoglu and R. Shimer, Holdups and efficiency with search frictions, *International Economic Review* **40** (1999), 827-850.
- [2] O.J. Blanchard and P.D. Diamond, The Beveridge curve, *Brookings Papers on Economic Activity* **1** (1989), 417-434.
- [3] K. Burdett and K. Judd, Equilibrium price distributions, *Econometrica* **51** (1983), 955-970.
- [4] K. Burdett and D. Mortensen, Wage differentials, employer size, and unemployment, *International Economic Review* **39** (1998), 257-273.
- [5] K. Burdett, S. Shi and R. Wright, Matching and pricing with frictions, *Journal of Political Economy* **109** (2001), 1060-1085.
- [6] M. Cao and S. Shi, Coordination, matching, and wages, *Canadian Journal of Economics* **33** (2000), 1009-1033.
- [7] P. Diamond, Aggregate demand management in search equilibrium, *Journal of Political Economy* **90** (1982), 881-894.
- [8] B. Jovanovic, Job matching and turnover, *Journal of Political Economy* (1979).
- [9] B. Julien, J. Kennes and I. King, Bidding for labor, *Review of Economic Dynamics* **3** (2000), 619-649.
- [10] N.M. Kiefer and G.R. Neumann, Wage dispersion with homogeneity: the empirical equilibrium search model, in: Butzel, et al. (eds), *Panel Data and Labour Market Dynamics* (pp. 57-74). Amsterdam: North Holland, 1993.
- [11] J.D. Montgomery, Equilibrium wage dispersion and interindustry wage differentials, *Quarterly Journal of Economics* **106** (1991), 163-179.
- [12] D. Mortensen, Property rights and efficiency in mating, racing, and related games, *American Economic Review* **72** (1982), 968-979.
- [13] M. Peters, Ex ante price offers in matching games: non-steady state, *Econometrica* **59** (1991), 1425-1454.
- [14] M. Peters, Limits of exact equilibria for capacity constrained sellers with costly search, *Journal of Economic Theory* **95** (2000), 139-168.
- [15] C. Pissarides, *Equilibrium Unemployment Theory*, Basil Blackwell, Cambridge, Massachusetts, 1990.
- [16] C. Pissarides, Search unemployment with on-the-job search, *Review of Economic Studies* **61** (1994), 457-475.
- [17] S. Shi, Frictional assignment, part I, efficiency, *Journal of Economic Theory* **98** (2001), 232-260.
- [18] S. Shi, Product market and the size-wage differential, *International Economic Review* **43** (2002a), 21-54.
- [19] S. Shi, A directed-search model of inequality with heterogeneous skills and skill-biased technology, *Review of Economic Studies* **69** (2002b), 467-491.
- [20] R. Shimer, The assignment of workers to jobs in an economy with coordination frictions, manuscript, Princeton University, 2001.

Appendix

A. Properties of $f(a)$ and $g(a)$

In this appendix we establish the following lemma.

Lemma A.1. Define $f(\cdot)$ as in (4.6) and $g(\cdot)$ as follows:

$$g(a) \equiv (r + \sigma + \lambda q) \left(\frac{1}{f(a)} - \frac{1}{h(a)} \right). \quad (\text{A.1})$$

For all $a > 0$, $f'(a) > 0$, $\frac{d}{da} \left(\frac{f(a)}{a} \right) > 0$, $\frac{d}{da} \left(\frac{1}{f(a)} - \frac{1}{aq(a)} \right) < 0$, and

$$(e^a - 1) \left[a(e^a - 1)^2 - (e^a - 1 - a)^2 \right] - (e^a - 1 - a)^3 > 0. \quad (\text{A.2})$$

Furthermore, if $(r + \sigma)/\lambda > f(a)/a$ then $g'(a) < 0$ for all $a > 0$.

Proof. First, we show that $\frac{d}{da} \left(\frac{f(a)}{a} \right) > 0$ implies $f'(a) > 0$ and $\frac{d}{da} \left(\frac{1}{f(a)} - \frac{1}{aq(a)} \right) < 0$. Since $\frac{d}{da} \left(\frac{f(a)}{a} \right) = \frac{1}{a^2} (af'(a) - f)$, then $\frac{d}{da} \left(\frac{f(a)}{a} \right) > 0$ implies $f'(a) > f(a)/a > 0$. Also, it implies that

$$\begin{aligned} \frac{d}{da} \left(\frac{1}{f(a)} - \frac{1}{aq(a)} \right) &= -\frac{f'}{f^2} + \frac{q+aq'}{(aq)^2} < -\frac{1}{af} + \frac{q+aq'}{(aq)^2} \\ &\sim -aq + \frac{f}{q}(q+aq') \\ &= -(1 - e^{-a}) + (e^a - 1 - a) \left(\frac{1-e^{-a}}{a} - \frac{1-(1+a)e^{-a}}{a} \right) \\ &= -(1 - e^{-a}) + (e^a - 1 - a)e^{-a} = -ae^{-a} < 0, \end{aligned}$$

where the symbol \sim means “having the same sign as”.

Now we show that $\frac{d}{da} \left(\frac{f(a)}{a} \right) > 0$. Substituting $q(a) = \frac{1-e^{-a}}{a}$, we have

$$\frac{f(a)}{a} = \frac{(e^a - 1)(e^a - 1 - a)}{a^2 e^a}.$$

Then, $\frac{d}{da} \left(\frac{f(a)}{a} \right) = \frac{f1(a)}{a^3 e^a}$, where

$$f1(a) = (a - 2)e^{2a} + (a + 4)e^a - (a^2 + 2a + 2).$$

Note that $f1(0) = 0$. Denote the n th order derivative of $f1(a)$ by $f1^{(n)}(a)$. We have:

$$f1^{(1)}(a) = (2a - 3)e^{2a} + (a + 5)e^a - (2a + 2), \text{ with } f1^{(1)}(0) = 0,$$

$$f1^{(2)}(a) = (4a - 4)e^{2a} + (a + 6)e^a - 2, \text{ with } f1^{(2)}(0) = 0,$$

$$f1^{(3)}(a) = e^a[(8a - 4)e^a + a + 7], \text{ with } f1^{(3)}(0) = 3 > 0,$$

$$\frac{d}{da} \left[e^{-a} f1^{(3)}(a) \right] = (8a + 4)e^a + 1 > 0 \text{ for all } a \geq 0.$$

The last two results imply that $f1^{(3)}(a) > f1^{(3)}(0) > 0$ for all $a > 0$, which in turn implies that $f1^{(2)}(a) > f1^{(2)}(0) = 0$, $f1^{(1)}(a) > f1^{(1)}(0) = 0$ and $f1(a) > f1(0) = 0$. Therefore, $\frac{d}{da} \left(\frac{f(a)}{a} \right) > 0$ for all $a > 0$.

We can establish (A.2) using the same procedure. Denote the left-hand side of (A.2) temporarily as $LHS(a)$. Then, $LHS^{(i)}(0) = 0$ for $i = 0, 1, 2$, and

$$LHS^{(3)}(a) = 27(a-1)e^{3a} + (16a+72)e^{2a} - (4a^2+31a+51)e^a + 6.$$

Because $(a-1)e^a + 1 > 0$ for all $a > 0$, $(a-1)e^{3a} > -e^{2a}$. Substituting this result for the first term in $LHS^{(3)}(a)$, we have $LHS^{(3)}(a) > (12a^2+30a-6)e^a + 6$. The last expression has a value 0 at $a = 0$ and a positive derivative for all $a > 0$. Thus, $LHS^{(3)}(a) > 0$. Then, for all $a > 0$, we have $LHS^{(2)}(a) > LHS^{(2)}(0) = 0$, $LHS^{(1)}(a) > LHS^{(1)}(0) = 0$ and $LHS(a) > LHS(0) = 0$.

