Directed Search for Equilibrium Wage-Tenure Contracts

Shouyong Shi*
Department of Economics, University of Toronto
150 St. George Street, Toronto, Ontario, Canada, M5S 3G7
(email: shouyong@chass.utoronto.ca)

August 2007

Abstract

I analyze the equilibrium in a labor market where firms offer wage-tenure contracts to direct the search of employed and unemployed workers. Each applicant observes all offers and there is no coordination among individuals. Workers’ applications (as well as firms’ recruiting decisions) are optimal. This optimality requires the equilibrium to be formulated differently from the that in the literature of undirected search. I provide such a formulation and show that the equilibrium exists. In the equilibrium, individuals explicitly tradeoff between an offer and the matching rate at that offer. This tradeoff yields a unique offer which is optimal for each worker to apply, and applicants are separated endogenously according to their current values. Despite such uniqueness and separation, there is a non-degenerate and continuous wage distribution of employed workers in the stationary equilibrium. The density of this distribution is increasing at low wages and decreasing at high wages. In all equilibrium contracts, wages increase with tenure, which results in quit rates to decrease with tenure. Moreover, the model makes novel predictions about individuals’ job-to-job transition and comparative statics.

Keywords: Directed search, On-the-Job, Wage-tenure contracts.

* I have received useful comments from the participants of the workshops at University of Pennsylvania, Rochester, Yale, UBC, the conference at UQAM on “Microfoundations of Markets with Frictions” (2007), and the NBER Summer Institute (2007). Financial support from the Bank of Canada Fellowship and from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged. The opinion expressed in the paper is my own and it does not reflect the view of the Bank of Canada. All errors are mine.
1. Introduction

Directed search is a matching process in which an individual can use his offer to affect his matching rate. The objective of this paper is to study the equilibrium in a labor market where firms offer wage-tenure contracts to direct workers’ search. A wage-tenure contract is a time profile of wages which describes how a worker’s wage will evolve with tenure. Employed workers can search on the job. I characterize the equilibrium and establish its existence. The equilibrium yields novel predictions about job-to-job transitions, the wage distribution, and effects of unemployment policy.

To see why directed search is interesting to study, it is useful to contrast it with the large literature on undirected search developed from Diamond (1982), Mortensen (1982), and Pissarides (1990). There are two classes of models in this literature. In one class, as in the three pioneering works, prices (wages) are a result of bargaining after individuals are matched. In the other class, the individuals on one side of the market post prices but searchers do not know who posted what prices before they are matched (e.g., Burdett and Mortensen, 1998, and Burdett and Coles, 2003). In both classes of models, search is undirected because prices play no role, ex ante, to direct workers to particular matches. Although undirected search has been a fertile ground of research, there are strong reasons for studying directed search. First, some search is directed rather than completely random, particularly for workers who search on the job. For example, applicants often have information about wages from job advertisement, word of mouth, or referrals. Second, undirected search generates an array of market inefficiencies, whose corrective policy depends on the details assumed for the matching and price determination processes (see Hosios, 1990). Directed search can eliminate most of these inefficiencies. Third, wage dispersion in undirected search models is sensitive to the assumption that a worker knows at most one wage before search. For example, the continuous wage distribution in the well-known model of Burdett and Mortensen (1998) would become degenerate if each applicant knew two or more wages. Directed search models do not have this sensitivity because they allow each applicant to observe all offers before the application.

During the last fifteen years or so, a literature has grown to analyze directed search. Peters (1984, 1991) and Montgomery (1991) provide two of the earliest formulations. Examples of the further exploration include Moen (1997), Acemoglu and Shimer (1999a,b), Julien, et al. (2000), Burdett, et al. (2001), Shi (2001, 2002), Coles and Eeckhout (2003), Galenianos and Kircher (2005), and Delacroix and Shi (2006). They have shown that an equilibrium with directed search and its efficiency properties are significantly different from
those with undirected search.

However, the literature of directed search has its own shortcomings. First, it has ignored wage-tenure contracts and on-the-job search, by assuming that each firm posts a fixed wage for the entire duration of a worker’s employment and that a worker must quit his job first before searching for another job. Without wage-tenure contracts, this literature is unable to explain the empirical regularities that wages rise and quit rates fall with tenure (e.g., Farber, 1999). Without on-the-job search, the literature cannot predict job-to-job transitions which constitute a large part of the flow of workers in the data. Second, directed search models generate only a few number (and often a singleton) of equilibrium wages among homogeneous workers, in contrast to the data and to the undirected search model by Burdett and Mortensen (1998). Given the appealing features of directed search discussed above, there is an urgency to advance the theory to bridge these gaps with the data.\footnote{Delacroix and Shi (2006) and Julien et al. (2006) are the exceptions that introduce on-the-job search into directed search models, but they do not examine wage-tenure contracts.}

To appreciate the challenge in characterizing a directed search equilibrium with contracts, let me compare the task with the one in undirected search, which is accomplished by Burdett and Coles (2003, BC henceforth). With undirected search, workers are assumed to send their applications randomly to a pool of recruiting firms. With directed search, however, each worker’s application must be optimal in the tradeoff between an offer and the likelihood of obtaining the offer. Similarly, each firm understands that it can raise the offer to entice more workers to apply to it. To describe this tradeoff, an equilibrium must determine two new objects in addition to optimal contracts. One is the employment rate function, which describes how the rate at which an applicant gets an offer varies with the offer. The other is the hiring rate function, which describes how the rate at which a recruiting firm successfully hires a worker varies with the offer. A challenge in characterizing the equilibrium is to show that these functions exist.

I formulate the equilibrium in an environment where all matches have the same productivity, and then establish the existence of the equilibrium. In equilibrium, the hiring rate associated with an offer is indeed an increasing function of the offer, and the employment rate is a decreasing function of the offer. Thus, the tradeoff between an offer and the matching rate is meaningful.

On wage-tenure contracts, the equilibrium with directed search preserves several attractive properties of the undirected search model by BC. First, wages increase and quit rates fall with tenure. Second, all equilibrium contracts are sections of a baseline contract. In the baseline contract, wages start at the lowest equilibrium level and then increase with
tenure. Any other equilibrium contract that starts at a different wage is identical to the remaining section of the baseline contract from that wage level onward. Third, on-the-job search and wage-tenure contracts together generate a continuous wage distribution in the equilibrium, even though all matches have the same productivity. On-the-job search creates jumps in wages when a worker changes the job, while wage-tenure contracts provide smooth increases in wages when a worker stays with a job.

