INTERNATIONAL ECONOMIC REVIEW
Vol. 47, No. 2, May 2006

DIRECTED SEARCH ON THE JOB AND THE WAGE LADDER*

BY ALAIN DELACROIX AND SHOUYONG SHI[1]

Département des sciences économiques, Université du Québec à Montréal, Canada; Department of Economics, University of Toronto, Canada

We model a labor market where employed workers search on the job and firms direct workers' search using wage offers and employment probabilities. Applicants observe all offers and face a trade-off between wage and employment probability. There is wage dispersion among workers, even though all workers and jobs are homogeneous. Equilibrium wages form a ladder, as workers optimally choose to climb the ladder one rung at a time. This is because low-wage applicants are relatively more sensitive to employment probability than to wage and thus forgo the opportunity to apply for a high wage, with a lower chance of success.

1. INTRODUCTION

We study a large labor market with on-the-job search and directed search. Employed workers search on the job. Search is directed in the sense that each firm takes into account how its offer will affect its own matching rate and workers' application, as opposed to undirected search where an exogenous-matching function determines agents' match rates. All workers and all jobs are identical. The number of workers is large and fixed, whereas the number of jobs is determined by free entry. The recruiting process in each period has two stages. In the first stage, each firm announces a wage level and an employment probability for the applicants. In the second stage, the applicants observe all offers and decide which job to apply to. After receiving the applicants, a firm selects one and pays the announced wage. Recruiting generates endogenous transition of workers between jobs, whereas exogenous separation sends workers into unemployment. We characterize the stationary equilibrium and study its properties.

The search literature (see Table 1 for a rough guide) offers very little knowledge about the equilibrium with directed search on the job. The original search models

* Manuscript received November 2003; revised July 2004.

[1] We would like to thank Dale Mortensen, Rob Shimer, and Audra Bowlus for comments. We also received useful comments from the participants of Minnesota Summer Workshop on Macroeconomic Theory (Minneapolis, 2002), National Bureau of Economic Research Summer Workshop (Cambridge, 2002), Canadian Theory Meeting (Toronto, 2002), Midwest Theory Meeting (Notre Dame, 2002), Midwest Macro Meeting (Chicago, 2003), the workshops at Northwestern University (2002), Purdue University (2002), and Queen's University (2003). Shi would like to acknowledge the Bank of Canada Fellowship and the Social Sciences and Humanities Research Council of Canada for financial support. The view expressed in this article is the authors' own and does not reflect the opinion of the Bank of Canada. All errors are ours alone. Please address correspondence to: S. Shi, Department of Economics, University of Toronto, 150 St. George Street, Toronto, Ontario, Canada, M5S 3G7. Phone: 416-978-4978. Fax: 416-978-6713. E-mail: *shouyong@economics.utoronto.ca.*

TABLE 1
SEARCH MODELS

	Whether Employed Workers Are Allowed to Search on the Job	
Type of Search	No	Yes
Undirected	Diamond (1982) Mortensen (1982) Pissarides (1990)	Burdett and Judd (1983) Burdett and Mortensen (1998) Pissarides (1994)
Directed	Peters (1991) Montgomery (1991)	This article

excluded both on-the-job search and directed search. The subsequent research has relaxed these two assumptions separately but not simultaneously. For example, Burdett and Mortensen (1998) and Pissarides (1994) have examined undirected search on the job, whereas Peters (1991) and Montgomery (1991) have examined directed search without on-the-job search.[2] The combined feature of directed search and on-the-job search will be the focus in this article.

Our study is also motivated by the following evidence on job/worker flows and wage inequality: (1) A large fraction of overall wage inequality is within-group inequality, i.e., among workers who have similar observed characteristics such as education, experience, and age (e.g., Juhn et al., 1993); (2) a large fraction of job changes are direct job-to-job transitions, and wage mobility is predominantly within-group mobility;[3] (3) wage mobility is limited in the sense that the probabilities of staying in the same quintile of wages and moving to adjacent quintiles account for most of wage changes annually (Buchinsky and Hunt, 1999); and (4) the density of the wage distribution is hump-shaped, with the hump occurring at a low-wage level (Kiefer and Neumann, 1993). A model generating wage dispersion among homogeneous workers is able to replicate fact (1). To account for fact (2), a model must incorporate on-the-job search. However, existing models of on-the-job search have difficulties accounting for facts (3) and (4).

To substantiate this claim, consider the influential model of on-the-job search by Burdett and Mortensen (1998; henceforth the BM model). The BM model successfully generates wage dispersion among homogeneous workers with homogeneous

[2] The model by Burdett and Judd (1983) is for the goods market, but the essence is carried over to the on-the-job search model by Burdett and Mortensen (1998). Similarly, the directed search model by Peters (1991) is for the goods market. Other examples of on-the-job search models are Coles (2001), Burdett and Coles (2003), and Postel-Vinay and Robin (2002). Other examples of directed search models are Acemoglu and Shimer (1999a,b), Cao and Shi (2000), Julien et al. (2000), Moen (1997), Peters (2000), Burdett et al. (2001), Shi (2001, 2002a,b), and Shimer (2001).

[3] Topel and Ward (1992) study U.S. workers' job changes using the (U.S.) Longitudinal Employee-Employer Data. They report that an average worker has changed jobs almost seven times by the 10th year after entering the labor market and, for workers with one or more years of experience, direct job-to-job transition dominates the transition from employment to nonemployment. Using the (U.S.) National Longitudinal Survey of Youth, Buchinsky and Hunt (1999) report that the index of within-group wage mobility is one order of magnitude larger than the index of between-group mobility.

jobs. The important insight is that on-the-job search creates heterogeneity among workers' reservation wages (i.e., their current wages), which supports a continuum of wages as an equilibrium. However, because search is undirected, each worker contacts all job openings randomly and with equal probability. To ensure that firms are indifferent between offering different equilibrium wages, there must be more firms offering high wages than low wages. Thus, the density function of offer wages is increasing (and convex). This induces an increasing (and convex) density function of employed wages, which does not accord well with fact (4). Moreover, because matching is exogenous, a worker's transition probabilities to higher wages (conditional on his current wage) are proportional to the distribution of offer wages.[4] Since more firms recruit at high wages than at low wages, then a worker is more likely to transit to a very high wage than to a wage just above his current wage. This pattern of wage mobility is opposite to that in fact (3).

Of course, one can introduce either observed or unobserved heterogeneity among workers or jobs into the existing models to make their predictions more realistic. For example, to generate the hump-shaped wage distribution, one may argue that workers differ in ability or that jobs differ in quality (e.g., van den Berg and Ridder, 1998). To generate limited wage mobility, 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 within-group wage dispersion is sufficiently large (fact (1)), it is useful to construct a theoretical model to generate wage dispersion among homogeneous workers.

Our model generates nondegenerate dispersion of ex ante optimal wage offers by homogeneous firms to homogeneous workers. On-the-job search is important for such dispersion, as in the BM model. If on-the-job search is not permitted, a single equilibrium wage will emerge as in previous models of directed search.[5] However, the equilibrium in our model is fundamentally different from that in the BM model.

In fact, directed search destroys the BM type of equilibrium, by requiring workers' application to be optimal. In the BM model, job application is exogenous, as a matching function assigns workers to jobs. Even though applying to a high wage yields higher expected payoff than applying to a low wage, workers are not allowed to choose which offers to apply to. Once this assumption is eliminated, the workers will choose to apply to only those offers that maximize their expected surplus, that is, the product of the ex post gain in value and the probability of obtaining

[4] More precisely, if $v(\cdot)$ is the density of offer wages and $V(\cdot)$ the corresponding cumulative distribution, then the density of transition for a worker employed at wage w to a higher wage w' is $v(w')/[1 - V(w)]$.

[5] Some directed search models, such as Julien et al. (2000), generate ex post wage dispersion by allowing firms or workers to auction the jobs. However, all firms use the same mechanism that generates the same expected wage. Other models of directed search generate wage dispersion by introducing heterogeneity. For example, in Shi (2002a), firms are different in size (i.e., the number of employees); in Shi (2002b) and Shimer (2001), the presence of high-skill workers induces partial sorting and generates a wage differential among low-skill workers.

the job upon application, denoted as *employment probability*. This choice makes the wage offers in the BM model unsustainable as an equilibrium.

The equilibrium with directed search on the job is a wage ladder, which comprises of a finite number of wage levels (or rungs). Firms trade off a higher wage for a greater probability of hiring a worker. In equilibrium, firms are indifferent between posting any wage on the support of the wage distribution, since they all yield the same expected surplus (or profit). Workers face a trade-off between a higher wage and a lower employment probability. However, applicants are *not* indifferent between applying to the different wages in equilibrium. Rather, an applicant chooses to apply only to such jobs that lie one rung above his current wage on the ladder, since these jobs provide him with the highest expected surplus. Thus, workers choose optimally to climb the wage ladder one rung at a time.

The wage ladder arises as an equilibrium because applicants' current wages affect their trade-off between employment probability and wage. The higher an applicant's current wage is, the lower the ex post surplus he can obtain from a given wage, and so a given amount of the wage gain represents a larger proportional increase in the expected surplus to such an applicant. Put differently, an applicant with a high current wage cares more about the wage gain, and less about the employment probability, than does an applicant with a low current wage. This single-crossing property allows firms to separate the applicants by offering a high wage with a low employment probability to those who have a high current wage, and a low wage with a high employment probability to those who have a low current wage. The separation produces the wage ladder in a stationary equilibrium. Obviously, this wage ladder does not rely on any exogenous heterogeneity between workers or jobs.

The wage ladder has the following strong implications: (i) Wage mobility is limited endogenously, as a worker's next move has only three possible outcomes—to stay at the current wage, to move up the ladder by one rung, or to transit into unemployment; (ii) the density of offer wages is strictly decreasing, which induces the density of employed wages to be either decreasing or nonmonotonic (numerical examples show that it is sharply decreasing); and (iii) the gap between two adjacent rungs on the ladder becomes smaller and smaller as a worker climbs up the ladder, and so the wage gain diminishes. These properties contrast sharply with those in the BM model. Strongly restricted wage mobility accords well with fact (3) listed in this introduction. The decreasing density of the wage distribution is consistent with only one part of the empirical distribution (see fact (4)), but it is the part that has been elusive in the BM model.

A decreasing density of offer wages is a necessary outcome of the wage ladder. Because the wage ladder implies that workers employed at low wages do not apply to very high wages, the source of applicants for high-wage recruiting firms is limited endogenously. However, in order to make firms indifferent between recruiting at different equilibrium wages, each high-wage firm must be more successful in recruiting than each low-wage firm. This is possible, given the limited source of applicants to a high wage, only if there are fewer firms competing against each other at a high wage than at a low wage, that is, only if the density of offer wages is decreasing. When the density of offer wages is sufficiently decreasing, the density of employed wages is also decreasing.

While the wage ladder generates the above novel properties, it retains the following realistic properties of the BM model. First, a worker's quit rate decreases with wage. Second, a worker's wage increases, on average, with his employment duration. Third, the average length of time an unemployed worker will take to return to his previous wage increases with that wage.[6]

In Section 2 we will describe the economy with directed search on the job and define an equilibrium. In Section 3 we will analyze agents' trade-off between the matching probability and wage, and argue that the equilibrium must be a wage ladder. Section 4 will construct a candidate equilibrium by restricting the deviations. Section 5 will remove such restrictions and find conditions under which the candidate is indeed an equilibrium. We will present the analytical properties of the equilibrium in Section 6 and provide numerical examples. Section 7 will contrast our model with BM. Section 8 will conclude the article and the appendix will collect the proofs.

2. A MODEL OF DIRECTED SEARCH ON THE JOB

2.1. *The Labor Market and Job Search.* Time is discrete. A labor market is populated by a large (exogenous) number, L, of risk-neutral and infinitely lived workers.[7] All workers are identical. When employed, a worker supplies one unit of labor and produces $y > 0$ units of output per period. When unemployed, a worker receives a 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 period is $C > 0$. Time is discrete. Workers and firms discount future with the discount factor $1/(1 + r)$, where $r > 0$.

The events in each period unfold as follows: At the beginning of the period, employed workers produce and obtain wages. Then, each employed worker receives a shock that has three realizations. The first realization is exogenous separation from the current job, which occurs with probability $\sigma > 0$ and sends the worker into unemployment. The second realization is a job application opportunity, which occurs with probability λ_1 and enables the worker to apply to other jobs. The third realization is an inactive state, in which the worker stays put in the current period. Also, each unemployed worker receives a shock that gives the worker a job

[6] Empirically, there exists a positive relationship between wage and firm size. Most directed search models do not address the issue of firm size explicitly, since applicants apply to a job and not to a firm (exceptions are Burdett et al., 2001; Shi, 2002a). However, our model can be reinterpreted as implying the same relationship, given that expected queue sizes increase with wage.

[7] The wage ladder is robust to the introduction of risk aversion. A worker's application choice is based on a trade-off between wage and the probability of obtaining the job. Even though risk-averse workers would give relatively higher value to employment probability at all wage levels, the trade-off would still depend on the current wage earned and thus a wage ladder would ensue.

application opportunity with probability λ_0 and keeps the worker inactive in the current period with probability $1 - \lambda_0$.

Next, firms and workers play a two-stage recruiting game. In the first stage, each potential firm chooses whether to incur the vacancy cost to become a recruiting firm in the period. Each entering firm announces a job description, which consists of a wage offer and a selection rule (described later). All recruiting firms announce the job descriptions simultaneously. In the second stage, the workers who have received application opportunities in the period observe all firms' announcements and choose which job opening to apply to.[8] To apply, an applicant must incur a small cost $S > 0$.[9] An applicant can apply to only one job, but the application can be mixed strategies over job openings. After receiving applicants, a firm selects one according to the announced criterion. Then, the period ends. If a vacancy is not filled in the period, the firm must incur the vacancy cost again next period in order to recruit.

To emphasize the new results that will arise from the combination of directed search and on-the-job search, we keep other aspects of our model as closely as possible to the BM model. Thus, we maintain the following three auxiliary assumptions of the BM model. First, firms commit to what they post (for a relaxation of this assumption, see Coles, 2001). Second, firms post wage levels rather than contracts or mechanisms (for a relaxation of this assumption, see Julien et al., 2000; Burdett and Coles, 2003).[10] Third, a worker's current employer does not match outside offers, and so the worker always quits upon receiving a higher wage (for a relaxation of this assumption, see Postel-Vinay and Robin, 2002). These auxiliary assumptions also simplify the analysis considerably.