Finally, we show that $g'(a) < 0$ for all $a > 0$ if $(r+\sigma)/\lambda > f(a)/a$. Compute:

$$g'(a) = (r+\sigma+\lambda q) \frac{d}{da} \left(\frac{1}{f(a)} - \frac{1}{aq(a)} \right) + \lambda \left(\frac{1}{f(a)} - \frac{1}{aq(a)} \right) q'.$$

Because $\frac{d}{da} \left(\frac{1}{f(a)} - \frac{1}{aq(a)} \right) < 0$, as shown above, the condition $(r+\sigma)/\lambda > f(a)/a$ implies

$$g'(a) < \left(\frac{\lambda f}{a} + \lambda q \right) \frac{d}{da} \left(\frac{1}{f} - \frac{1}{aq} \right) + \lambda \left(\frac{1}{f} - \frac{1}{aq} \right) q' \sim g1(a),$$

where $g1(a) = (3-a)e^{3a} - (5a+9)e^{2a} + (7a^2+9a+9)e^a - (2a^3+7a^2+7a+3)$ (correction to $(3-a)e^{3a} - (5a+9)e^{2a} + (7a^2+13a+9)e^a - (2a^3+7a^2+7a+9)$; the calculations should also be corrected accordingly.) Verify that $g1(0) = 0$, $g1^{(1)}(0) = -4 < 0$, $g1^{(2)}(0) = -8 < 0$, $g1^{(3)}(0) = -12 < 0$, and $g1^{(4)}(a) = e^a g2(a)$, where

$$g2(a) = (135-81a)e^{2a} - (80a+304)e^a + 7a^2 + 65a + 129.$$

Verify that $g2(0) = -40 < 0$, $g2^{(1)}(0) = -130 < 0$, $g2^{(2)}(0) = -234 < 0$ and $g2^{(3)}(a) = e^a g3(a)$, where

$$g3(a) = (108-648a)e^a - (80a+544).$$

Verify that $g3(0) = -436 < 0$ and $g3'(a) < 0$ for all $a \geq 0$. Tracing all the way back, we have $g1(a) < 0$ for all $a > 0$ and so $g'(a) < 0$ for all $a > 0$. **QED**

B. Proofs of Proposition Recursive and Lemma 4.2

We prove Proposition Recursive by induction. The argument preceding the proposition in the text has already established (4.11) and (4.12) for $j = 0$. To verify (4.13) and (4.14) for $j = 0$, set $i = M-1$ (and $w^d = w_{M-1}$) in (4.2) to obtain an equation for $J_e(w_{M-1})$. Using this equation and substituting $J_e(w_M)$, we get:

$$J_e(w_M) - J_e(w_{M-1}) = \frac{w_M - w_{M-1} + \lambda S}{r + \sigma + \lambda q_M}.$$

Combining this equation with (4.4) for $i = M-1$, we obtain (4.13) and (4.14) for $j = 0$.

Now suppose that (4.11) – (4.14) hold for an arbitrary $j \in \{0, 1, \dots, M-3\}$. We show that they also hold for $j+1$ and so, by induction, the proposition holds. For $j+1$, (4.11) comes from

setting $i = M - (j + 1)$ in (4.3), and (4.12) from the definitions of $h_{M-(j+1)}$ and $q_{M-(j+1)}$. To verify (4.13) and (4.14) for $j + 1$, set $i = M - j - 2$ (and $w^d = w_{M-j-1}$) in (4.2) to obtain an equation for $J_e(w_{M-j-2})$. Substituting this result, we get:

$$\begin{aligned} & J_e(w_{M-(j+1)}) - J_e(w_{M-(j+2)}) \\ = & \frac{1}{r+\sigma+\lambda q_{M-(j+1)}} \left[(r + \sigma) J_e(w_{M-(j+1)}) - w_{M-(j+2)} - \sigma J_u + \lambda S \right] \\ = & \frac{1}{r+\sigma+\lambda q_{M-(j+1)}} \left[w_M - w_{M-(j+2)} + \lambda S - (r + \sigma) C \sum_{t=0}^j \frac{1}{f_{M-t}} \right]. \end{aligned}$$

The second equality comes from substituting (4.14) for j , which holds by supposition. Combining the above result with (4.4) for $i = M - (j + 1)$, we obtain (4.13) and (4.14) for $j + 1$.

Finally, the zero-profit condition (4.3) must hold for a firm posting w_1 . By the above derivation, this implies that (4.11) and (4.12) must hold for $j = M - 1$. In contrast, (4.13) and (4.14) need be modified for $j = M - 1$. By definition, $w_0 = b$ and $J_e(w_0) = J_u$. To derive (4.15), use the wage ladder to simplify (2.11) as $rJ_u = b - \lambda_0 S + \lambda_0 q_1 [J_e(w_1) - J_u]$. Substituting $[J_e(w_1) - J_u]$ from (4.4) (with $i = 1$), we obtain (4.15). This completes the proof of Proposition Recursive.

Now, we prove Lemma 4.2. Because M is unknown, we start with an arbitrary but sufficiently large integer m and set $h_m = h^*$, where h^* is any value that satisfies (4.10). Compute the sequence $(a_{m-t})_{t \geq 0}$ according to Proposition Recursive and define

$$\delta_i(m) = \frac{w_m - b + \lambda_0 S}{r + \sigma} - \frac{C \lambda_0 q_i}{(r + \sigma) f_i} - C \sum_{t=0}^{m-i} \frac{1}{f_{m-t}}.$$

Note that $\delta_1(M) = \Delta(M, h^*)$, where Δ is given by (4.18). By Proposition Monotone in the text (proven later), $a_{m-t-1} < a_{m-t} \leq \bar{a} < \infty$ for all $t \geq 0$. Since $1/f_{m-t}$ and q_{m-t}/f_{m-t} are both decreasing functions of a_{m-t} , and $a_i < a_{i+1}$, we get:

$$\delta_{i+1}(m) - \delta_i(m) = \frac{C \lambda_0}{r + \sigma} \left(\frac{q_i}{f_i} - \frac{q_{i+1}}{f_{i+1}} \right) + \frac{C}{f_i} > \frac{C}{f_i} \geq \frac{C}{f(\bar{a})}.$$

Because $C/f(\bar{a})$ is bounded away above 0, the sequence δ_i decreases by a strictly positive amount each time when i decreases. If $\delta_m(m) \geq 0$, then there exists i^* such that $\delta_i \leq 0$ for all $i \leq i^*$ and $\delta_i > 0$ for all $i \geq i^* + 1$. In this case, let $M^* = m - (i^* - 1)$ and compute the sequence $\{\delta_i(M^*)\}$ by setting $h_{M^*} = h^*$. Then, $\delta_i(M^*) \leq 0$ for all $i \leq 1$ and $\delta_i(M^*) > 0$ for all $i \geq 2$. Moreover, for any integer $M' \neq M^*$, the sequence $\{\delta_i(M')\}$ computed by setting $h_{M'} = h^*$ satisfies $\delta_1(M') = \delta_{(M^* - M' + 1)}(M^*)$. From the properties of the sequence $\{\delta_i(M^*)\}$, we have $\delta_1(M') > 0$ for all $M' \leq M^* - 1$ and $\delta_1(M') < 0$ for all $M' \geq M^* + 1$. This is the result in Lemma 4.2.