Beyond these similarities, the equilibrium here has little in common with undirected search. A striking difference is that directed search creates the dichotomy that individuals’ optimal decisions, equilibrium contracts and the matching rate functions are independent of the distribution of workers. In undirected search models, in contrast, the distribution of workers is critical for determining individuals’ optimal decisions. The dichotomy generates novel results of comparative statics. For example, although an increase in the unemployment benefit affects the distribution of workers, it has no effect on an employed worker’s job-to-job transition rate and his future wage path, given the worker’s current wage. Moreover, the dichotomy makes the model tractable for business cycle research with on-the-job search, as I will elaborate in the concluding section.

The second difference is in job-to-job transitions and wage mobility. With directed search, each worker optimally chooses to apply to a unique target which is an increasing function of the worker’s state, i.e., the value of the worker’s current contract or unemployment benefit. As workers separate themselves this way, wage mobility is endogenously limited by workers’ current states despite that there is no difference in productivity across matches. Such limited wage mobility seems realistic (see Buchinsky and Hunt, 1999). In contrast, undirected search models (e.g., BC, 2003, and Burdett and Mortensen, 1998) predict that a worker can receive an offer that lies anywhere in the support of the wage distribution, and that a worker can transit to any wage that lies above his current wage.

The third difference from undirected search is the wage distribution. As said earlier, wage dispersion generated by directed search is robust because all applicants are allowed to observe all offers before they apply. Moreover, the density function of the wage distribution is decreasing at high wages. This feature is an empirical regularity (see Kiefer and Neumann, 1993) which cannot be captured by an undirected search model with homogeneous matches. The equilibrium in the latter model necessarily predicts that the density function of equilibrium wages is increasing at high wages. To modify this prediction, the literature of undirected search has introduced heterogeneity across matches (e.g., van den Berg and Ridder, 1998). It is important to show that directed search can generate a decreasing wage density without such heterogeneity.
Now let me return briefly contrast the current paper with Delacroix and Shi (2006), who also examine directed search on the job. That paper excludes wage-tenure contracts by assumption. Incorporating wage-tenure contracts allows me not only to explain the patterns of wages and quits over tenure, but also to produce a rich wage distribution. In Delacroix and Shi, the equilibrium wage structure is a wage ladder, the discreteness of which makes the analysis quite messy. Wage-tenure contracts fill in the gap between any two rungs of the ladder, because wages increase with tenure. Moreover, the smoothness enables me to characterize the equilibrium more generally.

To emphasize the differences between directed and undirected search, I maintain four assumptions imposed by BC. First, workers are risk averse; second, the capital market is imperfect so that workers cannot borrow against their future income. These two assumptions are important for generating the wage-tenure relationship, as discussed by BC. Third, a firm does not respond to the employee’s outside offers. How reasonable this assumption is varies across different types of markets. In any case, the assumption is commonly imposed in the literature, and it enables me to compare the results clearly with those in BC. For a model of undirected search without this assumption, see Postel-Vinay and Robin (2002). Finally, I assume that the productivity of a match is public information and deterministic. For private information or learning about productivity, see Jovanovic (1979), Harris and Holmstrom (1982), and Moscarini (2005). Such productivity differences between matches or over time are clearly important for wage dynamics and turnover, but abstracting from them enables me to have a clear exploration of search frictions.

2. The Model

Consider a labor market that lasts forever in continuous time. I normalize the rate of time preference to zero. There is a unit measure of risk averse workers whose utility function is $u(w)$, where $w$ is income. Workers are not able to borrow against their future income, and so the lower bound on wages is 0. All workers have the same productivity. When employed, each worker produces a flow of output, $y > 0$. When unemployed, a worker enjoys a flow of utility, $u(b)$, which is derived from leisure and other benefits in unemployment. However, I will refer to $b$ simply as the unemployment benefit.\footnote{In many occupations, workers choose to quit for outside offers rather than ask their employers to match the offers. Matching outside offers is more common in other occupations such as economists and professional athletes. However, in these occupations, a main motivation for matching offers might be the competition for workers’ ability. Because the current model abstracts from all heterogeneity across matches, the assumption of not responding to outside offers may not be unreasonable.}
Assumption 1. The utility function has the following properties: $0 < u'(w) < \infty$ and $-\infty < u''(w) < 0$ for all $w \in (0, \infty)$; $u'(0) = \infty$; and $u(0) = -\infty$.

Risk aversion in the above assumption is not common in the labor search literature, but it is critical for the main results in this paper. The assumption $u(0) = -\infty$ is imposed for the following reason discussed extensively by BC. That is, if $u(0)$ is large, wages may be zero at the beginning of a contract. Although this possibility does not pose serious difficulties to the analysis, it is cumbersome to be included and, hence, excluded by the sufficient condition $u(0) = -\infty$.

All workers face death at a Poisson rate $\delta \in (0, \infty)$ and a dead worker is replaced with a newborn who enters the labor market through unemployment. Death is the only exogenous separation, and so employment is permanent until either the worker dies or quits for a better job. Burdett and Coles (2003) also model exogenous separation as death, rather than separation into unemployment. The modelling simplifies the analysis by eliminating savings: because an employed worker never returns to unemployment, the worker has no incentive to save provided that wages increase with tenure.\(^3\)

There are also a large number of identical and risk-neutral firms that can enter the market. Entry is competitive: a firm can post a vacancy at a flow cost $k > 0$. As common in the literature, a firm has a production technology with constant returns to scale and considers different jobs independently. Normalize the production cost to 0. Recruiting firms announce wage-tenure contracts to compete for workers. Firms are assumed to commit to the contracts, although a worker can quit his job at any time. This assumption has two requirements. First, as commonly assumed in the search literature, a firm cannot continue to recruit for a filled job in an attempt to replace the current employee or to renegotiate the contract with him. Second, as discussed in the introduction, a firm cannot respond to the employee’s outside offers.

A contract specifies a path of wages as a function of the worker’s tenure, conditional on the worker’s employment in the firm. However, because only the lifetime utility generated by a contract matters to a worker, I express a contract alternatively as a path of such utilities, or values. To do so, consider a contract offered at time $s$. After a worker has worked under the contract for a length of tenure $t$, let $V(t, s)$ denote the remaining value of $V(t, s)$.