This labor market exhibits the following frictions that are familiar in previous search models. First, agents cannot coordinate their decisions, which creates the possibility of unmatched agents. Second, job application opportunities are not abundant, in the sense that $\lambda_1 < 1$ and $\lambda_0 < 1$. This is a proxy for the cost of gathering information about jobs. Third, each applicant can apply to only a small number of jobs at a time, which is set to be one in our article. This is a proxy for the constraint that an applicant can attend only one interview at a time.

The unique feature of our model is the combination of on-the-job search and directed search. Employed workers search on the job after receiving application opportunities. Search is directed because the applicants observe firms' offers before the application. By choosing the job description, a firm can intentionally affect how quickly it will get a match and what type of applicants it will attract. By contract, previous models of on-the-job search, like the BM model, have assumed undirected search in the sense that matching rates are exogenous to the agents. On the other hand, previous models of directed search, such as Acemoglu and Shimer (1999a) and Burdett et al. (2001), have excluded employed workers from

[8] For directed search to occur, it is sufficient to assume that each applicant observes two offers that are randomly drawn from all recruiting firms' offers. However, the analysis is much more complicated.

[9] The small cost S is needed to help the existence of an equilibrium in our model. We will discuss the role of S in our model in Section 4.2 and how it affects the BM equilibrium in Section 7.

[10] If each firm auctions the job by posting a reserve wage, as it does in Julien et al. (2000), then the reserve wage serves the role similar to the posted wage in our model. Then an argument similar to ours in Section 3.2 shows that the equilibrium must be a wage ladder.

search. By allowing for $\lambda_1 \neq \lambda_0$, we nest those previous directed search models as a special case where $\lambda_1 = 0 < \lambda_0$.

Let wage levels lie in the interval, $\mathcal{W} = [\underline{w}, y]$, where $\underline{w} \in (-\infty, y]$.[11] Let $N(\cdot)$ be the cumulative distribution function of workers over wages at the beginning of each period, with $n(\cdot)$ being the density.[12] Since we classify the unemployment benefit w_0 ($=b$) as a "wage" level, then $N(w_0) = u$. The distributional density of (employed) workers over wages in $\mathcal{W}\backslash\{b\}$ 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 $L \to \infty$. However, to detail agents' strategies, we will first analyze a market where L is large and finite, and then take the limit $L \to \infty$. The expected number of applicants currently employed at w is $\lambda_1 n(w)L$ (or $\lambda_0 uL$ if $w = w_0$). Assume that this number is an integer without loss of generality.

2.2. *Strategies and Payoffs.* Before describing the strategies, let us call a w-worker who has just received a job application opportunity a w-applicant. To unify the notation for employed and unemployed workers, let $\lambda(w) = \lambda_1$ for all $w \neq w_0$ and $\lambda(w_0) = \lambda_0$. 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 $J_f(w)$ be the value function of a firm that currently employs a worker at w and $J_v(w)$ be the value function of a vacancy recruiting at wage w. Let $\bar{J}_v = \max_w J_v(w)$. All these value functions are measured at the end of each period.

Let us describe the applicants' strategies first. A w'-applicant's strategy is a function $p(\cdot, w') : \mathcal{W} \to [0, 1]$, where $p(w, w')$ is the probability with which the applicant applies to *each* of the job openings at wage w.[13] An implicit assumption is that the applicant must assign equal probability to applying to all identical offers. Also, all applicants of the same type use the same strategy. This requirement of symmetry reflects the fact that it is difficult for agents to coordinate their actions in a large market.

We normalize p as follows:

$$a(w, w') = p(w, w')\lambda(w')n(w')L \tag{1}$$

Since $(\lambda(w'), n(w'), L)$ are all exogenous to an individual agent, the above normalization does not change the nature of p. Thus, a w'-applicant's strategy can be described alternatively by a function $a(\cdot, w')$: $\mathcal{W} \to \Re_+$. It is more convenient to

[11] The lower bound $\underline{w}$ can be set as $\underline{w} = b - \lambda_1[(y-b)/r - S]$. An unemployed worker will never accept wages below this $\underline{w}$ because the present value of accepting $\underline{w}$ and then facing the prospect of getting a job that pays y forever is equal to the present value of staying unemployed forever, b/r.

[12] Clearly, we do not presume that $N(\cdot)$ is a continuous distribution function. Throughout this article, we expand the meaning of a "density function" to include the frequency function of a discrete distribution, as well as the density of a continuous distribution.

[13] There is no need to specify the dependence of p on the distribution of offer wages, because an applicant can always set $p(w, w') = 0$ for such wages w that are not offered.

use a rather than p because $p(w, w') \to 0$ for all w in a symmetric equilibrium as the market becomes large, but $a(w, w')/n(w')$ remains strictly positive provided that $p(w, w') > 0$.

The variable $a(w, w')$ has another meaning in a symmetric equilibrium. When all type w applicants use the same probability $p(w', w)$ to apply to each opening at wage w', the expected number of such applicants whom the opening receives is equal to $a(w', w)$. For this reason, we also call $a(w', w)$ the *queue length* of w-applicants for a job opening at w'. Despite this coincidence, the use of $a(w', w)$ as a w-applicant's strategy does not mean that an individual w-applicant can choose the queue length or can influence other w-applicants' strategies.

Denote $A(w') = \{a(w, w')\}_{w\in\mathcal{W}}$. A w'-applicant's target set of wages is $T(w') = \{w : a(w, w') > 0\}$. Evidently, for each w'-applicant, the probabilities $p(w, w')$ must sum to one over w and over vacancies that offer these wages. This constraint can be written as follows:

$$\sum_w \left[\frac{a(w, w')}{n(w')} v(w)\right] = \frac{\lambda(w')}{k}$$

To specify an applicant's strategy and payoff, let $q(w)$ be the ex ante employment probability offered to an applicant by a firm recruiting at wage w—"ex ante" in the sense that the probability is computed before knowing the actual number and composition of the applicants whom the firm will receive (see more discussion later). Then, a w'-applicant's payoff of applying for a firm posting w is the expected surplus, $q(w)[J_e(w) - J_e(w')]$. Let $E(w')$ be the maximum expected surplus that a w'-applicant can get by applying to firms other than the one in the discussion. (The dependence of $E(w')$ on the distribution of wage offers is suppressed.) Clearly, a w'-applicant applies to the particular firm with probability one if the expected surplus of such application exceeds $E(w')$, with any probability in [0, 1] if the expected surplus is equal to $E(w')$, and with probability zero otherwise. Expressing this result in terms of $a(w, w')$ and taking the limit $L \to \infty$, we can characterize a w'-applicant's optimal strategy as follows:

$$a(w, w') \begin{cases} = \infty, & \text{if } q(w)\big[J_e(w) - J_e(w')\big] > E(w') \\ \in [0, \infty), & \text{if } q(w)\big[J_e(w) - J_e(w')\big] = E(w') \\ = 0, & \text{otherwise} \end{cases} \tag{2}$$

Note that the first case and the third case in the above equation will not arise when $(w, q(w))$ is an equilibrium offer. If $a(w, w') = \infty$, 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 $E(w')$, which contradicts 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. Thus, for every equilibrium offer, the expected surplus of applying to a firm making such an offer must be equal to $E(w')$. For this reason, we call $E(w')$ the *market surplus* for a w'-applicant.

When the size of the market approaches infinity, the effect of any individual agent (worker or firm) on the market surplus vanishes.[14] That is, each individual agent in a large market takes $E(w')$ as given for all w'.

Now, let us turn to a recruiting firm's strategies. A recruiting firm announces a job description, which consists of a wage level $w \in \mathcal{W}$ and a criterion that determines which applicant will get the job. This ex post selection criterion depends on the profile (i.e., the composition) of the applicants whom the firm will receive. The profile is stochastic in general, because the applicants can use mixed strategies to apply to identical jobs. For each realization of the profile, the selection criterion must specify the ex post employment probability with which each applicant is selected. This is cumbersome but, fortunately, it is not necessary. Since all agents are risk neutral, posting a job description is equivalent to posting a wage w and an "ex ante" employment probability for each applicant, q (see Delacroix and Shi, 2002, supplementary Appendix G, for a proof). This ex ante probability is calculated by aggregating ex post employment probabilities in the ex post selection criterion over all realizations of the profile of applicants whom the firm will receive. We refer to q simply as the *employment probability*. Because firms recruiting at different wages may offer different employment probabilities, we write q as $q(w)$ to index the employment probability by the wage which the firm offers.[15]

Therefore, a recruiting firm's strategy is to announce a wage level $w \in \mathcal{W}$ and an employment probability $q(w) \in [0, 1]$ for every applicant. Let us now refer to $(w, q(w))_{w\in\mathcal{W}}$ as the profile of job descriptions.

The payoff to a firm posting $(w, q(w))$ is the expected surplus, $h(w)[J_f(w) - \bar{J}_v]$, where $h(w)$ denotes the firm's probability of successfully hiring a worker. The only event in which the firm fails to hire is when the firm does not receive any applicant at all. In the limit $L \to \infty$, this probability is $\exp(-\sum_{w'} a(w, w'))$, where $\sum_{w'} a(w, w')$ is the total expected number of applicants for the firm.[16] Thus, the firm's hiring probability in a large market is

$$h(w) = 1 - \exp\left[-\sum_{w'} a(w, w')\right] \tag{3}$$

[14] This result is proven by Peters (2000), Cao and Shi (2000), and Burdett et al. (2001) in various settings of directed search. A similar proof applies here.

[15] In general, the ex ante employment probability offered by a firm posting wage w could differ across the applicants' types; that is, it could have the form $q(w, w')$, where w' is the applicant's type. However, it is optimal for a firm to set $q(w, w') = q(w)$ for all w' (see Delacroix and Shi, 2002, Appendix G). 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 (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. This implies that it is not even necessary to assume that the applicant's current wage is observed by firms.

[16] The derivation for this formula starts with a finite market and then takes the limit $L \to \infty$, using the fact that $(1 - p)^{1/p} \to e^{-1}$ when $p \to 0$. We omit this derivation, because it is familiar in directed search models (e.g., Peters, 1991; Burdett et al., 2001).

Since each firm hires at most one worker, $h(w)$ is also the expected number of workers hired by the firm. As a result, the employment probability for each applicant who applies to the firm is

$$q(w) = \left[1 - \exp\left(-\sum_{w'} a(w, w')\right)\right] \Big/ \sum_{w'} a(w, w') \tag{4}$$

Thus, the more applicants the firm attracts, the more likely the firm will succeed in filling the job, and the less likely each applicant will get the job. For this reason, q and h have the following negative relationship:

$$q(w) = -\frac{h(w)}{\ln[1 - h(w)]} \equiv \Psi(h(w)) \tag{5}$$

Let Ψ^{-1} be the inverse of Ψ, so that $h(w) = \Psi^{-1}(q(w))$. Clearly, $\Psi^{-1}(1) = 0$. That is, the only case in which a firm can guarantee employment for a potential applicant is when the firm does not have any applicant at all. Similarly, $\Psi^{-1}(0) = 1$.

A recruiting firm's optimal offer $(w, q(w))$ solves the following problem:

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

subject to

$$q(w)[J_e(w) - J_e(w')] \geq E(w'), \text{ for all } w' \text{ such that } T(w') \ni w \tag{6}$$

The firm takes as given other firms' decisions and the applicants' market surpluses. Since $\Psi^{-1}(1) = 0$ and $\Psi^{-1}(0) = 1$, offering $q = 1$ does not maximize the expected profit and offering $q = 0$ violates the constraint.

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 reduces the ex post value of the job to the firm (J_f), but it attracts more applicants and hence increases the success of hiring. Similarly, for an applicant, a higher wage offer yields a higher value of employment J_e, but it is more difficult to be obtained. The applicant applies only to those firms whose offers maximize the applicant's expected surplus, as the dual to $(\mathcal{P})$ suggests.

The trade-off between the matching probability and wage is "smooth," as in all directed search models. A marginally higher wage offer will not induce the applicants to increase the probability of application to that offer by a discrete amount. This is because all the applicants will observe the increase in the wage offer before applying to the job. If they all responded to the slightly higher offer by a discrete increase in the application probability, then the employment probability for each of them would be close to zero, which would not be optimal for the applicants. Such smoothness will enable us to use the first-order conditions to analyze the equilibrium.[17]

Now, we specify the value functions. Recall that the value functions are measured at the end of each period. They satisfy the following Bellman equations:

$$J_v(w) = \frac{1}{1+r}\{-C + h(w)J_f(w) + [1-h(w)]\bar{J}_v\} \tag{7}$$

$$rJ_f(w) = (y-w) - [\sigma + \rho(w)][J_f(w) - \bar{J}_v] \tag{8}$$

$$rJ_e(w) = \left[\begin{array}{l} w - \sigma[J_e(w) - J_u] - \chi\left(\sum_{w'} a(w', w)\right)\lambda_1 S \\ + \dfrac{k}{n(w)} \sum_{w'} q(w')[J_e(w') - J_e(w)]a(w', w)v(w') \end{array}\right] \tag{9}$$

$$rJ_u = b - \lambda_0 S + \frac{k}{u}\sum_{w'} q(w')[J_e(w') - J_u]a(w', b)v(w') \tag{10}$$

The function $\chi(\cdot)$ in (9) is an indicator function, with $\chi(\Sigma a) = 1$ if $\Sigma a > 0$ and $\chi(\Sigma a) = 0$ if $\Sigma a = 0$. The function $\rho(w)$ is the probability with which a w-worker quits his current job for another job, which is given as follows:

$$\rho(w) = \lambda(w)\sum_{w'} q(w')p(w', w)v(w')K = k\sum_{w'} q(w')v(w')\frac{a(w', w)}{n(w)} \tag{11}$$

To explain the above Bellman equations, consider (9) for example. This equation equates the permanent income of a worker employed at wage w, $rJ_e(w)$, to the expected "cash flow" to such employment in the next period. The cash flow consists of the 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 (9). If this difference is nonpositive, the worker will choose $\Sigma a = 0$, in which case the last two terms in the equation are zero. By construction, experiencing exogenous separation and receiving a job application opportunity are two distinct realizations of the same shock to an employed worker.