Finally, we show that $\delta_m(m) \geq 0$, indeed. After substituting $w_m = y - C(r + \sigma)/h_m$, the condition $\delta_m(m) \geq 0$ becomes

$$b \leq y + \lambda_0 S - \frac{C [(r + \sigma) e^{a_m} + \lambda_0]}{e^{a_m} - 1 - a_m}.$$

The right-hand of this inequality is an increasing function of a_m . Because a_m is bounded from below by $\bar{a} - \ln(1 + \bar{a})$ according to (4.10), a sufficient condition for the above inequality is that it holds for this lower bound of a_m , which is imposed as (4.19). **QED**

C. Proof of Proposition Monotone

First, we verify (5.3) and (5.4) by induction. To begin, we show that these conditions hold for $j = 0$. By (4.10), $a_M \leq \bar{a}$. Since $h(a)$ is an increasing function and $q(a)$ a decreasing function, all three inequalities in (5.3) are equivalent to each other, and so we show $h_{M-1} < h_M$ only. Since (5.2) holds for $j = 0$ after replacing the term q_{+1} by 0 and $\lambda q_{+1}/f_{+1}$ by $\lambda S/C$, $h_{M-1} < h_M$ if and only if

$$0 < \frac{r + \sigma + \lambda q_M}{f_M} - \frac{\lambda S}{C} - \frac{\lambda q_M}{h_M} = \lambda \left(\frac{1}{e^{a_M} - 1 - a_M} - \frac{1}{e^{\bar{a}} - 1 - \bar{a}} \right) + \left(\frac{r + \sigma}{f_M} - \frac{\lambda}{a_M} \right),$$

where we have used the definition of \bar{a} in (4.7) to replace S/C . Because $a_M \leq \bar{a}$ by construction (see (4.10)) and $(e^a - 1 - a)$ is an increasing function, the term in the first (\cdot) is positive. Also, $f(a)/a$ is an increasing function, as shown in Appendix A, and so Assumption Regularity implies $(r + \sigma)/\lambda > f(\bar{a})/\bar{a} \geq f_M/a_M$. That is, the term in the second (\cdot) above is also positive. Thus, $h_{M-1} < h_M$, verifying (5.3) for $j = 0$.

Now that $a_{M-1} < a_M \leq \bar{a}$, and that $f(a)/a$ is an increasing function of a , (4.20) implies $(r + \sigma)/\lambda > f(a_{M-1})/a_{M-1}$. That is, (5.4) holds for $j = 0$.

Suppose that (5.3) and (5.4) hold for an arbitrary $j \in \{1, 2, \dots, M-3\}$. We show that they hold for $j+1$. For (5.3), this amounts to proving $h_{-2} < h_{-1}$. Computing h_{-2} using (5.2), $h_{-2} < h_{-1}$ if and only if

$$0 < \frac{r + \sigma + \lambda q}{h_{-1}} - \frac{\lambda q}{f} + (r + \sigma + \lambda q_{-1}) \left(\frac{1}{f_{-1}} - \frac{1}{h_{-1}} \right).$$

Because $h_{-1} < h$ by supposition, a sufficient condition for the above inequality is:

$$0 < \frac{r + \sigma + \lambda q}{h} - \frac{\lambda q}{f} + (r + \sigma + \lambda q_{-1}) \left(\frac{1}{f_{-1}} - \frac{1}{h_{-1}} \right).$$

The last term is equal to $g(a_{-1})$. In Appendix A we showed that $g'(a) < 0$ if $(r + \sigma)/\lambda > f(a)/a$. Because $(r + \sigma)/\lambda > f(\bar{a})/\bar{a}$, we have $g'(a) < 0$ for all $a \leq \bar{a}$. Since $a_{-1} < a$ by supposition and $a \leq \bar{a}$, then $g(a_{-1}) > g(a)$. Thus,

$$\begin{aligned} & \frac{r + \sigma + \lambda q}{h} - \frac{\lambda q}{f} + g(a_{-1}) \\ & > \frac{r + \sigma + \lambda q}{h} - \frac{\lambda q}{f} + (r + \sigma + \lambda q) \left(\frac{1}{f} - \frac{1}{h} \right) = (r + \sigma)/f(a) > 0. \end{aligned}$$

That is, (5.3) holds for $j+1$. This in turn implies $a_{-2} < a_{-1}$. Because $f(a)/a$ is an increasing function of a (see Appendix A), the supposition $(r + \sigma)/\lambda > f(a_{-1})/a_{-1}$ implies $(r + \sigma)/\lambda > f(a_{-2})/a_{-2}$. That is, (5.4) also holds for $j+1$. By induction, (5.3) and (5.4) hold for all $j \in \{0, 1, \dots, M-2\}$.

Second, we prove (5.5), which is equivalent to $h_{-1} > h(a - \ln(1+a))$. By (5.2), this in turn is equivalent to:

$$\begin{aligned} 0 & > \frac{r + \sigma + \lambda q_{+1}}{h} - \frac{\lambda q_{+1}}{f_{+1}} + (r + \sigma + \lambda q) \left(\frac{1}{f} - \frac{1}{h(a - \ln(1+a))} \right) \\ & = \frac{r + \sigma + \lambda q_{+1}}{h} - \frac{\lambda q_{+1}}{f_{+1}} - \frac{r + \sigma + \lambda q}{h} = -\lambda \left(\frac{q - q_{+1}}{h} + \frac{q_{+1}}{f_{+1}} \right). \end{aligned}$$

The equalities follow from calculating f and $h(a - \ln(1 + a))$ explicitly. Because $q > q_{+1}$, the above inequality clearly holds, and so does (5.5).

Finally, we show that $da/dh_M > 0$ and $dw/dh_M > 0$ for any h_M that satisfies (4.10). From (4.13) it is easy to see that $da/dh_M > 0$ implies $dw/dh_M > 0$; so, we need to prove only $da/dh_M > 0$. Because $a_M = -\ln(1 - h_M)$, it is obvious that $da_M/dh_M > 0$. If $da_{+t}/dh_M \geq 0$ for all $t \geq 1$ implies $da/dh_M > 0$, then by induction, $da/dh_M > 0$. Suppose that $da_{+t}/dh_M \geq 0$ for all $t \geq 1$. By construction, $h = \lambda C \left(\frac{r+\sigma}{\lambda} + q \right) / (y - w)$. Totally differentiating this relationship with respect to h_M (where dw/dh_M can be calculated using (4.13)), we have

$$\begin{aligned} \frac{y-w}{\lambda C h} h' \frac{da}{dh_M} &= \frac{1}{\lambda C} \frac{dw_M}{dh_M} + \frac{r+\sigma}{\lambda} \sum_{t=2}^j \frac{f'_{+t}}{f_{+t}^2} \left(\frac{da_{+t}}{dh_M} \right) \\ &+ \left[q'_{+1} \left(\frac{1}{h} - \frac{1}{f_{+1}} \right) + \left(\frac{r+\sigma}{\lambda} + q_{+1} \right) \frac{f'_{+1}}{f_{+1}^2} \right] \left(\frac{da_{+1}}{dh_M} \right). \end{aligned}$$