\(^3\)If there is exogenous separation into unemployment, one way to eliminate savings is to assume that workers receive severance pays which are sufficiently increasing in wages. Because only the workers with sufficiently high wages will have incentive to save against the event of separation, the severance pay acts as a substitute for such savings. In this case, the analysis can be adapted accordingly. I thank Christopher Pissarides for making this suggestion.
the contract to the worker. This value is the discounted sum of the worker’s expected utility generated by the remaining wage path in the contract from time \((s + t)\) onward, conditional on the worker’s optimal quitting strategy in the future. I refer to the value \(V(0, s)\) as an offer at time \(s\) and the path of values, \(\{V(t, s)\}_{t=0}^{\infty}\), as the contract that delivers the offer. Let \(w(V(t, s))\) denote the wage level at tenure \(t\) according to this contract. I call the function \(w(V)\) the wage function.

Throughout this paper, \(t\) denotes a worker’s tenure rather than the calendar time. To unify the notation, I denote an unemployed worker’s “tenure” as \(t = \emptyset\). The value for an unemployed worker is denoted \(V_u = V(\emptyset, s)\) and the unemployment benefit is expressed alternatively as \(w(V(\emptyset, s)) = b\), for all \(s\).

All offers are bounded in \([\underline{V}, \bar{V}]\), where
\[
\bar{V} = u(\bar{w})/\delta, \quad \underline{V} = u(b)/\delta.
\]
\(\bar{w}\) is the highest wage which will be determined in Lemma 3.2 and \(\bar{V}\) is the lifetime utility of a worker who is employed at the highest wage permanently until death. The lower bound \(\underline{V}\) is the lifetime utility of an unemployed worker who is deprived of the opportunity of applying to jobs in the entire lifetime. Because an unemployed worker does have the opportunity to apply for jobs, all equilibrium offers will be strictly higher than \(\underline{V}\).

Both unemployed and employed workers can search for jobs. At any instant, a fraction \(\lambda_0\) of unemployed workers and a fraction \(\lambda_1\) of employed workers are randomly selected to receive job application opportunities. I allow for the possibility \(\lambda_0 = \lambda_1 = 1\) by letting \(\lambda_0\), \(\lambda_1 \in (0, 1]\).\(^4\) A worker who receives the application opportunity observes all firms’ offers instantly without any cost and then chooses the offer to which he applies. As in most search models, each worker can apply to only one offer.\(^5\)

Individuals cannot coordinate their actions. If two or more workers apply to the same offer, the firm randomly selects only one to employ. Thus, the coordination failure generates unemployment. Moreover, if an employed worker gets an offer, the worker must quit his current job before accepting the offer. As discussed in the introduction, the worker’s

\(^4\)Note that the \(\lambda\)’es are not Poisson rates but, rather, the fractions of workers who receive job application opportunities at any instant. As such, they are bounded above by one.

\(^5\)Let me clarify two assumptions here. One is that an applicant observes all offers. This assumption is not necessary, because the essential results are the same if each applicant is assumed to observe two offers that are randomly drawn from the offer distribution (see Acemoglu and Shimer, 1999b). The second assumption is that each applicant can apply to only one offer at a time (for multiple applications, see Galenianos and Kircher, 2005). In continuous time, this assumption is not as restrictive as it may sound. Although a worker in reality may be able to send out multiple applications, the probability with which two or more of his applications will be received by different firms at the same instant is negligible.
incumbent firm is assumed not to respond to the worker’s outside offers. A job is destroyed when the worker either accepts another firm’s offer or dies.

Because workers observe the offers before they apply, the offers can direct workers’ search. Workers and firms both make the tradeoff between an offer and the matching rate at that offer. When choosing a value to offer, a firm faces a hiring rate function, \( q(.) \). That is, a firm knows that it can change the offer to affect its hiring rate directly according to \( q(.) \). Similarly, an applicant understands that different offers are associated with different employment rates according to an employment rate function, \( p(.) \). Because \( p \) and \( q \) are Poisson rates instead of probabilities, they can exceed one.

The functions \( q(.) \) and \( p(.) \) are equilibrium objects, since they must satisfy two equilibrium requirements. First, they must be consistent with aggregation. That is, as firms and workers make their choices under these functions, the resulting matching rates must indeed be given by these functions. Second, the hiring rate function must ensure that the expected profit of recruiting be the same for all equilibrium offers. Delaying the second requirement to section 4, I specify the first requirement below.

Let me start with a matching function, \( M(\theta, 1) \), which specifies the measure of matches between a measure \( \theta \) of workers and a unit measure of firms. Refer to \( \theta \) as the tightness. Assume that \( M \) is linearly homogeneous. Given the two functions \( p(.) \) and \( q(.) \), individuals’ decisions induce the tightness, \( \theta(V) \), at each value \( V \). Aggregate consistency requires that the matching rates satisfy: \( q(V) = M(\theta(V), 1) \) and \( p(V) = M(\theta(V), 1)/\theta(V) \). Using these relationships to eliminate \( \theta \), I can express aggregate consistency as \( p(V) = M(q(V)) \).

The function \( M(q) \) embodies all essential properties of the matching function. From now on, I will take \( M(q) \) as a primitive of the model and refer to it as the matching function.\(^6\) To specify the properties of the matching function, let \( q(V) \in [q, \bar{q}] \) for all \( V \), with \( 0 < q < \bar{q} \), where \( \bar{q} \) is an exogenous upper bound on \( q \) given by the matching function and \( q \) will be restricted by (5.4) later.

**Assumption 2.** The matching function \( M(q) \) satisfies: (i) \( M(q) \) is continuous for all \( q \in [q, \bar{q}] \) and, for all \( q \) in the interior of \( (q, \bar{q}) \), the derivatives \( M'(q) \) and \( M''(q) \) exist and are finite; (ii) \( \bar{q} < \infty \) and \( M(\bar{q}) = 0 \); (iii) \( M'(q) < 0 \); (iv) \( -qM''(q)/M'(q) \leq 2 \).