The value function $J_e(w)$ must be strictly increasing for all $w \neq w_0$ and $w < y$. To see this, consider two workers who are employed at w' and w'', respectively, with $w'' < w' < y$. The worker employed at w' receives a job application opportunity as often as does the worker employed at the lower wage w'', and he can apply to all job openings that the worker at w'' can. Thus, the expected payoff to the w'-worker in the future application is at least as high as that to the w''-worker. In addition, a

[17] The BM model does not have such smoothness. There, if a mass of workers are employed at a wage level, then a firm can increase expected profit by a discrete amount by increasing the offer slightly above that wage. This is because all applicants at that wage will be exogenously matched to the deviator with a positive probability.

w'-worker gets a higher wage from his current job than a w''-worker does. Thus, $J_e(w') > J_e(w'')$. The only exception arises when we use $J_e(w_0)$ to stand for J_u. If $\lambda_0 < \lambda_1$, employment gives a worker a better access to higher wages in the future than unemployment. In this case, it is possible that $J_e(w') > J_e(w_0)$ for some $w' < w_0$.

As explained before, directed search implies that firms and workers have a smooth trade-off between the matching probability and wage. Thus, we maintain that the value functions are continuous. Again, the only exception is $J_e(w_0)$ when $\lambda_0 \neq \lambda_1$.

2.3. *Definition of Equilibrium.* The set of equilibrium wage offers is $\Omega \equiv \{w \in \mathcal{W} : v(w) > 0\}$. 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 cumulative 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 1. A symmetric Nash equilibrium in the labor market consists of the aggregate characteristics $(\Omega, N(\cdot), V(\cdot), k)$, recruiting firms' offers $(w, q(w))_{w\in\Omega}$, and the applicants' strategies $(A(w))_{w\in\Omega_0}$, where $A(w) = (a(w', w))_{w'\in\Omega}$, such that the following requirements are met: (i) Given the aggregate characteristics and given that other (current and future) recruiting firms post $(w, q(w))_{w\in\Omega}$, an individual firm's optimal strategy is to post $(w, q(w))$ with $w \in \Omega$; (ii) given the vacancy distribution over $(w, q(w))_{w\in\Omega}$ and the aggregate characteristics, each w-applicant's optimal strategy is $A(w)$; (iii) the strategies are symmetric within each type; (iv) there is free entry of firms: $J_v(w) = \bar{J}_v = 0$ for every recruiting firm; and (v) the aggregate characteristics are stationary.

Note that part (i) of the definition requires each firm to take as given that other firms post $(w, q(w))_{w\in\Omega}$ in the future, as well as in the current period. That is, when a firm deviates to a wage $w' \notin \Omega$, it views that other firms in the future will not post wages outside Ω to attract its prospect worker. Thus, the equilibrium defined above may not be Markovian. In Section G of the Appendix, we examine the Markov perfect equilibrium and discuss the difficulties of analytically characterizing it. We also provide a numerical example to show that the Nash equilibrium in our model can be very close to the Markov equilibrium.

As required by the equilibrium, we will set $J_v(w) = \bar{J}_v = 0$ in the remainder of this article.

3. CONFIGURATION OF THE EQUILIBRIUM

In this section, we will first show that an equilibrium must feature the separation of the applicants by their current wages. Then we will argue that an equilibrium must be a ladder.

3.1. *Separation of Applicants by Their Current Wages.* To search for the equilibrium configuration, we start with an applicant's trade-off between the employment probability and wage. This trade-off is summarized by the applicant's indifference curve. For an applicant who is currently employed at wage w', the indifference curve is the equality form of (6), which can be rewritten as follows:

$$q(w) = \frac{E(w')}{J_e(w) - J_e(w')} \text{ all } w' \text{ such that } a(w, w') > 0 \tag{12}$$

Because the value of employment is a strictly increasing function of wage, as we explained before, this is a negative relationship between the employment probability and wage. Thus, a low wage must be compensated by a high employment probability in order to induce an applicant to apply.

An applicant's current wage affects the trade-off between the employment probability and wage. In fact, the indifference curves of two different types of applicants have the single-crossing property. To see this, we calculate the slope of a w'-applicant's indifference curve at (w^*, q^*) as follows:

$$-\frac{dq}{dw} = \frac{q^* J_e'(w^*)}{J_e(w^*) - J_e(w')}$$

As we argued before, $J_e(\cdot)$ is an increasing function, and so $J'_e(w^*) > 0$. Thus, if an offer (w^*, q^*) lies on both indifference curves of type w_i and type w_j applicants, with $w_j > w_i$, then

$$-\frac{dq}{dw}\bigg|_{w'=w_j} > -\frac{dq}{dw}\bigg|_{w'=w_i} \quad \text{at } (w, q) = (w^*, q^*)$$

The single-crossing property reflects the difference between the two applicants' willingness to sacrifice the employment probability for a wage gain. The higher an applicant's current wage, the lower the ex post surplus he can obtain from a given wage. So, a given amount of the wage gain 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 applicant does.

The single-crossing property implies that the equilibrium will feature separation of applicants by their types, as stated in the following lemma (see Section A of the Appendix for a proof).

LEMMA 1. *If there is an equilibrium, then each equilibrium wage attracts at most one type of applicant. Precisely,* $a(w^*, w_i)a(w^*, w_j) = 0$ *for all* $w_i, w_j, w^* \in \Omega_0$ *with* $w_j > w_i$.

Figure 1 illustrates Lemma 1. Two indifference curves are drawn, one for the applicant of type $w' = w_i$ and the other for $w' = w_j > w_i$. The single-crossing property implies that the indifference curve of the high-wage applicant (w_j-applicant) crosses that of the low-wage applicant from above (at point A). We also draw the

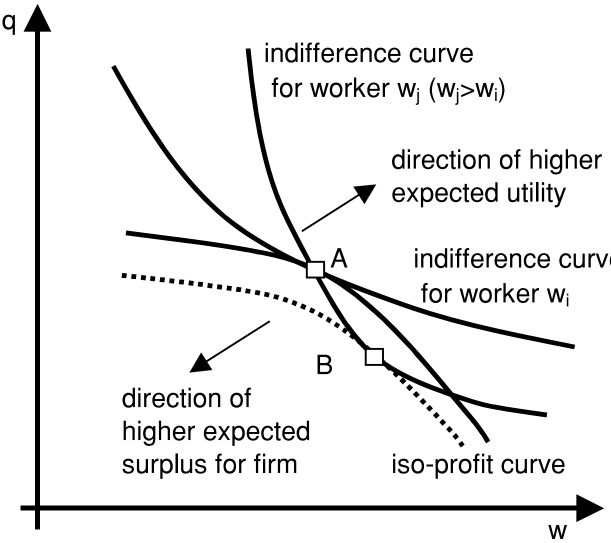

FIGURE 1

INDIFFERENCE CURVES IN $(w,\ f)$-SPACE

iso-profit curve of the recruiting firm, which summarizes different combinations of (w, q) that yield the same expected surplus to the recruiting firm. If the firm attracts both types of applicants, the firm's offer must be at the intersection of these two indifference curves. However, this offer does not maximize the firm's expected surplus. (In the situation depicted in Figure 1, the offer B yields higher expected surplus to the firm.)

Lemma 1 implies that the BM equilibrium cannot be an equilibrium when on-the-job search is directed, as we will show in the next subsection.

3.2. *The Equilibrium Must Be a Ladder.* To see what structure the equilibrium has, we assume that the recruiting firm's decision problem $(\mathcal{P})$ has a unique solution that is continuous in the applicant's current wage.[18] Then, the dual of $(\mathcal{P})$ also has a unique continuous solution. That is, for each type w of applicants, the target set of wages $T(w)$ is singleton and continuous in w. Use $T(w)$ to refer to 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 applies only to wages higher than his current wage. That is, $T(w) > w$ for all $w \neq w_0$, provided that $T(w)$ is nonempty. This is because

[18] When the solution to $(\mathcal{P})$ is not unique but finite, the number of equilibrium wage levels is still finite but there will be multiple ladders. The difficulty to establish the uniqueness and continuity of the solution lies in the fact that the properties of the solution to $(\mathcal{P})$ depend on the properties of the value functions $\{J_e(w), J_f(w)\}$. But, in turn, the properties of these functions depend on whether $(\mathcal{P})$ has a unique solution. We cannot verify such uniqueness analytically, but the solution was indeed unique in all numerical examples examined.

the expected surplus of application is nonpositive if an employed worker applies to a wage equal to or below his current wage.

The same result holds for an unemployed applicant if $\lambda_0 \geq \lambda_1$. However, if $\lambda_0 < \lambda_1$, then an employed worker receives job opportunities more often than an unemployed worker. In this case, an unemployed worker may be willing to accept a wage lower than the unemployment benefit so as to gain a better access to higher wages. Thus, $T(w_0) < w_0$ is possible only if $\lambda_0 < \lambda_1$.

Second, starting from any equilibrium wage w (including w_0), the path of future equilibrium wages 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 nonnegative 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. Note that $T^{-i}(w)$ is strictly decreasing in i for any given w and the difference $[T^{-(i-1)} \times (w) - T^{-i}(w)]$ 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 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. An applicant applies only to the wage one rung above his current wage. This happens not because the applicant does not observe job openings at higher wages as in the BM model, but rather because it is optimal to climb the ladder one rung at a time. The job openings one level above the applicant's current wage on the ladder provides a higher employment probability than jobs at higher wages and, as such, they provide a best trade-off between the employment probability and wage.

The above argument shows that the BM equilibrium cannot be an equilibrium with directed search. It is important to see how the above argument breaks down in the BM framework. There, workers do not choose which wage to apply to, because search is assumed to be undirected. Rather, each applicant is exogenously matched with job openings all with positive probability, despite the fact that a high-wage offer yields a higher expected surplus than a low-wage offer. As a result, a recruiting firm receives applicants at all wage levels with positive probability, irrespective of the wage offer. This feature of the BM model implies that if w_1 is the lowest equilibrium wage and if w_2 ($>w_1$) is another equilibrium wage, then any wage offer $w' \in (w_1, w_2)$ will receive some unemployed applicants who will accept the offer. This property in the BM model supports a continuum of wages as equilibrium wages. Directed search destroys this property by allowing applicants to choose which wage offer to apply to. In this case, only the wages that are optimal to the applicants will receive applications with positive probability. In the above example, wages in (w_1, w_2) will not attract any applicants and hence they cannot be equilibrium wages. In Section 7 we will discuss in more detail the differences between our model and the BM model.

3.3. *Simplifying the Notation and Imposing the Restriction Off the Equilibrium.* Each equilibrium wage on the ladder is necessarily a mass point of the wage distribution. Let the set of equilibrium wages be $\Omega = (w_i)_{i=1}^{M}$, where $w_i = T^i(w_0)$ and M is the (finite) number of rungs on the ladder. We depict the wage ladder in Figure 2, where the arrows indicate the directions in which workers apply for jobs. Exogenous separation is not depicted here. As discussed before, $w_1 < w_0$ is possible only when $\lambda_0 < \lambda_1$. Moreover, the ladder collapses into one rung if on-the-job is shut down, i.e., if $\lambda_1 = 0$.

With the ladder structure, we can simplify the notation. For each $w_i \in \Omega$, denote $n_i = n(w_i)$, $v_i = v(w_i)$ 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 (3) and (4) as follows:

$$h_i = h(a_i) \equiv 1 - e^{-a_i}, \quad q_i = q(a_i) \equiv (1 - e^{-a_i})/a_i \tag{13}$$

Clearly, $h'(a) > 0$ and $q'(a) < 0$. The probability with which a worker at w_{i-1} endogenously separates from the job is $\rho_{i-1} = \lambda_1 q_i$ for $i \geq 2$ and $\rho_0 = \lambda_0 q_1$. Moreover, because $p(w_i, w_{i-1}) = 1/(v_i K)$ for $i \geq 1$, (1) becomes

$$v_i = \lambda_1 n_{i-1}/(a_i k) \text{ for } i \geq 2, \text{ and } v_1 = \lambda_0 u/(a_1 k) \tag{14}$$

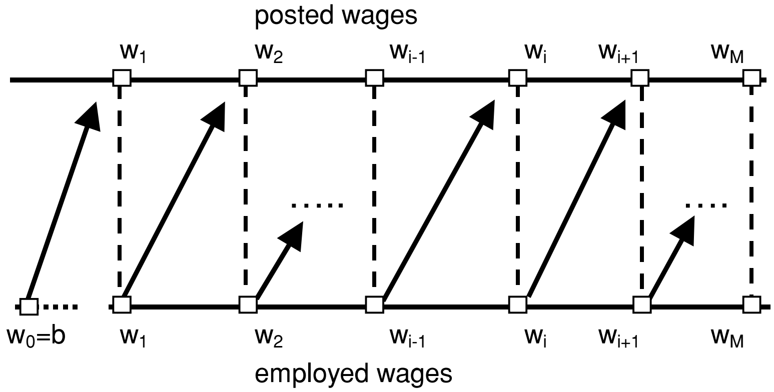

FIGURE 2

THE WAGE LADDER

Now that the equilibrium must be a wage ladder, we still need to show that such an equilibrium exists. This is not an easy task. With on-the-job search, an applicant for a job cares about what opportunities the job will generate in the future. Thus, current applicants' and recruiting firms' strategies depend on the strategies of future recruiting firms, which are an equilibrium object. We will explore the recursive structure of the wage ladder to find an equilibrium and to characterize the analytical properties of the equilibrium.

In any attempt to establish an equilibrium, we must check deviations from equilibrium strategies. Since one firm's deviation sends its prospective employee off the equilibrium path, it is necessary to specify what firms would do to the applicants whose wages happen to lie outside Ω.[19] Because all workers have the same productivity, it is natural to maintain the following restriction on beliefs off the equilibrium path:

Restriction (Off-eqm): The employment probabilities satisfy $q(w, w') = q(w)$ for all $w', w \in \mathcal{W}$, including $w' \notin \Omega_0$ and $w \notin \Omega$.

4. A CANDIDATE EQUILIBRIUM

In this section, we construct a candidate equilibrium by examining a restricted set of deviations. Suppose that all (current and future) recruiting firms, except one, post offers in the set $(w_i, q_i)_{i=1}^{M}$. The deviating firm offers (w^d, q^d), where $w^d \notin \Omega$. If the firm successfully hires a worker, it pays the wage w^d until the worker separates; if the firm fails to hire a worker, the firm returns to the equilibrium recruiting strategy next period. In the construction of the candidate equilibrium, we restrict the deviation as follows:

Temporary Restriction (One-rung): There exists $w^* \in \Omega$ such that w^d attracts the same type of applicants as does w^* and a worker who gets w^d will apply to the same wage in the future as will a worker who gets w^*, with the only exceptions being (i) $w^d > w_M$ and w^d attracts only w_M-applicants, and (ii) $w^d < w_1$ and the worker who gets w^d will apply to w_1 in the next application.

This temporary restriction makes the deviation comparable to the equilibrium offer $(w^*, q(w^*))$. (Try depicting the deviation in Figure 2.) After constructing the candidate, we will eliminate this restriction in Section 5 and find conditions under which the candidate is indeed an equilibrium.