Because $w_M = y - (r + \sigma)C/h_M$, $dw_M/dh_M > 0$. Because $da_{+t}/dh_M \geq 0$ for all $t \geq 1$, a sufficient condition for $da/dh_M > 0$ is that the following inequality holds for all j :

$$q'(a) \left(\frac{1}{h_{-1}} - \frac{1}{f} \right) + \left(\frac{r + \sigma}{\lambda} + q \right) \frac{f'(a)}{f^2} > 0.$$

To verify this inequality, temporarily denote the left-hand side of the inequality by LHS . Because $a_{-1} > a - \ln(1 + a)$, $q' < 0$, $(r + \sigma)/\lambda > f/a$ and $f' > 0$, we have

$$LHS > q'(a) \left(\frac{1}{h(a - \ln(1 + a))} - \frac{1}{f} \right) + \left(\frac{f}{a} + q \right) \frac{f'(a)}{f^2}.$$

After substituting (q, f, q', f') , the right-hand side of this inequality has the same sign as the expression, $(e^a - 1) \left[a(e^a - 1)^2 - (e^a - 1 - a)^2 \right] - (e^a - 1 - a)^3$, which is positive for all $a > 0$ as shown in Lemma A.1 in Appendix A. Thus, the required condition $LHS > 0$ holds. **QED**

D. Proofs of Propositions W-property and W-density

We prove Proposition W-property first. Using (4.13) and (4.14), it is easy to verify (i) in the Proposition. Property (ii) holds because $E(w_{-1}) = Cq/f > Cq_{+1}/f_{+1} = E(w)$ and $E(w_{M-1}) \geq S$ (see (4.9) or equivalently the first part of (4.10)).

To establish (iii), use (5.1) to rewrite it as

$$\frac{R + q}{f} - \frac{R + q_{+1}}{f_{+1}} - \frac{q_{+1}}{f_{+1}} + \frac{q_{+2}}{f_{+2}} > 0,$$

where $R = (r + \sigma)/\lambda$. For the computed sequence to be an equilibrium we need $q > q_{+1} (1 + f/f_{+1})$ (see (6.2)), as shown in section 6. Under this condition, the left-hand side of the above inequality is greater than the following expression:

$$(R + q_{+1}) \left(\frac{1}{f} - \frac{1}{f_{+1}} \right) + \frac{q_{+2}}{f_{+2}}.$$

This is clearly positive, because $a < a_{+1}$ and $f(\cdot)$ is an increasing function. Thus, Proposition W-property holds.

To prove Proposition W-density, recall that the density of offer wages is (v_i) and of employed wages $(n_i/(1-u))$, where $i = 1, 2, \dots, M$. So, the density of offer wages is a decreasing function iff $v_{-1} > v$ and the density of employed wages is decreasing iff $n_{-1} > n$. By (4.16) and (4.17), $n_{-1}/n = (\frac{\sigma}{\lambda} + q_{+1})/q$ and $v_{-1}/v = (n_{-2}a)/(n_{-1}a_{-1})$ for all $i \geq 3$. For all $i \geq 3$, we have:

$$\frac{v_{-1}}{v} = \left(a \frac{\sigma}{\lambda} + h \right) / h_{-1} > \frac{h}{h_{-1}} > 1.$$

Similarly, the result holds for $i = 2$; i.e., $v_1/v_2 > h_2/h_1 > 1$.

The density of employed wages is a decreasing function iff $\sigma/\lambda > q - q_{+1}$. Because $q_{M+1} = 0$, the density of employed wages is decreasing at the upper end of the wage support (i.e. $n_{M-1} > n_M$) iff $\sigma/\lambda > q_M$. Because $q(\cdot)$ is a decreasing function and $a_M \geq \bar{a} - \ln(1 + \bar{a})$ by (4.10), a sufficient condition for $n_{M-1} > n_M$ is $\sigma/\lambda > q(\bar{a} - \ln(1 + \bar{a}))$, which can be rewritten as (5.7). When r is sufficiently close to 0, this condition is satisfied iff $(r + \sigma)/\lambda > q(\bar{a} - \ln(1 + \bar{a}))$. Because $(r + \sigma)/\lambda \geq f(\bar{a})/\bar{a}$ by Assumption Regularity, (5.7) is satisfied if $f(\bar{a})/\bar{a} > q(\bar{a} - \ln(1 + \bar{a}))$, which is equivalent to $\bar{a} > 1.605$ and hence to $C/S > 2.373$. Similarly, because $a_M \leq \bar{a}$ by (4.10), a sufficient condition for $n_{M-1} < n_M$ is $\sigma/\lambda < q(\bar{a})$. **QED**

E. Proof of Lemma 6.2

To show that $\phi(a_{+1})$ is well-defined for each a_{+1} by the equality form of (6.2), we use the definition of f to rewrite the equality as

$$\frac{q(a)}{q(a_{+1})} \left[1 - \frac{e^a - 1 - a}{e^{a_{+1}} - 1 - a_{+1}} \right] = 1. \quad (\text{E.1})$$

The left-hand side of (E.1) is a decreasing function of a and an increasing function of a_{+1} (note that $a_{+1} > a$). If $\phi(a_{+1})$ is a solution for a , then the solution is unique and satisfies $\phi' > 0$, verifying part (i) of the lemma. When $a = a_{+1}$, the left-hand side of (E.1) is 0, which is less than the right-hand side. When $a \rightarrow 0$, the left-hand side approaches $1/q(a_{+1}) > 1$. Thus, the solution for a , $\phi(a_{+1})$, indeed exists and is unique. This argument also establishes the inequality $\phi(a_{+1}) < a_{+1}$ in part (ii) of the lemma.

For the inequality $\phi(a_{+1}) > a_{+1} - \ln(1 + a_{+1})$ in part (ii), we show that the left-hand side of (E.1) is greater than 1 (the right-hand side) when $a = a_{+1} - \ln(1 + a_{+1})$. Substituting this particular value of a and re-arranging terms, the condition to be established becomes $\ln(1 + a_{+1}) - \frac{a_{+1}}{1 + a_{+1}} > 0$. The left-hand side of this inequality is equal to 0 when $a_{+1} = 0$, and its derivative with respect to a_{+1} is $a_{+1}/(1 + a_{+1})^2 > 0$. Thus, the desired inequality holds for all $a_{+1} > 0$.

Before establishing part (iii), we claim that the following inequalities hold:

$$\frac{d}{da} \left[\frac{1}{f(a)} - \frac{1}{h(\phi(a))} \right] \leq 0, \quad (\text{E.2})$$

$$\left[\frac{f(a)}{a} + q(a) \right] \frac{d}{da} \left[\frac{1}{f(a)} - \frac{1}{h(\phi(a))} \right] + q'(a) \left[\frac{1}{f(a)} - \frac{1}{h(\phi(a))} \right] \leq 0. \quad (\text{E.3})$$

Because the expressions in these conditions are single-variable functions which do not have any parameter, we can graph them using a computer and show that the inequalities hold, indeed. (However, it is difficult to prove them using pen and paper.)