Part (i) is a regularity condition that is satisfied by many well-known matching functions. Part (ii) is imposed for the convenience of working with bounded functions. Part

---

\(^6\)I follow Moen (1997) and Acemoglu and Shimer (1999a) to take the matching function as given. In contrast, some directed search models have derived the matching function by aggregating agents’ strategies, e.g., Peters (1991), Burdett et al. (2001), Julien et al. (2000) and Delacroix and Shi (2006). The main results of the current paper continue to hold when the matching function is endogenized so.
(iii) is equivalent to $0 < \theta M_1 / M < 1$, which is satisfied by all matching functions of constant returns to scale that are strictly increasing in the arguments. I will use these assumptions to show that the equilibrium generates the desirable properties, $q'(V) > 0$ and $p'(V) < 0$. Part (iv) restricts convexity of $M(q)$, which will be useful for ensuring uniqueness of a worker’s optimal application decision.\footnote{For a general matching function, part (iv) of the assumption requires $1 - \theta M_1 / M \leq [(-\theta M_{11} / (2M_1)]^{1/2}$, where the left-hand side is the share of vacancies in the matching function.}

The following example gives two common matching functions and finds the restrictions for them to satisfy Assumption 2:

**Example 2.1.** One matching function is the so-called urn-ball matching function. Derived endogenously by Peters (1991) and Burdett et al. (2001), the function has the form: $M(\theta, 1) = \bar{q} (1 - e^{-\theta})$ with $\bar{q} < \infty$, which implies:

$$M(q) = -\frac{q}{\ln (1 - q/\bar{q})}.$$  

This function satisfies Assumption 2. Another matching function is the CES function: $M(\theta, 1) = [\alpha \theta^\rho + 1 - \alpha]^{1/\rho}$, where $\alpha \in (0, 1)$ and $-\infty < \rho < 1$. This function implies:

$$M(q) = q \left( \frac{q^\rho - 1}{\alpha} + 1 \right)^{-1/\rho}.$$  

With this function, parts (i) and (iii) of Assumption 2 are satisfied. Part (ii) is satisfied iff $-\infty < \rho < 0$. In this case, $\bar{q} = (1 - \alpha)^{1/\rho}$. Part (iv) is satisfied iff $\alpha \geq 1 - (1 - \rho)q^\rho / 2$. When $\rho \leq -1$, this condition is satisfied for all $\alpha > 0$. When $-1 < \rho < 0$, the condition puts a lower bound on $\alpha$. Note that, for $\rho < 0$, the derivative $M'(q)$ is unbounded at $q = \bar{q}$.

### 3. Workers’ and Firms’ Optimal Decisions

In this section, I characterize individuals’ optimal decisions and their value functions. Because this paper focuses on stationary equilibria, the time at which a contract is offered does not matter. Thus, unless it is necessary, I suppress the starting time of a contract, $s$, from the notation such as $w(t, s)$ and $V(t, s)$.

#### 3.1. Optimal Application

Workers’ search is directed by the employment rate function, $p(V)$. This equilibrium object gives the Poisson rate of obtaining an offer $V$ upon application. I will focus on equilibria
in which $p(.)$ has the properties described in the following claim. The procedure of the
analysis is to characterize individuals’ decisions first by assuming these properties of $p(.)$ and
then verify them in the equilibrium (as an implication of Lemma 5.1 later).

**Claim 1.** $p(V)$ is bounded, continuous and concave for all $V$. Moreover, $p(V)$ is continu-
ously differentiable and strictly decreasing for all $V < \bar{V}$, with $p(\bar{V}) = 0$.

Let me analyze the application decision of a worker who has tenure $t$. Refer to the
remaining value of the worker’s contract, $V(t)$, as the worker’s state. (For an unemployed
applicant, set $t = \emptyset$ and $V(\emptyset) = V_u$.) After receiving a job application opportunity, the
expected increase in the value for the worker is:

$$D(V(t)) = \max_{f \in [V(t), \bar{V}]} p(f) [f - V(t)].$$

Denote the solution as $f(t) = F(V(t))$. The following lemma describes the main features
of the optimal application (see Appendix A for a proof):

**Lemma 3.1.** $F(\bar{V}) = \bar{V}$. For all $V < \bar{V}$, the following results hold: (i) There is a unique
and interior solution to (3.1), $f = F(V)$; (ii) $F(.)$ is continuous and $D(V)$ is differentiable,
with $D'(V) = -p(F(V)) < 0$; (iii) $F(.)$ is strictly increasing; (iv) $F(V)$ obeys:

$$V = F(V) + \frac{p(F(V))}{p'(F(V))}.$$  \hspace{1cm} (3.2)

Moreover, $[F(V_2) - F(V_1)] / (V_2 - V_1) \leq 1/2$ for all $V_2 \neq V_1$; (v) If $p(.)$ is twice differentiable, then $F(V)$ is differentiable with $0 < F'(V) \leq 1/2$, and $D(V)$ is twice differentiable.

For a worker at a value $V$, the offer $F(V)$ is the only optimal target of application. Other offers are not optimal for this worker, despite that the worker observes all of them. Offers higher than $F(V)$ have too low employment probabilities to be optimal, while offers lower than $F(V)$ have too low values. Only the offer $F(V)$ provides the optimal tradeoff between the value and the probability of obtaining it.

Not only is a worker’s optimal target of application unique, it is also increasing in the
current value for the worker. That is, the higher the worker’s current value (state), the
higher the offer to which the worker will apply. Thus, the applicants choose to separate
themselves according to their states. This endogenous separation is optimal because an
applicant’s payoff function has the single-crossing property with respect to the worker’s
current value. An applicant’s current job is a backup for him when he fails to obtain the

9
applied job. The higher this backup value is, the more the worker can afford to “gamble” on the application and, hence, the higher the offer to which he will apply.

Figure 1 illustrates the single-crossing property in the $(f, p)$-space for two workers, 1 and 2. Worker 1 is at value $V_1$ and worker 2 at value $V_2$, where $V_2 > V_1$. Worker $i$’s indifference curve can be written as $f = V_i + D_i/p$, for $i = 1, 2$. Suppose that the two indifference curves cross each other at a point, $(f_0, p_0)$, where $f_0 > V_2$. At this point, the slope of worker $i$’s indifference curve is $df/dp < 0$, and the absolute value of this slope decreases with $V_i$. This implies that the worker with the higher value (worker 2) is more willing to tolerate a low employment probability than does the worker with the low value. Equivalently, to compensate for the same reduction in the probability of getting an offer, worker 2 needs a smaller increase in the offer than worker 1 does.