4.1. *Wages Lower Than the Highest Level.* Examine the equilibrium wage $w_i \in \Omega$, where $1 \le i \le M - 1$. Consider an individual firm's deviation $w^d \in (w_{i-1}, w_{i+1})$. Since the purpose of examining this deviation is to derive conditions on

[19] If 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.

(w_i, q_i), we take w_i to serve the role of w^* in the restriction (One-rung); that is, w^d attracts only w_{i-1}-applicants and a worker getting w^d will apply to w_{+1} in the next application.[20]

Let a^d be the queue length of the applicants at wage w_{i-1} whom w^d attracts. Then, the deviating firm's hiring probability is $h(a^d)$ and each applicant's employment probability is $q(a^d)$, where the functions $h(a)$ and $q(a)$ are defined by (13). After an applicant gets the job, his future quit rate is $\lambda_1 q_{i+1}$ (note that we are invoking 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 (8) and (9):

$$J_f(w^d) = \frac{y - w^d}{r + \sigma + \lambda_1 q_{i+1}} \tag{15}$$

$$J_e(w^d) = \frac{w^d + \sigma J_u - \lambda_1 S + \lambda_1 q_{i+1} J_e(w_{i+1})}{r + \sigma + \lambda_1 q_{i+1}} \tag{16}$$

Because the deviating firm takes future recruiting firms' strategies as given, it takes q_{i+1} and $J_e(w_{i+1})$ in the above formulas as given.

It is more convenient to formulate the deviator's choices as (w^d, a^d), rather than (w^d, q^d). The optimal choices (w^d, a^d) solve the following problem similar to $(\mathcal{P})$:

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

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 (15) and (16), respectively, with w_i replacing w^d. Note 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. This share decreases endogenously with the queue length a_i.

In addition, free entry of firms drives the value of a vacancy to zero. That is,

$$\frac{C}{h_i} = J_f(w_i) = \frac{y - w_i}{r + \sigma + \lambda_1 q_{i+1}} \tag{17}$$

[20] The same analysis applies to two other possible cases, with relabeling. The first is that $w^d \in (w_i, w_{i+1})$ and w_{i+1} serves the role of w^* in the restriction (One-rung). In this case, treat w^d as a downward deviation from w_{i+1} rather than a deviation from w_i. The second case is that $w^d \in (w_{i-1}, w_i)$ and w_{i-1} serves the role of w^* in the restriction (One-rung). In this case, treat w^d as an upward deviation from w_{i-1}.

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

$$J_e(w_i) - J_e(w_{i-1}) = C/f_i \tag{18}$$

$$E(w_{i-1}) = q_i\,[J_e(w_i) - J_e(w_{i-1})] = Cq_i/f_i \tag{19}$$

where

$$f_i = f(a_i) \equiv q(a_i)\,(e^{a_i} - 1 - a_i) \tag{20}$$

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 (19), this requirement becomes

$$a_i \le \bar{a}, \quad \text{where } e^{\bar{a}} - 1 - \bar{a} = C/S \tag{21}$$

4.2. *Highest Wage in Equilibrium.* Consider a deviation $w^d \in (w_M, y]$. This deviation may attract the applicants at w_{M-1}, in which case the deviation satisfies the restriction (One-rung). For this deviation to be not profitable, the highest wage w_M must satisfy (17)–(21) for $i = M$, with $q_{M+1} = 0$.

The deviation may also attract w_M-applicants, which no equilibrium wage does. (This is the exception (i) in restriction (One-rung).) To ensure that such a deviation is not profitable, we need an additional condition. Facing an opening $w^d > w_M$, the applicants employed at w_M will apply to it as long as the expected surplus from the application is equal to the fixed cost of application. Thus, the queue length of applicants for the deviating firm, a^d, satisfies

$$q(a^d)\big[J_e(w^d) - J_e(w_M)\big] = S \tag{22}$$

where $J_e(w) = (w + \sigma J_u)/(r + \sigma)$ for both $w = w^d$ and w_M. The deviator's expected surplus is $h(a^d)J_f(w^d)$, where $J_f(w^d) = (y - w^d)/(r + \sigma)$. This deviation is not profitable if and only if the deviator's maximum expected surplus is less than the vacancy cost C. Solving the deviator's maximization problem subject to (22), we can write this requirement as

$$w_M > y - (r + \sigma)Se^{\bar{a}} \tag{23}$$

where $\bar{a}$ is defined in (21). Clearly, there is any $w_M \le y$ satisfying the condition only if $S > 0$.

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, 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 (21) and (23) for $i = M$ as requirements on the hiring probability at wage w_M, as follows:[21]

$$1 - (1 + \bar{a})e^{-\bar{a}} < h_M \leq 1 - e^{-\bar{a}} \tag{24}$$

4.3. *Recursive Characterization of the Candidate Equilibrium.* The conditions in the two previous subsections provide a recursive characterization of the candidate equilibrium. Pick up a number h_M that satisfies (24). Then, $q_{M+1} = 0$. Moreover, setting $i = M$ and $w^d = w_M$ in (16) and (17), we get

$$a_M = -\ln(1 - h_M), \quad q_M = h_M / a_M$$
$$w_M = y - (r + \sigma)C / h_M, \quad J_e(w_M) = (w_M + \sigma J_u)/(r + \sigma)$$

Starting from this highest rung of the wage ladder, we can recursively compute the characteristics at lower rungs, as stated in the following proposition (see Section B of the Appendix for a proof).

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

$$h_{M-j} = \frac{(r + \sigma + \lambda_1 q_{M-j+1})C}{y - w_{M-j}} \tag{25}$$

$$a_{M-j} = -\ln(1 - h_{M-j}), \quad q_{M-j} = h_{M-j}/a_{M-j} \tag{26}$$

$$w_{M-j-1} = w_M + \lambda_1 S - \frac{\lambda_1 q_{M-j} C}{f_{M-j}} - (r + \sigma)C \sum_{t=0}^{j} \frac{1}{f_{M-t}} \tag{27}$$

$$J_e(w_{M-j-1}) = \frac{\sigma J_u + w_M}{r + \sigma} - C \sum_{t=0}^{j} \frac{1}{f_{M-t}} \tag{28}$$

In addition, (25) *and* (26) *hold for $j = M - 1$, and J_u satisfies*

$$J_u = \frac{1}{r}\left[b - \lambda_0 S + \lambda_0 C \frac{q_1}{f_1}\right] \tag{29}$$

ASSUMPTION 1. *Define $\bar{a}$ by* (21)*. Assume that the following conditions hold*

$$b \leq y + \lambda_0 S - C\left[\frac{(r + \sigma)e^a + \lambda_0}{e^a - 1 - a}\right]_{a = \bar{a} - \ln(1 + \bar{a})} \tag{30}$$

[21] For $i = M$, (21) is equivalent to $h_M \leq h(\bar{a}) = 1 - e^{-\bar{a}}$. To rewrite (23), note that $h_M J_f(w_M) = C$ in equilibrium and $J_f(w_M) = (y - w_M)/(r + \sigma)$.

$$(r+\sigma)/\lambda_1 > f(\bar{a})/\bar{a} \tag{31}$$

The condition (30) ensures that there is at least one wage level that yields a higher present value to the workers than unemployment, whereas (31) is a technical condition necessary for exploring Proposition 1.

The following proposition describes some important properties of the computed sequence (a, q, h) (see Section C of the Appendix for a proof).

PROPOSITION 2. *For any given h_M that satisfies* (24), *the sequence constructed in Proposition 1 has the following monotonicity properties for all* $2 \le i \le M$:

$$a_{i-1} < a_i \le \bar{a}, \; h_{i-1} < h_i, q_{i-1} > q_i \tag{32}$$

$$(r+\sigma)/\lambda_1 > f(a_{i-1})/a_{i-1} \tag{33}$$

$$a_{i-1} > a_i - \ln(1 + a_i) \tag{34}$$

$$da_i/dh_M > 0, \quad dw_i/dh_M > 0 \tag{35}$$

The recursive method in Proposition 1 generates a sequence $(h_i, a_i, q_i, w_i, J_e(w_i))_{i=1}^M$ for each given h_M. For the sequence to be an equilibrium, the values of M and h_M must be such that $J_e(w_1)$ satisfies (18) for $i = 1$. That is, $\Delta = 0$, where $\Delta \equiv J_e(w_1) - J_u - C/f_1$. Setting $j = M-2$ in (28) to obtain $J_e(w_1)$ and substituting (29) for J_u, we write Δ as

$$\Delta(M, h_M) = \frac{w_M - b + \lambda_0 S}{r+\sigma} - \frac{C\lambda_0 q_1}{(r+\sigma) f_1} - C\sum_{t=0}^{M-1}\frac{1}{f_{M-t}} \tag{36}$$

Note that Δ does not depend on J_u directly because the computed a sequence and the number w_M do not (see Proposition 1). Solving (M, h_M) involves iterations on Δ until $\Delta(M, h_M) = 0$. We describe the iteration procedure in Section B of the Appendix and prove the following proposition.

PROPOSITION 3. *Fix h_M at any value h^* that satisfies* (24) *and compute the a sequence according to Proposition* 1. *Then, there exists an integer $M(h^*) \ge 1$ such that $\Delta(M', h^*) \ge 0$ for all $M' \le M(h^*)$ and $\Delta(M', h^*) < 0$ for all $M' \ge M(h^*) + 1$. Let $M^{**} = M(1 - (1+\bar{a})e^{-\bar{a}}) + 1$. Then, there exists an equilibrium value of h_M if and only if*

$$\Delta(M^{**}, 1 - e^{-\bar{a}}) \ge 0 \tag{37}$$

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 $M^{**} + 1$.

With the equilibrium values of (h, a, q, M), we can calculate 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,

$$(38) \qquad \begin{aligned} (\sigma + \lambda_1 q_{i+1}) n_i &= \lambda_1 n_{i-1} q_i, \text{ for all } 2 \le i \le M \\ (\sigma + \lambda_1 q_2) n_1 &= \lambda_0 u q_1 = \sigma(1-u), \text{ and } du = 1 - \sum_{i=1}^{M} n_i \end{aligned}$$

These equations solve for u and $(n_i)_{i=1}^{M}$. Second, we can solve the distribution of vacancies $(v_i)_{i=1}^{M}$ from (14) and the requirement $\sum_{i=1}^{M} v_i = 1$. These equations also solve for k—the ratio of vacancies to workers. This completes the construction of the candidate equilibrium.

5. THE WAGE LADDER IS AN EQUILIBRIUM

We now find conditions under which the candidate equilibrium is indeed an equilibrium. This will be done in two steps. First, we consider only equilibrium wages and verify that workers indeed climb the ladder one rung at a time. This requirement is necessary for the value functions used in the construction of the candidate equilibrium to be valid. Second, we eliminate the restriction (One-rung) imposed in Section 4 and show that the candidate equilibrium can survive all deviations. The restriction (Off-eqm) is maintained throughout. Whenever possible, we suppress the index $i = M - j$ and use the subscript $\pm t$ to stand for $M - j \pm t$. The readers who are more interested in the properties of the equilibrium can skip this section and go directly to Section 6.

5.1. *Workers Climb the Ladder One Rung at a Time.* In this subsection, we confine wages to the equilibrium set Ω and find the *one* condition under which applicants employed at wage w_{-1} indeed apply only to job openings at wage w, for any wage w_{-1} on the ladder. We proceed as follows: We must verify that a w_{-1}-applicant does not have any incentive to apply to any wage strictly greater than w on the wage ladder. That is potentially a large number of deviations to check. The reader can refer to Table 2 to see how we can greatly simplify that task. Take a w_i-applicant, $i \in \{0, 1, \ldots, M-2\}$. To verify that he prefers to apply to w_{i+1} rather than any other wage on Ω, we verify that he prefers to apply to w_{i+j} than to w_{i+j+1} for $j \in \{1, 2, \ldots, M-i-1\}$. If all these conditions are met, workers only apply to the next wage on the ladder. The conditions are represented by the upper triangle in Table 2. We first derive Lemma 2 to show that it is enough to just check the conditions given by the diagonal of that triangle—the terms in brackets. If a condition on a diagonal term is satisfied, so are all the conditions in the corresponding column. Then we derive Lemma 3 to show that it is even enough to only check the very last condition on that diagonal—the underlined term.[22] This implies that a single condition, namely that a w_{M-2}-applicant prefers

[22] Lemma 4 follows directly from Lemma 3.

TABLE 2
POSSIBLE DEVIATIONS WITHIN THE EQUILIBRIUM SET Ω

Type	Prefers to Apply to w_i Than to w_j : w_i/w_j						
b	$[w_1/w_2]$	w_2/w_3	w_3/w_4	w_4/w_5	$\cdots$	$\cdots$	w_{M-1}/w_M
w_1		$[w_2/w_3]$	w_3/w_4	w_4/w_5	$\cdots$	$\cdots$	w_{M-1}/w_M
w_2			$[w_3/w_4]$	w_4/w_5	$\cdots$	$\cdots$	w_{M-1}/w_M
w_3				$[w_4/w_5]$	$\cdots$	$\cdots$	w_{M-1}/w_M
$\vdots$					$\ddots$	$\ddots$	$\vdots$
$\vdots$						$\ddots$	$\vdots$
w_{M-2}							$[w_{M-1}/w_M]$

to apply to wage w_{M-1} rather than to w_M is enough to ensure that no applicant will want to apply anywhere else than on the next rung of the ladder.

For the strategy of applying to the next rung on the wage ladder to be optimal, the expected surplus that a w_{-1}-applicant obtains from applying to wage w must be larger than or equal to that from applying to other higher wages. That is, for all $j \in \{1, 2, \ldots, M-1\}$ and all $t \in \{1, 2, \ldots, j\}$, the following inequality must hold

$$q_{M-j}[J_e(w_{M-j}) - J_e(w_{M-j-1})] \geq q_{M-j+t}[J_e(w_{M-j+t}) - J_e(w_{M-j-1})] \tag{39}$$

For a fixed wage w on the ladder, which applicants have the strongest incentive to leap over w? The applicants currently employed one level below w do. To leap over w for higher wages, the applicant must care about the wage gain sufficiently more than about the employment probability. Because the applicants' indifference curves have the single-crossing property, the applicants at wage w_{-1} care about the wage increase the most among all applicants employed below w on the ladder. Thus, if the applicants employed at wage w_{-1} do not leap over w, then neither do applicants employed at wages below w_{-1}. We formalize this intuition in the following lemma (see Section A of the Appendix for a proof).