Now, suppose $a \leq \phi(a_{+1})$. We show $a_{-1} < \phi(a)$ or, equivalently, $h_{-1} < h(\phi(a))$. Under (5.2), this desired condition is equivalent to:

$$\frac{r + \sigma + \lambda q_{+1}}{h} - \frac{\lambda q_{+1}}{f_{+1}} + (r + \sigma + \lambda q) \left(\frac{1}{f} - \frac{1}{h(\phi(a))} \right) > 0.$$

Because $a \leq \phi(a_{+1})$ by supposition, $h \leq h(\phi(a_{+1}))$, and so a sufficient condition for the above inequality is

$$\frac{r + \sigma + \lambda q_{+1}}{h(\phi(a_{+1}))} - \frac{\lambda q_{+1}}{f_{+1}} + (r + \sigma + \lambda q) \left(\frac{1}{f} - \frac{1}{h(\phi(a))} \right) > 0.$$

Under (E.2) and (E.3), we have

$$\begin{aligned} & \frac{d}{da} \left\{ (r + \sigma + \lambda q) \left[\frac{1}{f} - \frac{1}{h(\phi(a))} \right] \right\} \\ &= (r + \sigma + \lambda q) \frac{d}{da} \left[\frac{1}{f} - \frac{1}{h(\phi(a))} \right] + \lambda q'(a) \left[\frac{1}{f} - \frac{1}{h(\phi(a))} \right] \\ &< \lambda \left(\frac{f}{a} + q \right) \frac{d}{da} \left[\frac{1}{f} - \frac{1}{h(\phi(a))} \right] + \lambda q'(a) \left[\frac{1}{f} - \frac{1}{h(\phi(a))} \right] \\ &\leq 0. \end{aligned}$$

The first inequality comes from (E.2) and the result $(r + \sigma)/\lambda > f(a)/a$ in Proposition Monotone, and the second inequality from (E.3). Because $a < a_{+1}$, the above result implies

$$\begin{aligned} & \geq \left[\frac{r + \sigma + \lambda q_{+1}}{h(\phi(a_{+1}))} - \frac{\lambda q_{+1}}{f_{+1}} \right] + (r + \sigma + \lambda q) \left[\frac{1}{f} - \frac{1}{h(\phi(a))} \right] \\ &= \frac{r + \sigma}{f_{+1}} > 0. \end{aligned}$$

This is the desired result. **QED**

F. Proof of Lemma Type-II

Consider a type-II deviation $w^d \in (w_{-1}, w)$. This deviation induces the value $J_f^d(w^d)$ to the firm and $J_e^d(w^d)$ to the worker who gets the job, where J_f^d is given by (6.6) and J_e^d by (6.7). Suppose that the deviation is profitable. Then it must satisfy the following conditions:

- (IIa) By applying to w^d , a w_{-2} -applicant's expected surplus is equal to $E(w_{-2})$;
- (IIb) The deviating firm earns an expected surplus greater than C ;
- (IIc) A w^d -applicant's future application is indeed to w_{+1} instead of w .

Under (IIa), $q^d [J_e^d(w^d) - J_e(w_{-2})] = E(w_{-2}) = Cq_{-1}/f_{-1}$. That is,

$$\begin{aligned} \frac{q_{-1}}{q^d f_{-1}} &= \frac{1}{C} [J_e^d(w^d) - J_e(w_{-2})] \\ &= \frac{1}{C} [J_e(w) - J_e(w_{-1})] + \frac{1}{C} [J_e(w_{-1}) - J_e(w_{-2})] + \frac{1}{C} [J_e^d(w^d) - J_e(w)] \\ &= \frac{1}{f} + \frac{1}{f_{-1}} - \frac{(w-w^d)/C}{r+\sigma+\lambda q_{+1}}. \end{aligned}$$

The last equality comes from substituting $(J_e^d(w^d), J_e(w))$ and using (4.4). Solve for w^d :

$$w^d = w - C(r + \sigma + \lambda q_{+1}) \left(\frac{1}{f} + \frac{1}{f_{-1}} - \frac{q_{-1}}{q^d f_{-1}} \right). \quad (\text{F.1})$$

The deviating firm's expected surplus is $\pi(a^d) = h^d J_f^d(w^d)$. Substituting $(J_f^d(w^d), w^d)$ and using (4.3) for $(y - w)$, this surplus becomes:

$$\pi(a^d) = \frac{C [(y - w) + (w - w^d)]}{r + \sigma + \lambda q_{+1}} = C \left[h^d \left(\frac{1}{h} + \frac{1}{f} + \frac{1}{f_{-1}} \right) - \frac{a^d q_{-1}}{f_{-1}} \right].$$

The expected surplus $\pi(a^d)$ is maximized at $a^d = A$ that solves $\pi'(A) = 0$. So,

$$A = \ln \left[\frac{1}{q_{-1}} \left(1 + f_{-1} \left(\frac{1}{h} + \frac{1}{f} \right) \right) \right]. \quad (\text{F.2})$$

Because $a_{-1} > a - \ln(1 + a)$ by Proposition Monotone, it can be shown that $A > a_{-1}$. The unconstrained maximum of $\pi(a^d)$ is:

$$\pi(A) = C (e^A - 1 - A) / (e^{a_{-1}} - 1 - a_{-1}) > C.$$

Thus, a type II deviation is not profitable if and only if the constraint (IIc) keeps a^d a sufficient distance away from A .

The constraint (IIc) requires $q_{+1} [J_e(w_{+1}) - J_e^d(w^d)] \geq q [J_e(w) - J_e^d(w^d)]$. Using (4.4) to substitute $J_e(w_{+1})$, noting that $J_e(w) - J_e^d(w^d) = (w - w^d)/(r + \sigma + \lambda q_{+1})$, and substituting w^d from (F.1), we can rewrite (IIc) as follows:

$$q^d \leq q_{-1} / \left[1 + f_{-1} \left(\frac{1}{f} - \frac{q_{+1}}{(q - q_{+1}) f_{+1}} \right) \right]. \quad (\text{F.3})$$

Let β be the level of a^d that satisfies (F.3) as equality. Since $q^d = q(a^d)$ is a decreasing function of a^d , (F.3) is equivalent to $a^d \geq \beta$.

Because A is the unique maximizer of $\pi(a^d)$ and $\pi(A) > C$, there exist A_1 and A_2 , with $A \in (A_1, A_2)$, such that $\pi(A_i) = C$, for $i = 1, 2$. Clearly, $\pi(a^d) > C$ iff $a^d \in (A_1, A_2)$, and $\pi'(A_1) > 0 > \pi'(A_2)$. Because a type II deviation must satisfy $a^d \geq \beta$ (i.e., the constraint (IIc)), the deviation is not profitable if and only if either $\beta \geq A_2$ or $\beta \leq a^d \leq A_1$. In the remainder of this proof, we rewrite these conditions to obtain the condition (6.9) in the lemma. Let us denote $Y = q_{+1}/[(q - q_{+1}) f_{+1}]$ in this appendix.