![Figure 1. Monotonicity of a worker’s optimal application](image)

The optimality of the application decision is one of the key differences between this model and the BC model or, more generally, between directed search and undirected search. By assumption, models with undirected search have no counterpart to the above decision problem of an applicant. This contrast between the two models leads to sharply different predictions on job-to-job transitions and mobility. Directed search predicts a definite pattern of transition and endogenously limited mobility of workers between values or wages. For example, take two workers whose current contracts have remaining values $V_1$ and $V_2$, respectively, with $V_1 < V_2$. Let $V_A = F(V_1)$ and $V_B = F(V_2)$. For the two workers, the probability of transiting immediately to a value above $V_B$ is zero. Moreover, conditional on that both have transited to new jobs, the likelihood between worker 2’s and worker 1’s probability of having transited to a value $V' \in [V_A, V_B)$ is zero. With undirected search, in
contrast, the probability of transiting to values above $V_B$ is positive for both workers, and the likelihood ratio of transiting to $V' \in [V_A, V_B)$ is a positive and finite constant.

In addition to limited wage mobility, directed search also yields predictions on the gain to a worker from an application. First, the higher the value of an applicant’s current job, the less he gains from an application. The expected gain from an application, $D(V)$, and the actual gain in percentage, $(F - V)/V$, both fall as $V$ increases. However, this decreasing gain does not necessarily mean a decreasing gain in wages, because it partly reflects the worker’s decreasing marginal utility. Second, $D''(V) > 0$. That is, the decrease in the expected gain from an application slows down as the worker’s current value increases.

### 3.2. Value Functions of Workers and Firms

Throughout this paper, denote $\dot{x} = dx/dt$ for any variable $x$. Recall that $t$ denotes tenure and that an unemployed worker’s tenure is denoted as $t = \emptyset$. For an employed worker, the value can change over time for four possible reasons. The first is the change in wages with tenure. The second is the event that the worker obtains a better offer and quits the current job. The third is death. The fourth is the adjustment to the steady state, which is abstracted from by the focus on stationary equilibria. Because the rate of time preference is zero, the value for an employed worker evolves as follows:

$$\dot{V}(t) = \delta V(t) - u(w(V(t))) - \lambda_1 D(V(t)),$$

where $D(V)$ is given by (3.1). Since $t$ denotes tenure, $\dot{V}(t) = 0$ if wages are constant over tenure. In particular, because the unemployment benefit does not change over time, the value for an unemployed worker (denoted as $V_u$) obeys:

$$0 = \delta V_u - u(b) - \lambda_0 D(V_u).$$

Now consider the value function of a firm that is employing a worker whose tenure is $t$ and whose remaining contract has a value $V(t)$. Let $J(V(t))$ denote this firm’s value.\(^8\)

\(^8\)Delacroix and Shi (2006) establish similar features in a model with directed search on the job, but they restrict that offers must be a constant wage over time. Nevertheless, the similarity suggests that these features are common ones of directed search.

\(^9\)The worker can also choose to quit the job to become unemployed if the wage profile is sufficiently decreasing. However, this event will never occur in the equilibrium, because the optimal wage profile has increasing wages with tenure, as shown later.

\(^10\)Strictly speaking, the firm’s value at any given tenure $t$ depends on the remaining contract, $\{V(\tau)\}_{\tau \geq t}$, not just on $V(t)$. However, given $V(t)$, the (maximized) value of the firm under an optimal contract that delivers the value $V(t)$ to the worker is a function of $V(t)$ alone. In order to economize on the notation, I treat $J$ as the maximized value of the firm.
Because the worker quits at rate $\lambda_1 p(F(V(t)))$ and dies at rate $\delta$, then

$$\dot{J}(V(t)) = [\delta + \lambda_1 p(F(V(t)))] J(V(t)) - y + w(V(t)), \quad (3.5)$$

where $\dot{J}$ denotes the derivative with respect to $t$ rather than $V$. This equation has embodied the aforementioned assumptions that a firm commits to the contract and that it does not respond to the employee’s outside offers.

For dynamic optimization, let me express $J$ as a discounted sum of profits by integrating (3.5). Since the contracts are expressed in terms of $V$, it is convenient to use $V$ rather than $t$ as the integration variable. To do so, use (3.3) to derive:

$$dt = \frac{dV}{\delta V - u(w(V)) - \lambda_1 D(V)}. \quad (3.6)$$

Substituting into (3.5), I obtain a differential equation for $J(V)$. To integrate this equation, let $t_a$ be an arbitrary point in $[0, t]$ and let $V_a = V(t_a)$. Let $\gamma(V(t), V_a)$ be the probability that a match will survive to tenure $t$ conditional on that it has survived to tenure $t_a$. Because the separation rate at any given value $V$ is $[\delta + \lambda_1 p(F(V))]$, then

$$\gamma(V, V_a) = \exp \left[ - \int_{V_a}^{V} \frac{\delta + \lambda_1 p(F(m))}{\delta m - u(w(m)) - \lambda_1 D(m)} dm \right], \quad (3.7)$$

where I have used (3.6) to substitute $dt$. Equivalently, $\gamma$ is given by the solution to the following differential equation:

$$\frac{d}{dV} \gamma(V, V_a) = - \frac{[\delta + \lambda_1 p(F(V))] \gamma(V, V_a)}{\delta V - u(w(V)) - \lambda_1 D(V)}, \quad (3.8)$$

where the terminal conditions are $\gamma(V_a, V_a) = 1$ and $\gamma(\bar{V}, V_a) = 0$ for all $V_a < \bar{V}$. Because $J$ is bounded, it satisfies the transversality condition $\lim_{V \to \bar{V}} J(V) \gamma(V, V_a) = 0$ for all $V_a < \bar{V}$. Integrating (3.5) with respect to $V$ yields:

$$J(V_a) = \int_{V_a}^{\bar{V}} \frac{[y - w(V)] \gamma(V, V_a)}{\delta V - u(w(V)) - \lambda_1 D(V)} dV.$$  

For any $V_a$, this value is determined by the contract remaining at tenure $t_a$.  

### 3.3. Optimal Recruiting Decisions and Contracts

A firm’s recruiting decision contains two parts. The first part is to choose an offer $V(0)$ to maximize the expected value of recruiting, $q(V(0))J(0)$, taking the function $q(V)$ as given. As I will explain later, the solution to this problem is a value $V(0)$ from a continuum. The
second part of a firm’s decision is to choose a wage path \( \{w(V)\}_{V=V(0)} \) to maximize \( J(0) \) and to deliver the value \( V(0) \). I characterize this decision below.