LEMMA 2. *For all $j \in \{1, 2, \ldots, M-1\}$ 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.*

With the above lemma, the applicants indeed climb the wage ladder one rung at a time if the applicants on each rung do not leap over the next rung. That is, for each $j \in \{1, 2, \ldots, M-1\}$, it suffices to verify (39) for only $t = 1$. Using (18), we can rewrite (39) for $t = 1$ as

$$\frac{q(a)}{q(a_{+1})} - 1 - \frac{f(a)}{f(a_{+1})} \geq 0 \tag{40}$$

This contains $(M-1)$ inequality conditions in total.

The single-crossing property suggests that we might be able to further reduce the number of inequality conditions to one. Because an applicant's preference for a wage gain increases with the applicant's current wage, the incentive to leap over the next rung on the ladder strengthens with the applicant's current wage. For all workers to climb the ladder one rung at a time, it suffices to find the condition under which the applicants at wage w_{M-2} do not leap over the next rung.

To formalize the above intuition, let us define $\phi(a_{+1})$ as the solution for a to the equality form of (40) under given a_{+1}. Because the left-hand side of (40) is a decreasing function of a, the inequality is equivalent to $a \le \phi(a_{+1})$. We prove the following lemma in Section D of the Appendix:

LEMMA 3. *The function $\phi(\cdot)$ 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 $a \le \phi(a_{+1})$, then $a_{-1} < \phi(a)$.*

Property (iii) captures our intuition. Thus, to verify (40) for all $j \in \{1, 2, \dots, M-1\}$, it suffices to verify it for $j = 1$. For $j = 1$, Equation (40) becomes $a_{M-1} \le \phi(a_M)$, which is equivalent to $h_{M-1} \le h(\phi(a_M))$. Use (C.2) in Section C of the Appendix to substitute h_{M-1}; we can express $h_{M-1} \le h(\phi(a_M))$ as the condition in the following lemma.[23]

LEMMA 4. *Given the wage levels computed in Proposition 1, it is optimal for applicants to apply only to the next wage level, rather than higher ones, if and only if*

$$\frac{r+\sigma}{\lambda_1} \ge \left(\frac{q_M}{h(\phi(a_M))} - \frac{q_M}{f_M} + \frac{S}{C}\right) \Big/ \left(\frac{1}{h_M} + \frac{1}{f_M} - \frac{1}{h(\phi(a_M))}\right) \tag{41}$$

5.2. *Eliminating Restriction (One-rung).* Now we reconsider the deviation examined in Section 4 but eliminate the restriction (One-rung). It suffices to consider a deviation $w^d \in (w_{-1}, w)$. This deviation can violate the restriction (One-rung) either in the type of applicants it attracts and/or the wage to which the worker who gets w^d will apply in the future. Table 3 summarizes all such deviations. The cells marked with $\times$ indicate that the corresponding deviations conform with the restriction (One-rung).

[23] Using (C.2) with $j = 0$, we can write the condition $h_{M-1} \le h(\phi(a_M))$ as

$$0 \le \frac{r+\sigma}{\lambda_1}\left(\frac{1}{h_M} + \frac{1}{f_M} - \frac{1}{h(\phi(a_M))}\right) - q_M\left(\frac{1}{h(\phi(a_M))} - \frac{1}{f_M}\right) - \frac{S}{C}$$

Note that $\frac{1}{h} + \frac{1}{f} = [1 - (1+a)e^{-a}]^{-1}$. So, the fact $\phi(a) > a - \ln(1 + a)$ implies $\frac{1}{h(a_M)} + \frac{1}{f(a_M)} > \frac{1}{h(\phi(a_M))}$. Thus, we can rewrite the above condition as (41). Moreover, if the right-hand side of (41) is an increasing function of a_M (which seems true from numerical examples), then we can obtain a sufficient condition for (41) by replacing a_M with $\bar{a}$ and C/S with $(e^{\bar{a}} - 1 - \bar{a})$.

TABLE 3
DEVIATION w^d THAT VIOLATES RESTRICTION (ONE-RUNG)

	Destination of Prospective Workers at w^d		
Source of Applicants to w^d	w	w_{+1}	w_{+t} $(t \geq 2)$
w_{-1}	I	$\times$	III
w_{-2}	$\times$	II	III
w_{-t} $(t \geq 3)$	IV	IV	III

The following lemma shows that type III deviations are not profitable, whereas type IV deviations are not profitable if type II deviations are not (see Section A of the Appendix for a proof).

LEMMA 5. *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* $t \geq 3$.

The intuition for the above lemma is again the single-crossing property of the applicants' indifference curves. This is obvious for part (ii), because part (ii) extends Lemma 2 to wages outside the equilibrium support and Lemma 2 relies on the single-crossing property. For part (i), the single-crossing property implies that, because $w > w^d$, an applicant at w is more willing to sacrifice the employment probability for the wage gain than an applicant at w^d. Since 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.

Now we turn to type I and type II deviations. A type I deviation is profitable only when the support of the wage distribution is too sparse, whereas a type II deviation is profitable only when the support is too dense. (Try depicting these two types of deviations in Figure 2.) Note that the exception (ii) in restriction (One-rung) is a type I deviation. Also, a type II deviation is meaningful only when there are three or more rungs on the ladder.

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,

$$J_f^d(w^d) = \frac{y - w^d}{r + \sigma + \lambda_1 q}$$

$$J_e^d(w^d) = \frac{w^d + \sigma J_u - \lambda_1 S + \lambda_1 q J_e(w)}{r + \sigma + \lambda_1 q}$$

These functions are different from those in (15) and (16), because the worker's future application direction is different from the one depicted in Figure 2 (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 both be 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

$$q^d\big[J_e^d(w^d) - J_e(w_{-1})\big] = E(w_{-1}) = Cq/f \tag{42}$$

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

$$w^d = w_{-1} + (r + \sigma + \lambda_1 q)\,Cq/(fq^d) \tag{43}$$

The deviating firm's expected surplus is $\pi(a^d) = h(a^d)J_f^d(w^d)$. Substituting $(y - w_{-1})$ from (25) and $(w^d - w_{-1})$ from (43), $\pi(a^d)$ becomes

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

The surplus $\pi(a^d)$ is maximized at $a^d = a^*$, where $\pi'(a^*) = 0$. Thus, $a^* = \ln(f/qh_{-1})$, and the maximum expected surplus of the deviating firm is

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

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, for all $a^d > a$, we have $q^d < q$. Since $J_e^d(w^d) < J_e(w)$, then w^d yields a lower expected surplus to a w_{-1}-applicant than w does, which contradicts (42).

Therefore, 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 enough to cover the vacancy cost.

Now consider a type II deviation $w^d \in (w_{-1}, w)$. The deviating firm's ex post value $J_f^d(w^d)$ and the employee's value $J_e^d(w^d)$ are

$$J_f^d(w^d) = \frac{y - w^d}{r + \sigma + \lambda_1 q_{+1}} \tag{44}$$

$$J_e^d(w^d) = \frac{w^d + \sigma J_u - \lambda_1 S + \lambda_1 q_{+1} J_e(w_{+1})}{r + \sigma + \lambda_1 q_{+1}} \tag{45}$$

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.

It can be shown that the deviator's maximum expected surplus exceeds C if (IIa) is the only constraint. For the deviation to be not profitable, the constraint (IIc) must be binding sufficiently. The following lemma gives the corresponding requirement (see Section E of the Appendix for a proof).

LEMMA 6. *Define β^* by the following equation:*

$$h(\beta^*) = 1 \Bigg/ \left[\frac{1}{h} + \frac{q_{+1}}{(q - q_{+1})\, f_{+1}}\right] \tag{46}$$

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

$$\beta^* - h(\beta^*)e^{a_{-1}} + \frac{r + \sigma + \lambda_1 q_{+1}}{r + \sigma + \lambda_1 q}\,(e^{a_{-1}} - 1 - a_{-1}) \geq 0 \tag{47}$$

We summarize the results on existence as follows:

PROPOSITION 4. *Maintain Assumption* 1 *and Restriction (Off-eqm). An equilibrium exists if and only if the following conditions hold:* (37) *in Proposition* 3*,* (41) *in Lemma* 4*, and (if $M \geq 3$)* (47) *in Lemma* 6*.*

Let us make two remarks. First, the conditions in the above proposition may fail to hold in certain parameter regions, in which case a stationary equilibrium does not exist. This does not necessarily imply that a Nash equilibrium does not exist in such parameter regions. For example, there might be nonstationary equilibria, which we do not know how to formulate. Second, when an equilibrium exists, it may not be unique. As Proposition 3 indicates, the hiring probability at the highest wage may lie in a range of values.

6. PROPERTIES OF THE EQUILIBRIUM

6.1. *Analytical Properties.* Proposition 2 reveals the following realistic properties of the equilibrium:

- A firm is more likely to succeed in hiring at a higher wage than at a lower wage, whereas an applicant is more successful in 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's wage, on average, increases with his employment duration, because a longer duration indicates that he has likely climbed more rungs on the ladder. Also, the higher wage a worker had just before becoming unemployed, the longer it will take for him to return to that wage.
- A worker's quit rate decreases with wage. The quit rate of a worker at wage w is $\rho(w) = \lambda_1 q(a_{+1})$. Because the employment probability for the next wage ($q(a_{+1})$) decreases as wage increases, the quit rate falls. By the previous property, a worker's quit rate also decreases on average with the employment duration. This also implies that there is a positive correlation between wage and the average tenure at that wage. As the workers employed at high wages quit less frequently than the workers employed at low wages, the average tenure at high wages must be longer than at low wages in the steady state.[24]

These properties are similar to those in the BM model. To explore the properties that are unique to the wage ladder, we establish the following proposition in Section F of the Appendix.

PROPOSITION 5. *For any given h_M that satisfies* (24)*, the computed sequence satisfies: (i) $E(w_{-1}) > E(w)$ $(\geq S)$, and if* (41) *in Lemma* 4 *holds, then (ii) $w - w_{-1} > w_{+1} - w$.*

Because $E(w_{-1}) \geq S$, all workers except those at the highest wage are willing to incur the fixed cost to apply to higher wages. Moreover, $E(w_{-1}) > 0$ implies $J_e(w) > J_e(w_{-1})$, so that the value of employment to a worker indeed increases with wage.

The ladder structure has the following novel implications on wage mobility:

- Wage mobility is limited: In each period a worker either climbs one rung up on the ladder or stays at the current wage or transits into unemployment.
- The gap between two adjacent rungs on the ladder shrinks as a worker climbs the ladder. So, the higher the wage, the smaller the wage gain in the next job change.[25]

[24] Of course, this is not necessarily true in the transition to the steady state. In particular, if the number of workers employed at low wages in the initial state of the economy exceeds the number in the steady state, then many high-wage workers are ones who had just transited into those wages. This transitional effect disappears in the steady state.

[25] A referee, whom we thank, made us aware of recent work by Barlevy (2003) who finds, using record theory and a Burdett–Mortensen on-the-job search model, that the null hypothesis that all wage gains are constant in percentage terms cannot be rejected.

- 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 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.

Limited wage mobility is broadly consistent with the evidence documented by Buchinsky and Hunt (1999). As discussed in the introduction, the BM model generates the opposite pattern that a worker's transition probability increases with the target wage. It is worth repeating that the limitation on wage mobility is endogenous in our model, because the workers choose optimally to climb the wage ladder gradually.

Our model also generates wage distributions that differ from those in the BM model, as stated in the following proposition (see Section F of the Appendix for a proof):

PROPOSITION 6. *The density 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*

$$\frac{\sigma}{\lambda_1} > \frac{1-(1+\bar{a})\,e^{-\bar{a}}}{\bar{a}-\ln(1+\bar{a})} \tag{48}$$

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 $\sigma/\lambda_1 < q(\bar{a})$.*

To appreciate the results in the above proposition, recall that the BM model of undirected search generates an increasing and convex density function of offer wages, which leads further to an increasing and convex density function of employed wages. Sufficient heterogeneity among workers and/or jobs is needed to generate a decreasing density at high wages in the BM model (e.g., van den Berg and Ridder, 1998).

Our model generates a decreasing density function of offer wages among homogeneous workers. This result is easy to understand, given the ladder structure. 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, that is, 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 $h_{+1} > h$. Therefore, the density of offer wages necessarily decreases with wage.

Note that the above explanation uses the equilibrium feature that a high-wage recruiting firm recruits more successfully than a low-wage firm. At this point, one must ask the following question: Since the BM model also has this feature, why does that model generate an increasing density of offer wages instead? The answer is that the BM model does not generate limited wage mobility as our model does. In particular, since search is undirected in the BM model, a firm's matching rate is independent of the firm's wage offer and is the same for all firms. Because a high wage is more likely to be accepted by a randomly selected applicant, a high-wage recruiting firm is more successful in hiring. This higher matching success, together with the higher probability of retaining a worker, is more than compensating for the higher wage offer. To make firms indifferent between posting different wages, there must be more firms competing against each other by offering high wages. This generates the increasing density function of offer wages in the BM model.

To make the density function of offer wages decrease at high wages, one must sufficiently reduce the advantage of recruiting at high wages. An arbitrary way to do so in the BM model is to assume that firms' matching rates are sharply decreasing in wages, which is essentially what Postel-Vinay and Robin (2002) have done.[26] It is questionable whether it is optimal in the BM model for a high-wage recruiting firm to reduce its matching rate. In our model, such decreasing matching rates are not optimal—each high-wage recruiting firm optimally uses a strategy to induce a high-matching rate for itself. Despite this high-matching rate for each high-wage firm, the base of applicants for high-wage firms is smaller than for low-wage firms, because low-wage workers choose not to apply to high-wage firms. This is the reason why fewer firms recruit at a high wage than at a low wage.

Finally, we turn to the density of employed wages. The density of employed wages can also be decreasing, but it is not always so. The ambiguity arises 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 from a low wage is also higher than from 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 even be monotonic with respect to wages. The numerical examples in the next subsection will display some decreasing density functions of employed wages.

6.2. *Numerical Examples.* Consider the following parameter values: $y = 1000$, $b = 0$, $C = 60$, $S = 1$, $r = 0.02$, $\lambda_1 = 0.025$, $\lambda_0 = 0.2$, and $\sigma = 0.125$. These parameter values satisfy all equilibrium requirements specified in Proposition 4. Under these parameter values, there is a unique equilibrium and the wage ladder has four rungs. The unemployment rate is $u = 40.9\%$ and the overall vacancy-worker ratio is $k = 0.41$. Other characteristics of this equilibrium are summarized in Table 4.