First, we show that $\beta > A_1$, and so the case $\beta \leq a^d \leq A_1$ never occurs. The inequality $\beta > A_1$ holds iff $q(\beta) < q(A_1)$ and hence iff

$$q(A_1) > \frac{q_{-1}}{1 + f_{-1} \left(\frac{1}{f} - Y \right)} = \frac{q(A_1) \left[1 + f_{-1} \left(\frac{1}{h} + \frac{1}{f} \right) \right] - f_{-1}/A_1}{1 + f_{-1} \left(\frac{1}{f} - Y \right)}.$$

Here we have used the definition of $q(\beta)$ first and then the definition of A_1 to substitute for q_{-1} . Re-arranging terms and using the definition of β^* in (6.8), the above inequality is equivalent to $h(A_1) < h(\beta^*)$. So, $\beta > A_1$ is equivalent to $\beta^* > A_1$. Because $a < \phi(a_{+1})$, $Y < 1/f$ and so

$$h(\beta^*) > \left(\frac{1}{h} + \frac{1}{f} \right)^{-1} = 1 - (1 + a) e^{-a}.$$

A sufficient condition for $\beta^* > A_1$ is then $A_1 < a - \ln(1 + a)$. Because $a - \ln(1 + a) < a_{-1}$ by Proposition Monotone and $a_{-1} < A$ as shown in the text, $a - \ln(1 + a) < A$. Because $\pi'(a^d) > 0$ for all $a^d < A$ and $\pi(A_1) = C$, then $A_1 < a - \ln(1 + a)$ iff $\pi(a - \ln(1 + a)) > C$. Calculating $\pi(a - \ln(1 + a))$ and re-arranging terms, the latter condition becomes $q(a - \ln(1 + a)) > q_{-1}$, which is satisfied because $q(\cdot)$ is a decreasing function and $a - \ln(1 + a) < a_{-1}$. Now that $\beta > A_1$, a type II deviation is not profitable iff $\beta \geq A_2$.

Second, we show that $\beta \geq A_2$ iff $\beta^* \geq \beta$. Similar to the above procedure that showed $\beta > A_1$ iff $\beta^* > A_1$, we can show that $\beta \geq A_2$ iff $\beta^* \geq A_2$. Because $\beta^* > A_1$, as shown above, and $\pi(A_2) = C$, the inequality $\beta^* \geq A_2$ holds iff $\pi(\beta^*) \leq C$. Substituting $\pi(\beta^*)$, we rewrite the latter condition as

$$\begin{aligned} 0 &\leq \frac{1}{h(\beta^*)} - \left(\frac{1}{h} + \frac{1}{f} + \frac{1}{f_{-1}} \right) + \frac{\beta^* q_{-1}}{h(\beta^*) f_{-1}} \\ &= \left(\frac{1}{h} + Y \right) - \left(\frac{1}{h} + \frac{1}{f} + \frac{1}{f_{-1}} \right) + \frac{q_{-1}}{q(\beta^*) f_{-1}} \\ &= \frac{q_{-1}}{q(\beta^*) f_{-1}} - \left(\frac{1}{f_{-1}} + \frac{1}{f} - Y \right). \end{aligned}$$

Using the equation that defines β to substitute for q_{-1} , we can rewrite the above inequality further as $q(\beta^*) \leq q(\beta)$. Thus, $\beta \geq A_2$ holds iff $\beta^* \geq \beta$.

Finally, we show that $\beta^* \geq \beta$ is equivalent to (6.9). To do so, rewrite (5.2) as

$$\frac{1}{f} - Y = \frac{1}{h_{-1}} - \frac{r + \sigma + \lambda q_{+1}}{r + \sigma + \lambda q} \left(\frac{1}{h} + Y \right).$$

Then, $\beta^* \geq \beta$ iff $1/q(\beta^*) \geq 1/q(\beta)$, and hence iff

$$\begin{aligned} 0 &\leq \frac{1}{q(\beta^*)} - \frac{1}{q_{-1}} \left\{ 1 + f_{-1} \left[\frac{1}{h_{-1}} - \frac{r + \sigma + \lambda q_{+1}}{r + \sigma + \lambda q} \left(\frac{1}{h} + Y \right) \right] \right\} \\ &= \frac{1}{q(\beta^*)} - \frac{1}{q_{-1}} \left(1 + \frac{f_{-1}}{h_{-1}} \right) + \left(\frac{r + \sigma + \lambda q_{+1}}{r + \sigma + \lambda q} \right) \frac{f_{-1}}{q_{-1} h(\beta^*)} \\ &= \frac{1}{q(\beta^*)} - e^{a_{-1}} + \left(\frac{r + \sigma + \lambda q_{+1}}{r + \sigma + \lambda q} \right) (e^{a_{-1}} - 1 - a_{-1}) / h(\beta^*). \end{aligned}$$

The inequality comes from substituting the definition of β and the term $\left(\frac{1}{f} - Y \right)$; the two equalities come from substituting the definitions of $h(\beta^*)$ and f . Multiplying the last expression by $h(\beta^*)$ yields (6.9). **QED**

Supplementary Appendices for the Manuscript:
“Directed Search On the Job and the Wage Ladder”

by

Alain Delacroix and Shouyong Shi

G. Simplifying a Recruiting Firm’s Strategy

In the text of the paper, we claimed that a recruiting firm’s decision can be modelled as announcing a wage level and an ex ante employment probability, rather than ex post selection criteria that the firm will use to select one worker out of the received applicants. We also claimed that the optimal choice of employment probabilities must be the same for all types of applicants. In this supplementary appendix, we prove these results.

Consider a firm posting wage w and examine its ex post selection criteria. Let $R(w')$ be the number of w' -applicants whom this firm will receive and denote $R = (R(w'))_{w'}$ as the vector of such numbers. The firm’s selection rule is a matrix of probabilities, $(Z(w', R))_{w', R}$; The element $Z(w', R)$ describes the probability with which the selected worker is a w' -worker, conditional on the composition of received applicants R . Conditional on R and that the selected worker is a w' -applicant, each particular w' -applicant who applied to the firm is chosen with probability $Z(w', R)/R(w')$. Clearly, the firm cannot select a w' -worker if no such applicant has applied to the firm, and the selection probabilities must add up to one if the firm has received one or more applicant. That is, the following restrictions must hold:

$$Z(w', R) = 0 \text{ if } R(w') = 0, \tag{G.1}$$

$$\sum_{w'} Z(w', R) = 1 \text{ if } R \neq 0, \text{ and } 0 \text{ otherwise.} \tag{G.2}$$

In addition, we exclude lexicographic selection rules by imposing the following restriction:

$$\begin{aligned} &\text{If } Z(w', R) > 0 \text{ for a particular } R \text{ such that } R(w') > 0, \\ &\text{then } Z(w', R) > 0 \text{ for all such } R \text{ that } R(w') > 0. \end{aligned} \tag{G.3}$$

That is, if a firm selects a type of applicants with positive probability in some cases, then it must select such applicants with positive probability in all cases, provided that they show up at the match. The reason for imposing this restriction is that lexicographic selection rules may induce pure strategy equilibria which feature implicit coordination. Note that the firm can rank the applicants probabilistically. In particular, we do *not* restrict the selection probabilities to be equal for all types of received applicants, although this is a result we will establish later.

Let $q(w, w')$ denote the (ex ante) employment probability that the firm recruiting at wage w offers to each w' -applicant. Do not impose $q(w, w') = q(w)$, because this is a result we need to prove in this supplementary appendix.

Proposition G.1. *Suppose that firms announce (w, Z) , where Z satisfies (G.1) – (G.3). Agents’ decision problems can be formulated equivalently using $(q(w, w'))_{w'}$ in lieu of Z . Moreover, the*

following relationship holds:

$$h(w) = \sum_{w'} q(w, w') a(w, w'). \quad (\text{G.4})$$

Proof. Suppress the particular firm's wage offer w in various notation. Let \mathcal{R} be the set containing the realizations of R , which is the composition of applicants whom the firm receives. Let $\Gamma(\cdot)$ be the distributional function of R , with a density $\gamma(\cdot)$. Because a w' -applicant applies to the firm with probability $p(w')$ and the number of such applicants is $\lambda(w')n(w')L$, we have:

$$\gamma(R) = \prod_{w'} \left[\binom{R(w')}{\lambda(w')n(w')L} [p(w')]^{R(w')} [1 - p(w')]^{\lambda(w')n(w')L - R(w')} \right],$$

where the expression in $[\cdot]$ is the probability with which the firm receives exactly a number $R(w')$ of w' -applicants.