The optimal wage profile, \( \{w(V)\}_{V=V(0)} \), solves:

\[
(P) \quad \max_{J(V(0))}, \text{ given } V(0).
\]

As said earlier, the maximized value of \( J \) is indeed a function of \( V(0) \). Use the short-hand notation \( V_0 = V(0) \). I characterize this decision below.

The optimal wage profile, \( \{w(V)\}_{V=V(0)} \), solves:

\[
(P) \quad \max_{J(V(0))}, \text{ given } V(0).
\]

Following a similar argument to that in BC, it can be shown that the assumption \( u(0) = -\infty \) implies \( w(V(t)) > 0 \) for almost all \( t \) in all optimal contracts. In Appendix B, I show that the optimality conditions of the Hamiltonian problem yield: \( \Lambda = J \) and

\[
J'(V) = -\frac{1}{u'(w(V))} < 0. \tag{3.9}
\]

Optimal contracts have three important properties. First, all optimal contracts provide optimal sharing of the value between a firm and its worker, as described by (3.9). That is, the amount of wage increase that is needed to increase the worker’s utility by a marginal unit must be equal to the reduction in the firm’s profit. Note that (3.9) can be written equivalently as \( -\dot{J} = \dot{V}/u'(w) \). For the analysis later, it is useful to substitute \( \dot{V} \) from (3.3) and \( \dot{J} \) from (3.5) to rewrite this equation as:

\[
u'(w)(y - w) + u(w) = u'(w) [\delta + \lambda_1 p(F(V))] J(V) + [\delta V - \lambda_1 D(V)]. \tag{3.10}
\]

This equation can be directly explained by viewing a match as a joint asset. The left-hand side of the equation measures the flow of “dividends” to the asset, which consists of the firm’s profit, evaluated with the worker’s marginal utility, and the worker’s utility. The right-hand side is the “permanent income” in utils generated by the asset. In particular, the permanent income to the firm is \( [\delta + \lambda_1 p(F)]J \), which is translated into units of utility with the marginal utility of the worker. The permanent income to the worker is \( [\delta V - \lambda_1 D(V)] \).

The optimal contract requires that the flow of dividends to the joint asset should be equal to the permanent income of the asset.

Second, wages increase with tenure in all optimal contracts. This feature and the bounds on wages are stated as follows (see Appendix B for a proof):

\[\text{It can be shown that the program } (P) \text{ is concave in terms of Gâteaux derivatives in a neighborhood of the optimal contract, and so the optimum is characterized by the optimality conditions.}\]
Lemma 3.2. For all \( V(t) < \bar{V} \), wages in an optimal contract satisfy:

\[
0 < \frac{dw(V(t))}{dt} = \frac{[u'(w(V))]^2}{u''(w(V))} \lambda_1 J(V) \left[ \frac{dp(F(V))}{dV} \right].
\] (3.11)

Moreover, \( \bar{w} = y - \delta k/\bar{q} < y \), \( \bar{V} = u(\bar{w})/\delta \), \( J(\bar{V}) = k/\bar{q} > 0 \), and \( q(\bar{V}) = \bar{q} (< \infty) \).

There are two forces that make an optimal wage profile increase smoothly with tenure. The first is a firm’s incentive to retain a worker by backloading wages, which appears in (3.11) through the feature \( dp(F(V))/dV < 0 \); the second force is risk aversion, which appears in (3.11) as \( u'' < 0 \). These forces work as follows. Because a worker cannot commit to the job, the firm can increase the worker’s opportunity cost of quitting by backloading wages. As wages rise with tenure, the probability with which the employee can find a better offer elsewhere falls, and so the worker’s quit rate falls with tenure. Thus, a rising wage profile is less costly to the firm than a constant wage profile that provides the same expected value to the worker. However, if workers are risk neutral, then the best way for a firm to backload wages is to offer zero wage initially with a promised jump in wages in the future (see Stevens, 2004). Because this jump is not desirable for risk averse workers, the optimal contract has smoothly increasing wages over tenure.

Because wages are increasing with tenure and bounded above, wages in all optimal contracts increase toward the upper bound \( \bar{w} \) as \( t \to \infty \). Accordingly, the value for an employed worker converges to \( \bar{V} \). This convergence in the value is also monotonic, as I will show later in Corollary 5.3. As a result, a firm’s value falls with tenure.

The third property is that all optimal contracts are sections of a baseline contract but with different initial values or wages. More precisely, denoting the baseline contract as \( \{V_b(t)\}_{t=0}^{\infty} \), then the entire set of optimal contracts is:

\[
\{\{V(t)\}_{t=0}^{\infty} : V(t) = V_b(t + s) \text{ for all } t, \text{ where } s \in [0, \infty) \}.
\]

That is, relative to the baseline contract, any other optimal contract amounts effectively to crediting the worker with a length of tenure before the start of the contract and, hence, with a higher initial value (or wage). From this initial value, the contract traces out the remaining section of the baseline contract.\(^{12}\)

The above property follows from the principle of dynamic optimality. To explain why, compare the baseline contract with another contract \( c \) given by \( \{V_c(t)\}_{t=0}^{\infty} \), with \( V_c(0) > \)

\(^{12}\)It is worth repeating that an offer with a higher value is not necessarily more attractive to an applicant, because the probability of obtaining the offer is lower. A worker’s optimal target of application is unique, given the remaining value of his current contract.
Because the values in both contracts increase with tenure toward \( \bar{V} \), there exists a time \( s \) such that \( V_b(s) = V_c(0) \). If contract \( c \) is optimal for delivering the value \( V_c(0) \) but if it is not the same as the section of the baseline contract from that value onward, the firm that offers the baseline contract can replace the remaining part of the contract from \( V_b(s) \) onward by contract \( c \), with \( V_b(t) = V_c(t - s) \) for all \( t \geq s \). This change will increase the firm’s expected value, which contradicts the optimality of the baseline contract.

With the above property, characterizing the entire set of optimal contracts at any time is equivalent to tracing out the baseline contract over tenure. From now on, I focus on the baseline contract and suppress the subscript \( b \). In particular, \( V(t) \) denotes the remaining value of the baseline contract at tenure \( t \), and the wage function \( w(V) \) denotes the wage level at the point where the remaining value of the baseline contract is \( V \). Denote the inverse function of \( V(t) \) as \( T(V) \); that is,

\[
T(V(t)) = t \quad \text{for all } t \geq 0. \tag{3.12}
\]

\( T(x) \) records the length of tenure required for a worker to increase the value from \( V(0) \) to \( x \) according to the baseline contract. Clearly, \( T'(V) = dt/dV \), which is given by (3.6). Finally, the set of equilibrium offers, denoted as \( V \), is given by \( V = \{ V(t) : t \geq 0 \} \).