The results in Table 4 confirm the properties in Propositions 2 and 5. In addition, a notable feature is that the small difference between wages induces large

[26] Postel-Vinay and Robin (2002) assume that firms' matching (contact) rates are sharply decreasing over firms' productivity. Since high-productivity firms are also likely to offer high wages, firms' matching rates (on average) are sharply decreasing over wages.

TABLE 4

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	0.2	1.0	1.77	46.9	83.0
4	990.8	<0.1	0.1	2.94	32.2	94.7

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 sharply decreasing function of wages in this example. 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.

On-the-job search is important for the wage ladder. We illustrate this importance in Table 5, by changing λ_1 while fixing other parameters (including λ_0). When $\lambda_1 = 0$, on-the-job search is shut down and, as previous models of directed search predict, the wage distribution is degenerate. However, the wage ladder becomes nondegenerate as soon as workers can search while employed. Even for a very small value $\lambda_1 = 10^{-4}$, the support of equilibrium wages is almost as dispersed as when $\lambda_1 = 0.025$. Thus, the support of the wage distribution does not seem upper hemi-continuous in λ_1 at $\lambda_1 = 0$. If one uses the range of equilibrium wages to measure wage dispersion, then such dispersion will change dramatically when λ_1 approaches 0.

However, the equilibrium seems continuous at $\lambda_1 = 0$ in terms of wage distributions. When λ_1 is close to 0, the distributions of offer wages and employed wages both have almost a unity of mass at the lowest wage. As λ_1 increases, the densities of wage distributions become more spread out to higher wages, and so the

TABLE 5

VARIOUS DEGREES OF ON-THE-JOB SEARCH

	$\lambda_1 = 0$	$\lambda_1 = 10^{-4}$	$\lambda_1 = 0.01$	$\lambda_1 = 0.025$		$\lambda_1 = 0$	$\lambda_1 = 10^{-4}$	$\lambda_1 = 0.01$	$\lambda_1 = 0.025$
w_1	953.2	953.2	951.0	947.6	a_1	0.21	0.21	0.21	0.21
w_2		984.0	983.4	982.6	a_2		0.79	0.78	0.78
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
%	$\lambda_1 = 0$	$\lambda_1 = 10^{-4}$	$\lambda_1 = 0.01$	$\lambda_1 = 0.025$	%	$\lambda_1 = 0$	$\lambda_1 = 10^{-4}$	$\lambda_1 = 0.01$	$\lambda_1 = 0.025$
v_1	100	99.9	98.2	95.7	$\frac{n_1}{1-u}$	100	99.9	94.7	87.8
v_2		<0.1	1.8	4.0	$\frac{n_2}{1-u}$		<0.1	5.1	11.2
v_3		<0.1	<0.1	0.2	$\frac{n_3}{1-u}$		<0.1	0.2	1.0
v_4		<0.1	<0.1	<0.1	$\frac{n_4}{1-u}$		<0.1	<0.1	0.1

distributions become less skewed. Some other characteristics of the equilibrium, such as the support of the distribution Ω, the unemployment rate u, and market tightness k, change very little with λ_1 in the range we experimented with.

7. THE ROLE OF THE APPLICATION COST AND THE CONTRAST WITH BURDETT AND MORTENSEN

In Section 3.2, we explained why directed search destroys the BM equilibrium and creates a wage ladder. The difference between directed search in our model and undirected search in the BM model is the fundamental one between the two models. However, our model differs from the BM model in another characteristic—our model has a fixed cost of application whereas the BM model does not have such a cost. In this section, we will first clarify the role of the application cost in our model and then compare our model with the BM model in the presence of the application cost.

We explained in Section 3.2 that *any* positive application cost S is enough to restrict attention to a wage ladder type of equilibrium. We also explained in Section 4.2 that a small application cost is necessary for obtaining an equilibrium by preventing a profitable deviation to posting a wage higher than w_M. Notice, however, that having costly applications does *not* affect the application decisions at other rungs, since $E(w_i) > S$ for $i < M$ (see Proposition 5). That is, the constraint $E(w_i) \geq S$ does not bind for any application except for the application to wages equal to or higher than w_M.

Because of this limited role of the application cost, we can make most applications free without affecting the results. For example, consider the following assumption of differential costs of application: "Every applicant who is currently employed at wages higher or equal to an exogenous level w_H must incur a fixed cost S of application, while the applicants who are currently employed at wages lower than w_H do not need to incur the cost." If w_H is set at the level w_M, then most applicants do not need to incur the application cost, and yet the equilibrium is the same as the one in our model.[27]

Now, we explore the equilibrium in the BM model when the application cost exists. Suppose that every applicant must incur a positive cost S to apply to a job. Then, it is well known that the BM model exhibits no wage dispersion at all, no matter how small the cost is (an argument similar in spirit to Diamond, 1971). To show this result, suppose, to the contrary, that there exists a nondegenerate wage distribution in the presence of the application cost. It is easy to show that the worker's optimal strategy dictates that above a certain wage w_{ns} (close enough to the upper end of the support $\bar{w}$), employed workers are not willing to engage in costly search to improve their condition. Because the application cost is strictly positive, that stopping wage w_{ns} is strictly less than the upper support of the distribution $\bar{w}$. But then, an individual firm at $\bar{w}$ could increase profits by decreasing

[27] Postel-Vinay and Robin (2004), in a random search environment, allow for some outside offers to cost effort and for some other offers to arrive at no cost, while on the job. Their model also differs from ours due to firm heterogeneity and because they give firms the possibility to match outside offers. Depending on the distribution of productivity across firms, a decreasing wage distribution may result.

the wage it offers by $\varepsilon < \bar{w} - w_{ns}$, without affecting the probability of having its offer accepted. Hence, an equilibrium with wage dispersion in a BM world with fixed application cost is not possible.

This discontinuity of the equilibrium in the BM model with respect to S complicates the comparison between our model and the BM model. However, there are still two ways to differentiate the two models. First, the BM model does not generate a wage ladder, regardless of whether there is an application cost, whereas our model does with a small application cost. Second, it is possible to compare the results of the two models directly under the above-mentioned differential costs of application, that is, under the assumption that the application cost S must be incurred only when the applicant is currently employed at a wage equal to or higher than $w_H = w_M$. Suppose that S is sufficiently small (but positive), so that w_H is sufficiently close to y. In this case, the BM equilibrium is unaffected by such application costs thus still exhibiting a nondegenerate wage distribution, because the highest equilibrium wage in the BM model is lower than w_H. Since the equilibrium in our model is also unaffected by the differential application costs, the two models differ from each other only in whether search is directed or undirected. All the comparisons that we have made between the two models remain valid. In particular, the wage distribution in the BM model does not constitute a wage ladder and it generates wage mobility that is unrealistically high.

Finally, it should be clear from the above discussion that our model will not start resembling the BM model even when the application cost becomes smaller. As S falls, the highest equilibrium wage gets closer and closer to the worker's productivity, but the number of rungs on the wage ladder is still bounded above. In contrast, the BM model generates either a degenerate wage distribution or a continuum of equilibrium wages, depending on the details of the application cost.

8. CONCLUSION

In this article, we have studied the equilibrium in a large labor market where employed workers search on the job and firms direct workers' search using wage offers and employment probabilities. All applicants observe all offers before the application. There is wage dispersion among workers, despite the assumption that all workers and all jobs are homogeneous. Moreover, equilibrium wages form a ladder. Because workers optimally choose to climb the ladder one rung at a time, wage mobility is limited endogenously. Also, wage gains diminish when a worker climbs up the ladder, because the gap between two adjacent rungs diminishes with wage. Furthermore, the density of the offer wage distribution is strictly decreasing, and the density of employed wages can be either decreasing or nonmonotonic. The equilibrium generates these new features without compromising on other familiar features, such as a worker's quit rate decreasing with wage and a worker's wage increasing with the employment duration.

The wage ladder arises here without any exogenous factors that hinder wage increases. For example, there is no gradual revelation of workers' productivity or job quality, no learning-by-doing and no match-specific productivity, as all workers have the same productivity on all jobs that is observable before match. There is no differential information regarding job openings, as all applicants observe all job openings before they apply. Also, firms do not discriminate the applicants, as they select all received applicants with equal probability. Rather, a worker chooses to climb the ladder gradually because his current wage affects his trade-off between the employment probability and wage. The jobs one level about the worker's current wage offer a significantly higher employment probability than other jobs at higher wages and, as such, they provide the best trade-off between the employment probability and wage.

We view the omission of the above factors as a strength, rather than a weakness, of the model. With purely homogeneous workers and jobs, the ability of the model to generate a wage ladder demonstrates the robustness of wage dispersion and the importance of search frictions for such wage dispersion. The results can be useful for explaining within-group wage inequality and wage mobility observed in the data, as we alluded to in the introduction.

The structure of the model permits many extensions and modifications. We are currently pursuing two. The first is to allow unemployment benefits to depend on workers' wages prior to unemployment, say, through the so-called replacement ratio. Because unemployed workers at each possible level of unemployment benefits form a source of a wage ladder, there will be many wage ladders in equilibrium. This will generate many more (and perhaps more dispersed) equilibrium wage levels than in the current article, and hence will help the model match the data better. The extension will also allow us to examine the equilibrium effects of changing the replacement ratio. The second line of research is to consider the efficiency properties of this model of directed search on the job, an issue considered in Acemoglu and Shimer (1999a,b), Moen (1997), Moen and Rosen (2004), and Shi (2001, 2002b) in various setups.

APPENDIX

A. *Proofs of Lemmas 1, 2, and 5.* We prove Lemma 1 first. 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 with w^*, $q(w)$ with $q(w^*)$, etc. Shorten the notation $q(w^*)$ to q^*. Suppose that (w^*, q^*) is an equilibrium offer. As concluded earlier, the constraint (6) must hold as equality for all types of applicants whom the firm attracts. Moreover, the constraint (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 nonbinding constraint.

Suppose that the offer (w^*, q^*) attracts both applicants at w_i and w_j, with $w_i < w_j$, and that the applicants' constraint binds for $w' = w_i$. 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. Since $J_e(\cdot)$ is a strictly increasing function, $\hat{q} < q^*$. Thus,

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

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 offer (w^*, q^*) induces the applicant's constraint to bind for $w' = w_j$. 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. This completes the proof of Lemma 1.

For Lemma 2, suppose that the w_{-1}-applicants prefer to apply to w relative to w_{+1}, that is,

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

For all $t \geq 2$, we have

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

The difference in $\{\cdot\}$ is nonnegative 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}.

Now turn to Lemma 5. For part (i) in the lemma, 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 (16). 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:

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

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

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

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:

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

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. ■

B. *Proofs of Propositions 1 and 3.* Before proving the propositions, we provide the following lemma (the proof is omitted; see Delacroix and Shi, 2002, Appendix A):

LEMMA B.1. *Define* $f(\cdot)$ *as in* (20) *and* $g(\cdot)$ *as follows:*

$$g(a) \equiv (r + \sigma + \lambda_1 q)\left(\frac{1}{f(a)} - \frac{1}{h(a)}\right) \tag{B.1}$$

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*

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

Furthermore, if $(r + \sigma)/\lambda_1 > f(a)/a$, *then* $g'(a) < 0$ *for all* $a > 0$.

We prove Proposition 1 by induction. The argument preceding the proposition in the text has already established (25) and (26) for $j = 0$. To verify (27) and (28) for $j = 0$, set $i = M - 1$ (and $w^d = w_{M-1}$) in (16) to obtain an equation for $J_e(w_{M-1})$. Using this equation and substituting $J_e(w_M)$, we get

$$J_e(w_M) - J_e(w_{M-1}) = \frac{w_M - w_{M-1} + \lambda_1 S}{r + \sigma + \lambda_1 q_M}$$

Combining this equation with (18) for $i = M - 1$, we obtain (27) and (28) for $j = 0$.

Now, suppose that (25)–(28) 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 all j. For $j + 1$, Equation (25) comes from setting $i = M - (j + 1)$ in (17), and (26) from the definitions of $h_{M-(j+1)}$ and $q_{M-(j+1)}$. To verify (27) and (28) for $j + 1$, set $i = M - j - 2$ (and $w^d = w_{M-j-1}$) in (16) to obtain an equation for $J_e(w_{M-j-2})$. Substituting this result, we get

$$\begin{aligned}
&J_e(w_{M-(j+1)}) - J_e(w_{M-(j+2)}) \\
&\quad = \frac{1}{r + \sigma + \lambda_1 q_{M-(j+1)}} \big[(r + \sigma) J_e(w_{M-(j+1)}) - w_{M-(j+2)} - \sigma J_u + \lambda_1 S\big] \\
&\quad = \frac{1}{r + \sigma + \lambda_1 q_{M-(j+1)}} \left[w_M - w_{M-(j+2)} + \lambda_1 S - (r + \sigma) C \sum_{t=0}^{j} \frac{1}{f_{M-t}} \right]
\end{aligned}$$

The second equality comes from substituting (28) for j, which holds by supposition. Combining the above result with (18) for $i = M - (j + 1)$, we obtain (27) and (28) for $j + 1$.

Finally, the zero-profit condition (17) must hold for a firm posting w_1. By the above derivation, this implies that (25) and (26) must hold for $j = M - 1$. By contrast, (27) and (28) need to be modified for $j = M - 1$. By definition, $w_0 = b$ and $J_e(w_0) = J_u$. To derive (29), use the wage ladder to simplify (10) 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 (18) (with $i = 1$), we obtain (29). This completes the proof of Proposition 1.

Now, we prove Proposition 3. First, for any fixed h^* that satisfies (24), we construct $M(h^*)$ in the proposition. Start with an arbitrary but sufficiently large integer m and set $h_m = h^*$. Compute the sequence $(a_{m-t})_{t \geq 0}$ according to Proposition 1 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 (36). By Proposition 2, $a_{m-t-1} < a_{m-t} \leq \bar{a} < \infty$ 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 since $a_i < a_{i+1}$ (see Proposition 2), 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.

Suppose that $\delta_m(m) \geq 0$. Then there exists i^* such that $\delta_i(m) \leq 0$ for all $i \leq i^*$ and $\delta_i(m) > 0$ for all $i \geq i^* + 1$. Let $M(h^*) = m - (i^* - 1)$ and compute the sequence $\{\delta_i(M(h^*))\}$ by setting $h_{M(h^*)} = h^*$. Then, $\delta_i(M(h^*)) \leq 0$ for all $i \leq 1$ and $\delta_i(M(h^*)) > 0$ for all $i \geq 2$. Moreover, for any integer $M' \neq M(h^*)$, the sequence $\{\delta_i(M')\}$ computed by setting $h_{M'} = h^*$ satisfies $\delta_1(M') = \delta_{(M(h^*)-M'+1)}(M(h^*))$. From the properties of the sequence $\{\delta_i(M(h^*))\}$, we have $\delta_1(M') > 0$ for all $M' \leq M(h^*) - 1$ and $\delta_1(M') < 0$ for all $M' \geq M(h^*) + 1$. This is the property described for $M(h^*)$ in Proposition 3.