We reformulate the firm's problem. Under (w, Z) , the firm's objective function, i.e., the expected surplus, is as follows:

$$\sum_{w^*} \sum_{R \in \mathcal{R}} Z(w^*, R) [J_f(w) - J_v] \gamma(R) = [J_f(w) - J_v] \sum_{w^*} \sum_{R \in \mathcal{R}} Z(w^*, R) \gamma(R).$$

The equality follows from the fact that the wage to be paid is independent of the type of the applicant whom the firm will select ex post and of the realization of R . The double summation in the last expression is the firm's hiring probability, as demonstrated below:

$$\begin{aligned} \sum_{w^*} \sum_{R \in \mathcal{R}} Z(w^*, R) \gamma(R) &= \sum_{R \in \mathcal{R}} \gamma(R) \left[\sum_{w^*} Z(w^*, R) \right] \\ &= \sum_{R \in \mathcal{R} \setminus \{0\}} \gamma(R) \left[\sum_{w^*} Z(w^*, R) \right] = \sum_{R \in \mathcal{R} \setminus \{0\}} \gamma(R) = 1 - \Gamma(0) = h. \end{aligned} \quad (\text{G.5})$$

The first equality comes from switching the order of the two summations. The second and third equalities come from (G.2), i.e., that the sum of $Z(w^*, R)$ over w^* is 0 when $R = 0$ and 1 when $R \neq 0$. The fourth equality comes from the definition of $\Gamma(0)$ and the last equality from the meaning of the hiring probability.

Therefore, the firm's objective function is $h[J_f(w) - J_v]$, as in subsection 2.2. Suppose that (G.4) holds, which we will established later. Then, (G.4) and (2.3) implies:

$$1 - \exp \left[- \sum_{w'} a(w, w') \right] - \sum_{w'} q(w, w') a(w, w') = 0. \quad (\text{G.6})$$

This is a general version of (2.4). We can formulate the firm's decision problem as choosing (a, q, w) to solve:

$$\begin{aligned} (\mathcal{P}') \quad & \max h(w) [J_f(w) - J_v] \\ & \text{s.t. (2.3), (G.6), and (2.6),} \end{aligned}$$

where in (2.6), $q(w)$ should be replaced by $q(w, w')$. As claimed, this problem involves the ex ante employment probabilities rather than the ex post selection criteria.

Next, we reformulate the applicant's decision problem. Consider a particular w^* -applicant who contemplates applying to the firm. Shorten the notation $q(w, w^*)$ to $q(w^*)$, which stands

for the applicant's ex ante probability of getting the job. (Note that $q(w^*)$ here is not the short form for $q(w^*, w)$ as we used in the text.) To calculate q , we need the distribution function of the composition of the firm's received applicants other than the particular w^* -applicant in discussion. This is the distribution function of R conditional on that the particular w^* -applicant applies to the firm, and hence it is different from the unconditional distribution Γ . Denote the composition of the firm's received applicants other than the particular w^* -applicant by R^- . Let \mathcal{R}^- be the set of vectors containing the possible values of R^- and $\gamma^-(\cdot, w^*)$ the density function of R^- . Conditional on that the particular w^* -applicant applies to the firm, the composition of applicants whom the firm receives is R , where $R(w') = R^-(w')$ if $w' \neq w^*$ and $R(w^*) = R^-(w^*) + 1$ otherwise. With a realization of R^- in addition to the particular w^* -applicant, the firm chooses a w^* -applicant with probability $Z(w^*, R)$, out of which the particular w^* -applicant is the chosen one with probability $1/R(w^*)$. Therefore,

$$q(w^*) = \sum_{R^- \in \mathcal{R}^-} \frac{Z(w^*, R)}{R(w^*)} \gamma^-(R^-, w^*). \quad (\text{G.7})$$

When applying to the job, a w^* -applicant's expected surplus is

$$\begin{aligned} & \sum_{R^- \in \mathcal{R}^-} \frac{Z(w^*, R)}{R(w^*)} [J_e(w) - J_e(w^*)] \gamma^-(R^-, w^*) \\ &= [J_e(w) - J_e(w^*)] \sum_{R^- \in \mathcal{R}^-} \frac{Z(w^*, R)}{R(w^*)} \gamma^-(R^-, w^*) \\ &= q(w^*) [J_e(w) - J_e(w^*)]. \end{aligned}$$

The first equality follows from the fact that the wage the applicant gets is independent of the realization of R^- , and the second equality from (G.7). The above result shows that what matters to an applicant is the wage offer and the ex ante employment probability q , as we claimed.

Finally, we show that h satisfies (G.4). To do so, we find the relationship between the two densities, γ^- and γ . Compute

$$\begin{aligned} \gamma^-(R^-, w^*) &= \prod_{w' \neq w^*} \left[\binom{R(w')}{\lambda(w')n(w')L} [p(w')]^{R(w')} [1 - p(w')]^{\lambda(w')n(w')L - R(w')} \right] \\ &\quad \times \binom{R^-(w^*)}{\lambda(w^*)n(w^*)L - 1} [p(w^*)]^{R^-(w^*)} [1 - p(w^*)]^{\lambda(w^*)n(w^*)L - 1 - R^-(w^*)}. \end{aligned}$$

Here we have isolated w^* -applicants and used $R(w') = R^-(w')$ for all $w' \neq w^*$. Rewrite:

$$\begin{aligned} & \binom{R^-(w^*)}{\lambda(w^*)n(w^*)L - 1} [p(w^*)]^{R^-(w^*)} [1 - p(w^*)]^{\lambda(w^*)n(w^*)L - 1 - R^-(w^*)} \\ &= \frac{R^-(w^*) + 1}{\lambda(w^*)n(w^*)L p(w^*)} \binom{R^-(w^*) + 1}{\lambda(w^*)n(w^*)L} [p(w^*)]^{R^-(w^*) + 1} [1 - p(w^*)]^{\lambda(w^*)n(w^*)L - (1 + R^-(w^*))} \\ &= \frac{R(w^*)}{a(w^*)} \binom{R(w^*)}{\lambda(w^*)n(w^*)L} [p(w^*)]^{R(w^*)} [1 - p(w^*)]^{\lambda(w^*)n(w^*)L - R(w^*)}. \end{aligned}$$

Here we have used the result, $a(w^*) = p(w^*)\lambda(w^*)n(w^*)L$, and the fact $R(w^*) = R^-(w^*) + 1$. Substituting the above result into the formula of γ^- and using the formula of γ , we have:

$$\gamma^-(R^-, w^*) = \frac{R(w^*)}{a(w^*)} \gamma(R).$$

Since $a(w^*)$ is independent of the realization of R , (G.7) then implies

$$q(w^*) = \sum_{R \in \mathcal{R}} \left[\frac{Z(w^*, R)}{R(w^*)} \cdot \frac{R(w^*)}{a(w^*)} \gamma(R) \right] = \frac{1}{a(w^*)} \sum_{R \in \mathcal{R}} Z(w^*, R) \gamma(R).$$

Therefore, (G.4) holds, as shown below:

$$\sum_{w^*} a(w^*) q(w^*) = \sum_{w^*} \sum_{R \in \mathcal{R}} Z(w^*, R) \gamma(R) = h.$$