4. Configuration and Definition of the Equilibrium

To illustrate the configuration of the equilibrium, denote:

\[
v_0 = V_u \text{ and } v_j = F^{(j)}(v_0), \quad j = 1, 2, \ldots, \tag{4.1}
\]

where \( F^{(0)}(v_0) = v_0 \) and \( F^{(j)}(v_0) = F(F^{(j-1)}(v_0)) \). The set of offers in the equilibrium is \( V = [V(0), \bar{V}] \). Note that \( V(0) = v_1 \) and \( V(\infty) = \bar{V} \). Because only unemployed applicants apply to \( v_1 \), then \( v_1 = F(V_u) \). Moreover, note that \( F(V_u) > V_u = V + \lambda_0 D(V_u)/\delta > \underline{V} \).

Thus, \( v_1 = F(V_u) > \underline{V} \), which shows that \( V \) is a strict subset of \([\underline{V}, \bar{V}]\).

Figure 2 depicts the career paths of some workers in the equilibrium. For example, take a worker who is employed at value \( v_j \), where \( j \geq 1 \). The worker applies only to \( v_{j+1} \). If he gets the new job, the jump in the value is depicted by the arched arrow from \( v_j \) to \( v_{j+1} \). If the worker does not get the new job, the worker’s current contract provides increases in the value, which are depicted by the horizontal arrows. After the increase in the value, say to \( v_j + a \), the worker updates the target of the next application to \( F(v_j + a) \). This process continues until death. Therefore, the value for a worker increases over time as a result of
both the jumps created by successful applications and the smooth increases provided by the contracts. There are workers employed at every value in the interval \([v_1, \bar{V}]\).

If on-the-job search were prohibited, only one value, \(v_1\), would be offered in the equilibrium, as in most models of directed search cited in the introduction. If wages were restricted to be constant over tenure, then workers would be employed only in the discrete set \(\{v_1, v_2, ..., \bar{V}\}\) because there would be no increase in values to fill the gap between two adjacent levels. Such a model is analyzed by Delacroix and Shi (2006). By filling the gaps, wage-tenure contracts generate an interval of equilibrium offers. Thus, wage-tenure contracts are useful for a directed search model to generate a continuous distribution of wages among homogeneous workers, as well as the wage-tenure relationship.

![Career paths of some workers](image)

**Figure 2. Career paths of some workers**

Let \(n\) be the fraction of employed workers and \(G\) the cumulative distribution function of employed workers over values. As said earlier, wages and values refer to the baseline contract. An equilibrium is a set of lifetime utilities, \(\mathcal{V} = \{V(t) : t \geq 0\}\), a Poisson rate of employment, \(p(.)\), an application strategy, \(F(.)\), a value function \(J(.)\), a wage function \(w(.)\), and the distribution of workers, \((G, n)\), that satisfy the following requirements:

(i) \(F(V)\) solves (3.1), given \(p(.)\);

(ii) Given \(F(.)\) and \(p(.)\), each value \(V(t) \in \mathcal{V}\) is delivered by a contract that solves \((\mathcal{P})\) with a starting value \(V(t)\), and the resulting value function of the firm is \(J(V(t))\);

(iii) Zero expected profit of recruiting: \(q(V)J(V) = k\) for all \(V \in [V_0, \bar{V}]\), and \(q(V)J(V) < k\) otherwise, where \(q(V) = M^{-1}(p(V))\);

(iv) \(G\) and \(n\) are stationary.
Most elements of this definition are self-explanatory, except (iii). This requirement asks the function \( q(V) \) to induce zero expected profit of recruiting for all \( V \in [V, \bar{V}] \). Because \( V \) is a strict subset of \([V, \bar{V}]\), this requirement imposes a restriction on beliefs out of the equilibrium. The reason for imposing this restriction is as follows. For a non-equilibrium value \( V \notin V \), there can be two different reasons why the value is not in the equilibrium set. One is the self-fulfilling expectation that no worker will apply to that value: This expectation induces firms not to offer that value, in which case no worker will apply to that value, indeed. The second reason is that, even after firms offer that value, workers still find it optimal not to apply to it. The first reason for a “missing market” may not be robust to trembling that exogenously puts some recruiting firms at the value \( V \). Requirement (iii) excludes such non-robust equilibria and, hence, refines the set of equilibria.\(^{13}\)

Requirement (iii) determines the hiring rate function and, hence, the employment rate function. For given \( J(.) \), the requirement yields \( q(V) = k/J(V) \), and so \( p(V) = M(k/J(V)) \) for all \( V \in [V, \bar{V}] \). For all \( V > \bar{V} \), (iii) requires that a firm recruiting at \( V \) should make an expected loss. This part of the requirement is always satisfied, because Lemma 3.2 implies \( q(V)J(V) \leq \bar{q}J(V) < \bar{q}J(\bar{V}) = k \).

### Table 1. Comparing equilibrium definitions

<table>
<thead>
<tr>
<th>Application</th>
<th>Directed search</th>
<th>Undirected search</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>optimal: given ( p(.) ), choose ( F(V) );</td>
<td>random: acceptable offers arrive at rate ( \lambda_1 [1 - O(V)] );</td>
</tr>
<tr>
<td>Hiring rate</td>
<td>( q(V) = M^{-1}(p(V)) );</td>
<td>( q(V) = \lambda_1 nG(V) + \lambda_0 (1 - n) );</td>
</tr>
<tr>
<td>Free-entry</td>
<td>determining ( p(.) );</td>
<td>relating ( O(.) ) to ( G(.) );</td>
</tr>
<tr>
<td>Stationarity</td>
<td>partial differential equation of ( G );</td>
<td>partial differential equation of ( G );</td>
</tr>
</tbody>
</table>

The above equilibrium definition differs substantially from that in undirected search models, such as Burdett and Mortensen (1998) and BC. Table 1 compares the components of the equilibrium definitions in the two classes of models, where \( O(.) \) is the distribution of offers. One difference is in workers’ application decisions. While the decisions are random with undirected search, they are optimal with directed search.