The condition $\delta_m(m) \geq 0$ is guaranteed by Assumption 1. To see this, substitute $w_m = y - C(r + \sigma)/h_m$ to rewrite the condition $\delta_m(m) \geq 0$ as

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

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 (24), a sufficient condition for the above inequality is that it holds for this lower bound of a_m, which is imposed as (30).

Next, we find the equilibrium values of h_M and M. To compute the lowest equilibrium value of h_M, choose $h^* = 1 - (1 + \bar{a})e^{-\bar{a}}$ (the lower end of the interval given by (24)) and use the above procedure to obtain the corresponding $M(h^*)$. Then, $\Delta(M(h^*), h^*) \geq 0$. If $\Delta(M(h^*), h^*) = 0$, then h^* is the lowest equilibrium value of h_M and $M(h^*)$ is the equilibrium number of rungs on the ladder. Suppose $\Delta(M(h^*), h^*) > 0$. Then $\Delta(M(h^*) + 1, h^*) < 0$. Set $M^{**} = M(h^*) + 1$. By Proposition 2, the a sequence is an increasing function of h_M. So is w_M. Thus, $\Delta(M^{**}, h)$ is an increasing function of h. 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 (24). The necessary and sufficient condition for such crossing to exist is (37). The first crossing gives the lowest equilibrium value of h_M, where $M^{**} = M(h^*) + 1$ is the equilibrium number of rungs on the ladder.

We can also compute 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^* and compute the corresponding $M(h^*)$. If $\Delta(M(h^*), h^*) = 0$, then h^* is the highest equilibrium value of h_M. If $\Delta(M(h^*), h^*) > 0$, then fix $M = M(h^*)$. 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 (24). The first crossing gives the highest equilibrium value of h_M, where $M = M(h^*)$ is the equilibrium number of rungs on the ladder.

Similarly, we can compute all equilibrium values of h_M, which must lie in $[h_L, h_H]$. ■

C. *Proof of Proposition 2.* Before proving the proposition, we derive some useful equations from Proposition 1. Subtracting (27) for j and $j + 1$, we get

$$w - w_{-1} = \frac{(r + \sigma + \lambda_1 q)C}{f} - \frac{\lambda_1 C q_{+1}}{f_{+1}} \tag{C.1}$$

The suppressed index is $i = M - j$, and the above equation holds for all $j \in \{1, 2, \ldots, M-2\}$. For $j = 0$, the equation also holds once the last term is replaced with $\lambda_1 Cq(\bar{a})/f(\bar{a}) = \lambda_1 S$. Moreover, for all $j \in \{1, \ldots, M-2\}$, we have

$$\begin{aligned}\frac{r+\sigma+\lambda_1 q}{h_{-1}} &= \frac{1}{C}(y - w_{-1}) = \frac{1}{C}(y-w) + \frac{1}{C}(w - w_{-1}) \\ &= \frac{r+\sigma+\lambda_1 q_{+1}}{h} + \frac{1}{C}(w-w_{-1}) \\ &= \frac{r+\sigma+\lambda_1 q_{+1}}{h} + \frac{r+\sigma+\lambda_1 q}{f} - \frac{\lambda_1 q_{+1}}{f_{+1}}\end{aligned}$$

The first equality comes from using (25) for $j+1$, the second equality from rewriting, the third equality from using (25) for j, and the last equality from substituting (C.1). Thus, the following equation holds for all $j \in \{1, \ldots, M-2\}$

$$h_{-1} = (r+\sigma+\lambda_1 q) \Big/ \left(\frac{r+\sigma+\lambda_1 q_{+1}}{h} + \frac{r+\sigma+\lambda_1 q}{f} - \frac{\lambda_1 q_{+1}}{f_{+1}} \right) \tag{C.2}$$

For $j = 0$, replace the first term involving q_{+1} with 0 and the term $\lambda_1 q_{+1}/f_{+1}$ with $\lambda_1 q(\bar{a})/f(\bar{a}) = \lambda_1 S/C$.

We now prove Proposition 2. First, we verify (32) and (33) by induction. Since $h(a)$ is an increasing function and $q(a)$ a decreasing function, all three inequalities in (32) are equivalent to each other. To begin, we show that (32) and (33) hold for $j = 0$. By (24), $a_M \le \bar{a}$. With (C.2) for $j = 0$, we rewrite the condition $h_{M-1} < h_M$ as follows:

$$\begin{aligned}0 < \frac{r+\sigma+\lambda_1 q_M}{f_M} - \frac{\lambda_1 S}{C} - \frac{\lambda_1 q_M}{h_M} &= \lambda_1 \left(\frac{1}{e^{a_M} - 1 - a_M} - \frac{1}{e^{\bar{a}} - 1 - \bar{a}} \right) \\ &\quad + \left(\frac{r+\sigma}{f_M} - \frac{\lambda_1}{a_M} \right)\end{aligned}$$

where we have used the definition of $\bar{a}$ in (21) to replace S/C. Because $a_M \le \bar{a}$ by construction (see (24)) and $(e^a - 1 - a)$ is an increasing function, the term in the first $(\cdot)$ above is positive. Also, since $f(a)/a$ is an increasing function (see Lemma B.1), Assumption 1 implies $(r+\sigma)/\lambda_1 > f(\bar{a})/\bar{a} \ge f_M/a_M$. That is, the term in the second $(\cdot)$ above is also positive. Thus, $h_{M-1} < h_M$, verifying (32) for $j = 0$.

Now that $a_{M-1} < a_M \le \bar{a}$, and that $f(a)/a$ is an increasing function of a, (31) implies $(r+\sigma)/\lambda_1 > f(a_{M-1})/a_{M-1}$. That is, (33) holds for $j = 0$.

Suppose that (32) and (33) hold for an arbitrary $j \in \{1, 2, \ldots, M-3\}$. We show that they hold for $j+1$. For (32), this amounts to proving $h_{-2} < h_{-1}$. Computing h_{-2} using (C.2), $h_{-2} < h_{-1}$ if and only if

$$0 < \frac{r+\sigma+\lambda_1 q}{h_{-1}} - \frac{\lambda_1 q}{f} + (r+\sigma+\lambda_1 q_{-1}) \left(\frac{1}{f_{-1}} - \frac{1}{h_{-1}} \right)$$

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

$$0 < \frac{r+\sigma+\lambda_1 q}{h} - \frac{\lambda_1 q}{f} + (r+\sigma+\lambda_1 q_{-1})\left(\frac{1}{f_{-1}} - \frac{1}{h_{-1}}\right)$$

The last term is equal to $g(a_{-1})$. Note that $g'(a) < 0$ if $(r+\sigma)/\lambda_1 > f(a)/a$ (see Lemma B.1). Because $(r+\sigma)/\lambda_1 > f(\bar{a})/\bar{a}$, we have $g'(a) < 0$ for all $a \le \bar{a}$. Since $a_{-1} < a$ by supposition and $a \le \bar{a}$, then $g(a_{-1}) > g(a)$. Thus,

$$\begin{aligned}&\frac{r+\sigma+\lambda_1 q}{h} - \frac{\lambda_1 q}{f} + g(a_{-1}) > \frac{r+\sigma+\lambda_1 q}{h} - \frac{\lambda_1 q}{f} + (r+\sigma+\lambda_1 q)\left(\frac{1}{f} - \frac{1}{h}\right)\\ &\quad = (r+\sigma)/f(a) > 0\end{aligned}$$

That is, (32) holds for $j+1$. This in turn implies $a_{-2} < a_{-1}$. Because $f(a)/a$ is an increasing function of a (see Lemma B.1), the supposition $(r+\sigma)/\lambda_1 > f(a_{-1})/a_{-1}$ implies $(r+\sigma)/\lambda_1 > f(a_{-2})/a_{-2}$. That is, (33) also holds for $j+1$. By induction, (32) and (33) hold for all $j \in \{0, 1, \ldots, M-2\}$.

Second, we prove (34), which is equivalent to $h_{-1} > h(a - \ln(1+a))$. By (C.2), this in turn is equivalent to

$$\begin{aligned}0 &> \frac{r+\sigma+\lambda_1 q_{+1}}{h} - \frac{\lambda_1 q_{+1}}{f_{+1}} + (r+\sigma+\lambda_1 q)\left(\frac{1}{f} - \frac{1}{h(a-\ln(1+a))}\right)\\ &= \frac{r+\sigma+\lambda_1 q_{+1}}{h} - \frac{\lambda_1 q_{+1}}{f_{+1}} - \frac{r+\sigma+\lambda_1 q}{h} = -\lambda_1\left(\frac{q-q_{+1}}{h} + \frac{q_{+1}}{f_{+1}}\right)\end{aligned}$$

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 (34).

Finally, we show that $da/dh_M > 0$ and $dw/dh_M > 0$ for any h_M that satisfies (24). From (27) 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 $da_{+t}/dh_M \ge 0$ for all $t \ge 1$ implies $da/dh_M > 0$, then by induction, $da/dh_M > 0$. Suppose that $da_{+t}/dh_M \ge 0$ for all $t \ge 1$. By construction, $h = \lambda_1 C(\frac{r+\sigma}{\lambda_1} + q)/(y-w)$. Totally differentiating this relationship with respect to h_M (where dw/dh_M can be calculated using (27)), we have

$$\begin{aligned}\frac{y-w}{\lambda_1 C h} h' \frac{da}{dh_M} &= \frac{1}{\lambda_1 C}\frac{dw_M}{dh_M} + \frac{r+\sigma}{\lambda_1}\sum_{t=2}^{j}\frac{f'_{+t}}{f^2_{+t}}\left(\frac{da_{+t}}{dh_M}\right)\\ &\quad + \left[q'_{+1}\left(\frac{1}{h} - \frac{1}{f_{+1}}\right) + \left(\frac{r+\sigma}{\lambda_1} + q_{+1}\right)\frac{f'_{+1}}{f^2_{+1}}\right]\left(\frac{da_{+1}}{dh_M}\right)\end{aligned}$$

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

$$q'(a)\left(\frac{1}{h_{-1}}-\frac{1}{f}\right)+\left(\frac{r+\sigma}{\lambda_1}+q\right)\frac{f'(a)}{f^2}>0$$

To verify this inequality, temporarily denote the left-hand side of the inequality by LHS. Because $a_{-1} > a - \ln(1+a)$, $q' < 0$, $(r+\sigma)/\lambda_1 > f/a$, and $f' > 0$, we have

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

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

D. *Proof of Lemma 3.* To show that $\phi\ (a_{+1})$ is well defined for each a_{+1} by the equality form of (40), we use the definition of f to rewrite the equality as

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

The left-hand side of (D.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 (D.1) is 0, which is less than the right-hand side. When $a \to 0$, the left hand 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 (D.1) is greater than 1 (the right-hand side) at $a = a_{+1} - \ln(1+a_{+1})$. Substituting this particular value of a and rearranging 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:

$$\text{(D.2)}\qquad \frac{d}{da}\left[\frac{1}{f(a)}-\frac{1}{h(\phi(a))}\right]\le 0$$

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

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 (C.2), this desired condition is equivalent to

$$\frac{r+\sigma+\lambda_1 q_{+1}}{h} - \frac{\lambda_1 q_{+1}}{f_{+1}} + (r+\sigma+\lambda_1 q)\left(\frac{1}{f} - \frac{1}{h(\phi(a))}\right) > 0$$

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

$$\frac{r+\sigma+\lambda_1 q_{+1}}{h(\phi(a_{+1}))} - \frac{\lambda_1 q_{+1}}{f_{+1}} + (r+\sigma+\lambda_1 q)\left(\frac{1}{f} - \frac{1}{h(\phi(a))}\right) > 0$$

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

$$\begin{aligned}
&\frac{d}{da}\left\{(r+\sigma+\lambda_1 q)\left[\frac{1}{f} - \frac{1}{h(\phi(a))}\right]\right\} \\
&= (r+\sigma+\lambda_1 q)\frac{d}{da}\left[\frac{1}{f} - \frac{1}{h(\phi(a))}\right] + \lambda_1 q'(a)\left[\frac{1}{f} - \frac{1}{h(\phi(a))}\right] \\
&= \lambda_1\left(\frac{f}{a} + q\right)\frac{d}{da}\left[\frac{1}{f} - \frac{1}{h(\phi(a))}\right] + \lambda_1 q'(a)\left[\frac{1}{f} - \frac{1}{h(\phi(a))}\right] \leq 0
\end{aligned}$$

The first inequality comes from (D.2) and the result $(r+\sigma)/\lambda_1 > f(a)/a$ in Proposition 2, and the second inequality from (D.3). Because $a < a_{+1}$, the above result implies

$$\begin{aligned}
&\left[\frac{r+\sigma+\lambda_1 q_{+1}}{h(\phi(a_{+1}))} - \frac{\lambda_1 q_{+1}}{f_{+1}}\right] + (r+\sigma+\lambda_1 q)\left[\frac{1}{f} - \frac{1}{h(\phi(a))}\right] \\
&\geq \left[\frac{r+\sigma+\lambda_1 q_{+1}}{h(\phi(a_{+1}))} - \frac{\lambda_1 q_{+1}}{f_{+1}}\right] + (r+\sigma+\lambda_1 q_{+1})\left[\frac{1}{f_{+1}} - \frac{1}{h(\phi(a_{+1}))}\right] = \frac{r+\sigma}{f_{+1}} > 0
\end{aligned}$$

This is the desired result. ■

E. *Proof of Lemma 6.* 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 (44) and J_e^d by (45). 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.