The second equality uses (G.5). This completes the proof of the proposition. **QED**

Lemma G.2. (Equal-employment) *It is optimal for a firm to give equal employment probability to all applicants. That is, $q(w, w') = q(w)$ for all such w' that $a(w, w') > 0$.*

Proof. Consider the recruiting firm's decision problem (\mathcal{P}') formulated in the proof of the previous Proposition. For all such w' that $a(w, w') > 0$, the optimal choice of $a(w, w')$ satisfies the following first-order condition:

$$q(w, w') = \left[\frac{J_f(w) - J_v}{\mu} + 1 \right] \exp \left(- \sum_{w'} a(w, w') \right),$$

where μ is the Lagrangian multiplier of (G.6). Clearly, the above equation implies that $q(w, w')$ is independent of w' . **QED**

H. Another Characterization of the Equilibrium

In the text, we used Figure 2*b* to illustrate the equilibrium heuristically. We did not translate Figure 2*b* into a mathematical formulation but, instead, claimed that such a formulation yields very little about the analytical properties of the equilibrium. In this supplementary appendix, we formulate Figure 2*b* mathematically, discuss the analytical difficulties of using this formulation, and then use a numerical example to show that this formulation yields the same equilibrium as the recursive formulation in the text.

Figure 2*b* indicates two steps in the characterization of the equilibrium. First, given a function $q(\cdot)$ that determines the employment probability at each wage level, the applicants choose the target set of wages $T(\cdot)$. Second, anticipating the dependence of the applicants' decisions on $q(\cdot)$, recruiting firms compete by offering $(w, q(w))$ and the competition drives each recruiting firm's net profit to zero, i.e., the expected surplus from recruiting to the level of the vacancy cost.

To formulate the first step, pick any decreasing employment probability function, $q(\cdot)$. For each $w \in \mathfrak{R}$, $q(w)$ gives the probability with which an applicant to a firm offering w will be selected. Note that this employment probability must be specified for all w , not only for equilibrium wages. For convenience, when the target application set is empty, set $q(\emptyset) = 0$. Define:

$$F(w) = w + \lambda(w) \max\{E(w) - S, 0\},$$

where $E(w)$ is a w -applicant's market surplus. We can use (2.10) to write a worker's value function as

$$J_e(w) = \frac{1}{r + \sigma} [F(w) + \sigma J_u].$$

The expected surplus of an applicant for a job at w' is

$$q(w')[J_e(w') - J_e(w)] = \frac{1}{r + \sigma} q(w')[F(w') - F(w)].$$

The applicant applies to w' only if the expected surplus is greater than or equal to S . Therefore, a w -applicant's market surplus is:

$$E(w) = \max \left\{ \frac{1}{r + \sigma} \max_{w'} \{q(w')[F(w') - F(w)]\}, S \right\}.$$

In the inner maximization, the applicant takes the function $q(\cdot)$ as given. If the inner maximization generates a value greater than $(r + \sigma)S$, the applicant's target set $T(w)$ is non-empty. Otherwise, $T(w) = \emptyset$. With this notation, the case $E(w) = S$ means that a w -applicant does not apply.

An applicant's decision depends on the function F and, by definition, F depends on the applicant's market surplus. Thus, the three objects $(F(\cdot), E(\cdot), T(\cdot))$ must be solved simultaneously for the given function $q(\cdot)$. We can combine the tasks by substituting the above formula of $E(w)$ into the definition of F to obtain:

$$F(w) = w + \lambda(w) \max \left\{ \frac{1}{r + \sigma} \max_{w'} \{q(w')[F(w') - F(w)]\} - S, 0 \right\}. \quad (\text{H.1})$$

This is a fixed-point problem for F . Under reasonable conditions we can show that the right-hand side of (H.1) is a contraction mapping, and so there is a unique function $F(\cdot)$ that satisfies the functional equation. However, it is difficult to find a meaningful condition to ensure that the set of maximizers $T(w)$ is singleton. Even if we assume that $T(w)$ is singleton, it is difficult to complete the second step of the characterization of the equilibrium.

The second step determines the employment probability function through the free-entry condition, (3.2). Rewrite (3.2) as $q(w) = Pq(w)$, where P is the following mapping:

$$Pq(w) = \Psi \left(\frac{r + \sigma + \lambda(w)q(T(w))}{y - w} C \right). \quad (\text{H.2})$$

Then, $q(\cdot)$ is a fixed point of the mapping P .

It is difficult to examine the fixed-point problem for q , because the maximizer $T(w)$ to the fixed-point problem (H.1) appears in the mapping P . To see the difficulty, note that we need the given function $q(\cdot)$ in (H.1) to be decreasing in order to ensure well-behaved fixed point F . In turn, this requires that the fixed point of P be decreasing, and hence that P maps decreasing functions into decreasing functions. However, we cannot find meaningful conditions to guarantee that P has this property. All such conditions involve T , which in turn involves the very object $q(\cdot)$ that we need to determine in equilibrium.

None of these analytical difficulties would be present if there were no on-the-job search, i.e., if $\lambda = 0 < \lambda_0$. If $\lambda = 0$, then $q(w) = \Psi \left(\frac{r + \sigma}{y - w} C \right)$, which is indeed a decreasing (and concave)

function of w . For all $w \neq w_0$, $F(w) = w$. For $w = w_0$, $F(w_0)$ is given by (H.1), with $F(w')$ on the right-hand side of that equation being replaced by w' and $q(\cdot)$ by the function just obtained.

The above formulation shows why it is difficult to analytically examine on-the-job search in a directed search environment. The recursive procedure presented in the text is useful for tackling this problem and to obtain analytical properties of the equilibrium.

Nevertheless, the above formulation suggests the following procedure to compute an equilibrium numerically. Start with a decreasing function $q(\cdot)$ and find the fixed point F in (H.1). Substitute the maximizer T into (H.2) to compute $Pq(w)$. Then, use this solution $Pq(w)$ to serve the role of $q(w)$ in (H.1). Repeat the process until $Pq(\cdot) = q(\cdot)$.

To see whether the above formulation and the recursive formulation in the text generate the same equilibrium, consider the following parameter values: $r = 0.02$, $y = 1000$, $b = 0$, $C = 60$, $S = 1$, $\lambda = \lambda_0 = .025$, and $\sigma = .125$ (this is the example of section 5.2, except for $\lambda = \lambda_0$). Discretize the interval between b and y , and set the number of points on the grid to be 25000. The equilibrium number of rungs on the wage ladder is $M = 4$ in both formulations. Refer to the above formulation as “fixed-point” and the recursive formulation “recursive”, we list the results as follows:

	w_1	w_2	w_3	w_4
recursive	929.005	980.004	988.339	990.679
fixed-point	929.043	980.401	988.640	990.760
discrepancy	0.004%	0.040%	0.031%	0.008%

The results under the two formulations are very close to each other. The maximum discrepancy in equilibrium wages between the two formulations is about 0.04 percent. The discrepancy arises for two reasons. First, the two formulations discretize different objects in the numerical procedure. With the recursive formulation, we discretize the interval of the hiring probability h_M and, with the fixed-point formulation, we discretize the interval of the wage level. Second, the function $q(w)$ is highly nonlinear. It remains flat at low wages but sharply declines at high wages, with a slope approaching $-\infty$ as w approaches the upper bound. Such nonlinearity reduces the accuracy of the numerical results. Taking such sources of discrepancy into account, we conclude that the two formulations yield the same equilibrium.