Condition (IIa) requires $q^d[J_e^d(w^d) - J_e(w_{-2})] = E(w_{-2}) = Cq_{-1}/f_{-1}$. Using (18) and substituting $(J_e^d\,(w^d), J_e(w))$, we have

$$\begin{aligned}
&J_e^d(w^d) - J_e(w_{-2}) \\
&\quad = [J_e(w) - J_e(w_{-1})] + [J_e(w_{-1}) - J_e(w_{-2})] + \left[J_e^d(w^d) - J_e(w)\right] \\
&\quad = \frac{C}{f} + \frac{C}{f_{-1}} - \frac{(w - w^d)}{r + \sigma + \lambda_1 q_{+1}}
\end{aligned}$$

Thus, (IIa) requires:

$$\text{(E.1)} \qquad w^d = w - C(r + \sigma + \lambda_1 q_{+1}) \left(\frac{1}{f} + \frac{1}{f_{-1}} - \frac{q_{-1}}{q^d f_{-1}}\right)$$

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 (18) to substitute $J_e(w_{+1})$, noting that $J_e(w) - J_e^d\,(w^d) = (w - w^d)/(r + \sigma + \lambda_1\, q_{+1})$, and substituting w^d from (E.1), we can rewrite (IIc) as follows:

$$\text{(E.2)} \qquad q^d \leq q_{-1} \Bigg/ \left[1 + f_{-1}\left(\frac{1}{f} - \frac{q_{+1}}{(q - q_{+1})\, f_{+1}}\right)\right]$$

Let β be the level of a^d that satisfies (E.2) as equality. Since $q^d = q(a^d)$ is a decreasing function of a^d, (E.2) is equivalent to $a^d \geq \beta$.

A type II deviating firm chooses (w^d, a^d) to maximize $\pi(a^d) = h^d J_f^d(w^d)$, subject to (E.1) and (E.2). Substituting w^d from (E.1), $J_f^d(w^d)$ from the text, and $(y - w)$ from (17), we write the deviator's expected surplus as follows:

$$\pi(a^d) = \frac{C[(y - w) + (w - w^d)]}{r + \sigma + \lambda_1 q_{+1}} = C\left[h^d\left(\frac{1}{h} + \frac{1}{f} + \frac{1}{f_{-1}}\right) - \frac{a^d q_{-1}}{f_{-1}}\right]$$

If the constraint (E.2) does not bind, then $\pi(a^d)$ is maximized at $a^d = A$ that solves $\pi'(A) = 0$, that is,

$$\text{(E.3)} \qquad A = \ln\left[\frac{1}{q_{-1}}\left(1 + f_{-1}\left(\frac{1}{h} + \frac{1}{f}\right)\right)\right]$$

Because $a_{-1} > a - \ln(1 + a)$ by Proposition 2, it can be shown that $A > a_{-1}$. The maximum of $\pi(a^d)$ without the constraint (E.2) is

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

Thus, for the deviation to be not profitable, the constraint (E.2) must bind to keep a^d a sufficient distance away from A. We make this requirement more explicit below.

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 (47) in the lemma. Let us denote $Y = q_{+1}/[(q - q_{+1}) f_{+1}]$ in this section of the 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}. Rearranging terms and using the definition of β^* in (46), 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 2 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 rearranging 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

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

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 (47). To do so, rewrite (C.2) as

$$\frac{1}{f} - Y = \frac{1}{h_{-1}} - \frac{r+\sigma+\lambda_1 q_{+1}}{r+\sigma+\lambda_1 q}\left(\frac{1}{h} + Y\right)$$

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

$$\begin{aligned} 0 &\leq \frac{1}{q(\beta^*)} - \frac{1}{q_{-1}}\left\{1 + f_{-1}\left[\frac{1}{h_{-1}} - \frac{r+\sigma+\lambda_1 q_{+1}}{r+\sigma+\lambda_1 q}\left(\frac{1}{h} + Y\right)\right]\right\} \\ &= \frac{1}{q(\beta^*)} - \frac{1}{q_{-1}}\left(1 + \frac{f_{-1}}{h_{-1}}\right) + \left(\frac{r+\sigma+\lambda_1 q_{+1}}{r+\sigma+\lambda_1 q}\right)\frac{f_{-1}}{q_{-1}h(\beta^*)} \\ &= \frac{1}{q(\beta^*)} - e^{a_{-1}} + \left(\frac{r+\sigma+\lambda_1 q_{+1}}{r+\sigma+\lambda_1 q}\right)\left(e^{a_{-1}} - 1 - a_{-1}\right)/h(\beta^*) \end{aligned}$$

The inequality comes from substituting the definition of β and the term $(\frac{1}{f} - Y)$; the two equalities come from substituting the definitions of $h(\beta^*)$ and f. Multiplying the above inequality by $h(\beta^*)$ yields (47). ■

F. *Proofs of Propositions 5 and 6.* We prove Proposition 5 first. Property (i) holds because $E(w_{-1}) = Cq/f > Cq_{+1}/f_{+1} = E(w)$ and $E(w_{M-1}) \geq S$ (see (23) or equivalently the first part of (24)). To establish (ii), use (C.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 $R = (r+\sigma)/\lambda_1$. For the computed sequence to be an equilibrium, we need $q > q_{+1}(1 + f/f_{+1})$ (see (40)). 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 expression is clearly positive, because $a < a_{+1}$ and $f(\cdot)$ is an increasing function. Thus, Proposition 5 holds.

To prove Proposition 6, recall that the density of offer wages is (v_i) and of employed wages $(n_i/(1-u))$, where $i = 1, 2, \ldots, M$. 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 (38) and (14), $n_{-1}/n = (\frac{\sigma}{\lambda_1} + q_{+1})/q$ and $v_{-1}/v = (n_{-2}\, a)/(n_{-1}\, a_{-1})$ for all $i \geq 3$. For all $i \geq 3$, we have

$$\frac{v_{-1}}{v} = \left(a\frac{\sigma}{\lambda_1} + h\right)\Big/ h_{-1} > \frac{h}{h_{-1}} > 1$$

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 $\sigma/\lambda_1 > q - q_{+1}$. 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 $\sigma/\lambda_1 > q_M$. Because $q(\cdot)$ is a decreasing function and $a_M \geq \bar{a} - \ln(1 + \bar{a})$ by (24), a sufficient condition for $n_{M-1} > n_M$ is $\sigma/\lambda_1 > q(\bar{a} - \ln(1 + \bar{a}))$, which can be rewritten as (48). When r is sufficiently close to 0, this condition is satisfied iff $(r + \sigma)/\lambda_1 > q(\bar{a} - \ln(1 + \bar{a}))$. Because $(r + \sigma)/\lambda_1 \geq f(\bar{a})/\bar{a}$ by Assumption 1, (48) 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 (24), a sufficient condition for $n_{M-1} < n_M$ is $\sigma/\lambda_1 < q(\bar{a})$. ■

G. *Markov Perfect Equilibrium.* In this section, we formulate the Markov perfect equilibrium, discuss the analytical difficulties of using this formulation, and then use a numerical example to show the discrepancy between the perfect equilibrium and the Nash equilibrium constructed in the text.

A Markov perfect equilibrium in the described environment has three requirements. First, given any distribution of workers, recruiting firms' optimal strategy is to use the function $q(\cdot)$ to describe employment probabilities for all possible wages. Second, given the distribution of job openings, the applicants' optimal strategy is a function $T(\cdot)$; that is, for every w, $T(w)$ describes the optimal target wages for an applicant currently employed at wage w. Third, the players' strategies in the period and the distribution of workers at the beginning of the period imply the distribution of workers at the beginning of next period.

The third requirement is easy to implement, as we focus on equilibria where the distribution of workers is stationary (see (38)). For the second requirement, we examine only those subgames where the distribution of job openings is consistent with the free-entry condition. This requirement determines the function $q(\cdot)$, as we will describe later. Thus, the main task of characterizing a Markov equilibrium is to characterize the function $T(\cdot)$.

To characterize T, take any decreasing probability function $q(\cdot)$ and formulate the applicants' problem. Note that the function $q(w)$ must be specified for all w, not only for equilibrium wages. For convenience, 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 (9) to write a worker's value function as $J_e(w) = [F(w) + \sigma J_u]/(r + \sigma)$. 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 nonempty. Otherwise, $T(w)=\emptyset$. With this notation, the case $E(w)=S$ means that a w-applicant does not apply.

Substituting the above formula of $E(w)$ into the definition of F, we obtain

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

This is a fixed-point problem for F, and the maximizer for the inner maximization gives $T(\cdot)$. Under reasonable conditions, we can show that the right-hand side of (G.1) is a contraction mapping, and so there is a unique function $F(\cdot)$ that satisfies the functional equation.

Next, to find the function $q(\cdot)$, we use the free-entry condition: $h(w)J_f(w)=C$. Substituting J_f from (8), $\rho\ (w)=\lambda\ (w)q(T(w))$ and $J_v=0$, we can write this condition as $q(w)=Pq(w)$, where P is the following mapping:

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

The function Ψ, defined in (5) is a decreasing function. Thus, $q(\cdot)$ is a fixed point of P.

It is difficult to examine the fixed-point problem for q, because the maximizer $T(w)$ to the fixed-point problem (G.1) appears in the mapping P. This difficulty exists even when we assume that $T(w)$ is singleton for every w and that $T(\cdot)$ is continuous. The source of the difficulty is that we need the given function $q(\cdot)$ in (G.1) to be decreasing to ensure well-behaved fixed point F (and to make economic sense). 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. Such a difficulty would not arise if there were no on-the-job search, because then $Pq(w)=\Psi(\frac{r+\sigma}{y-w}C)$ is clearly a decreasing function.

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

To see the discrepancy between the Markov equilibrium and the Nash equilibrium constructed in the text, consider the following parameter values: $r=0.02$, $y=1000$, $b=0$, $C=60$, $S=1$, $\lambda_1=\lambda_0=.025$, and $\sigma=.125$ (this is the example of Section 6.2, except for $\lambda_1=\lambda_0$). Discretize the interval between b and y and set the number of points on the grid to be 25,000. The equilibrium number of rungs on the wage ladder is $M=4$ in both equilibria. We list the results in Table G.1.

The two equilibria are very close to each other. The maximum discrepancy in equilibrium wages between the two types of equilibria is about 0.04%. This discrepancy is small, especially when we consider the following two factors. First,

TABLE G.1
NASH VERSUS MARKOV EQUILIBRIUM DISTRIBUTIONS

	w_1	w_2	w_3	w_4
Nash	929.005	980.004	988.339	990.679
Markov	929.043	980.401	988.640	990.760
Discrepancy	0.004%	0.040%	0.031%	0.008%

the objects to be discretized in the numerical procedures are different for the two equilibria. For the Nash equilibrium, we discretized the interval of the hiring probability h_M and, for the Markov equilibrium, we discretized 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.

REFERENCES

ACEMOGLU, D., AND R. SHIMER, "Holdups and Efficiency with Search Frictions," *International Economic Review* 40 (1999a), 827–50.
——, AND ——, "Efficient Unemployment Insurance," *Journal of Political Economy* 107 (1999b), 893–928.
BARLEVY, G., "Estimating Models of On the Job Search Using Record Statistics," NBER Working Paper No. w10146, 2003.
BUCHINSKY, M., AND J. HUNT, "Wage Mobility in the United States," *Review of Economics and Statistics* 81 (1999), 351–68.
BURDETT, K., AND M. G. COLES, "Equilibrium Wage-Tenure Contracts," *Econometrica* 71 (2003), 1377–1404.
——, AND K. JUDD, "Equilibrium Price Distributions," *Econometrica* 51 (1983), 955–70.
——, AND D. MORTENSEN, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review* 39 (1998), 257–73.
——, S. SHI, AND R. WRIGHT, "Matching and Pricing with Frictions," *Journal of Political Economy* 109 (2001), 1060–85.
CAO, M., AND S. SHI, "Coordination, Matching, and Wages," *Canadian Journal of Economics* 33 (2000), 1009–33.
COLES, M. G., "Equilibrium Wage Dispersion, Firm Size, and Growth," *Review of Economic Dynamics* 4 (2001), 149–87.
DELACROIX, A., AND S. SHI, "Directed Search On the Job and the Wage Ladder," Working Paper, available at: http://www.economics.utoronto.ca/ecipa/archive/UT-ECIPA-SHOUYONG-03-04.html, University of Toronto, 2002.
DIAMOND, P., "A Model of Price Adjustment," *Journal of Economic Theory* 3 (1971), 156–68.
——, "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy* 90 (1982), 881–94.
JOVANOVIC, B., "Job Matching and Turnover," *Journal of Political Economy* 87 (1979), 972–90.
JUHN, C., K. M. MURPHY, AND B. PIERCE, "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy* 101 (1993), 410–42.

JULIEN, B., J. KENNES, AND I. KING, "Bidding for Labor," *Review of Economic Dynamics* 3 (2000), 619–49.

KIEFER, N. M., AND G. R. NEUMANN, "Wage Dispersion with Homogeneity: The Empirical Equilibrium Search Model," in Bunzel et al., eds., *Panel Data and Labour Market Dynamics* (Amsterdam: North Holland, 1993).

MOEN, E., "Competitive Search Equilibrium," *Journal of Political Economy* 105 (1997), 385–411.

——, AND A. ROSEN, "Does Poaching Distort Training?" *Review of Economic Studies* 71 (2004), 1143–62.

MONTGOMERY, J. D., "Equilibrium Wage Dispersion and Interindustry Wage Differentials," *Quarterly Journal of Economics* 106 (1991), 163–79.

MORTENSEN, D., "Property Rights and Efficiency in Mating, Racing, and Related Games," *American Economic Review* 72 (1982), 968–79.

PETERS, M., "Ex Ante Price Offers in Matching Games: Non-Steady State," *Econometrica* 59 (1991), 1425–54.

——, "Limits of Exact Equilibria for Capacity Constrained Sellers with Costly Search," *Journal of Economic Theory* 95 (2000), 139–68.

PISSARIDES, C., *Equilibrium Unemployment Theory*, 1st ed. (Cambridge, MA: Basil Blackwell, 1990).

——, "Search Unemployment with On the Job Search," *Review of Economic Studies* 61 (1994), 457–75.

POSTEL-VINAY, F., AND J-M. ROBIN, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica* 70 (2002), 2295–2350.

——, AND ——, "To Match or Not to Match? Optimal Wage Policy with Endogenous Worker Search Intensity," *Review of Economic Dynamics* 7 (2004), 297–330.

SHI, S., "Frictional Assignment, Part I: Efficiency," *Journal of Economic Theory* 98 (2001), 232–60.

——, "Product Market and the Size-Wage Differential," *International Economic Review* 43 (2002a), 21–54.

——, "A Directed-Search Model of Inequality with Heterogeneous Skills and Skill-Biased Technology," *Review of Economic Studies* 69 (2002b), 467–91.

SHIMER, R., "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," Manuscript, Princeton University, 2001.

TOPEL, R. H., AND M. P. WARD, "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics* 107 (1992), 439–79.

VAN DEN BERG, G., AND G. RIDDER, "An Empirical Equilibrium Search Model of the Labor Market," *Econometrica* 66 (1998), 1183–1221.