\(^{13}\)Similar refinements have been used in directed search models, e.g., Acemoglu and Shimer (1999b) and Delacroix and Shi (2006). In Delacroix and Shi, the refinement restricts the applicants’, rather than firms’, expected payoff off the equilibrium path. It requires a worker’s expected surplus from applying to every offer (including a non-equilibrium offer) to be the same. In an environment with homogeneous workers, this alternative restriction achieves the same purpose as requirement (iii) does. However, when the applicants are heterogeneous as in the current model, the alternative restriction is not useful because it is not possible to find one function \( p(.) \) or \( q(.) \) that induces all applicants to be indifferent between equilibrium and non-equilibrium offers.
Another difference is in the role of the distribution of workers. When search is undirected, the distribution plays a critical role in the equilibrium. There, the employment rate and the hiring rate are functions of the distribution of offers, which is tied to the distribution of workers by the free-entry condition. As a result, individuals’ optimal decisions depend on the distribution of workers. In turn, this distribution obeys a non-linear partial differential equation which involves individuals’ optimal decisions. The task of determining the equilibrium is challenging. In contrast, with directed search, the workers who apply to any given offer \( V \) and the firms that recruit at \( V \) form a submarket, and the economy consists of a continuum of such submarkets. The only connection between the submarkets is the employment rate function, \( p(\cdot) \), or equivalently, the hiring rate function, \( q(\cdot) \). The free-entry condition of firms ties the loop by determining the employment rate function. Therefore, optimal decisions, optimal contracts, and the matching rate functions can all be determined by invoking (i) – (iii) in the above equilibrium definition, without any reference to the distribution of workers or offers. Thus, individuals’ decisions affect the distribution of workers, but the distribution plays no role in individuals’ decisions.\(^{14}\)

In the next section, I will determine the equilibrium functions, \( p(\cdot), q(\cdot), w(\cdot) \) and \( J(\cdot) \). Then, section 6 will solve for the distribution of workers.

### 5. Equilibrium Employment Rate and the Wage Function

The main step of determining an equilibrium is to determine the employment rate function, \( p(V) \). However, it is more convenient to organize the analysis around the wage function, \( w(V) \). The following procedure develops a mapping on \( w \) and obtains \( (p, J, F) \).

Start with an arbitrary function \( w(\cdot) \) and add the subscript \( w \) to other functions constructed with this given function. First, I integrate (3.9) and use \( J(\bar{V}) = \frac{k}{\bar{q}} \) to get:

\[
J_w(V) = \frac{k}{\bar{q}} + \int_V^V \frac{1}{u'(w(z))} \, dz.
\]  

(5.1)

Second, the zero-profit condition for recruiting yields \( q_w(V) = k/J_w(V) \) and, hence,

\[
p_w(V) = M \left( \frac{k}{J_w(V)} \right).
\]

(5.2)

Third, using \( p_w(V) \) as the employment rate, I can express an applicant’s optimal decision as \( f = F_w(V) \) and the expected gain as \( D_w(V) \).

\(^{14}\)Competitive entry of firms is important for this feature. If the number of firms is fixed, instead, then the expected value of recruiting is endogenous. Even in this case, the distribution of workers plays only a limited role in the equilibrium, because the expected value of recruiting is the only channel through which the distribution affects applications and contracts in the equilibrium.
Fourth, I explore (3.10) to obtain a mapping on \( w \). Treat \( w \) on the left-hand side of (3.10) as a variable but substitute the given function \( w(V) \) for \( w \) on the right-hand side. To avoid confusion, use \( w_1 \) instead of \( w \) on the left-hand side. Then,

\[
u(w_1) + u'(w_1)(y - w_1) = u'(w(V)) \left[ \delta + \lambda_1 p_w(F_w(V)) \right] J_w(V) + \delta V - \lambda_1 D_w(V).
\] (5.3)

Denote the solution for \( w_1 \) as \( w_1(V) = (\psi w)(V) \). Equilibrium wage function, \( w(V) \), is a fixed point of \( \psi \). That is, \( w(V) = (\psi w)(V) \) for all \( V \). This fixed point is independent of the distribution of workers. So are the functions \( p(\cdot), q(\cdot), J(\cdot) \) and \( F(\cdot) \).

To characterize the fixed point for \( w \), let me specify the bounds on various functions. First, using the constant \( \bar{w} \) to replace the function \( w(V) \) in (5.1) and (5.2), I obtain \( J_{\bar{w}}(V) \) and \( p_{\bar{w}}(V) \). Because \( J_w(\cdot) \) and \( p_w(\cdot) \) are monotone in \( w \), then \( J_w(V) \leq J_{\bar{w}}(V) \) and \( p_w(V) \leq p_{\bar{w}}(V) \) for all \( V \). Second, define

\[
q = k/J_{\bar{w}}(V).
\] (5.4)

Since \( J_{\bar{w}}(V) \) is decreasing in \( V \), \( q \in (0, \bar{q}) \) and \( q(V) \in [q, \bar{q}] \) for all \( V \). This lower bound on \( q \) is the one used in Assumption 2. Similarly, \( p(V) \) is bounded in \([0, M(q)]\). Third, let \( w \) be a strictly positive number that is sufficiently close to 0.

**Assumption 3.** Assume that \( b, V \) and \( w \) satisfy:

\[
0 < b < \bar{w} = y - \delta k/\bar{q},
\] (5.5)

\[
J_{\bar{w}}(V) \left[ \delta + \lambda_1 p_{\bar{w}}(V) \right] < y,
\] (5.6)

\[
u(w) + u'(w) \left[ y - w - J_w(V)(\delta + \lambda_1 p_w(V)) \right] \geq u(b).
\] (5.7)

The condition (5.5) is necessary for there to be any worker employed. (5.6) and (5.7) are sufficient conditions to ensure that \( \psi w(V) \geq \underline{w} \) for all \( V \), given that \( w(V) \geq \underline{w} \). The condition (5.6) requires that the permanent income of a job to a firm be less than output even when the firm is providing the lowest value to the worker. Because \( J_{\bar{w}}(V) \) and \( p_{\bar{w}}(V) \) are decreasing functions and because \( V \) is an increasing function of \( b \), (5.6) puts a lower bound on \( b \). This lower bound is strictly lower than \( \bar{w} \) because (5.6) is satisfied when \( b = \bar{w} \). (5.7) imposes an upper bound on \( w \). Because \( \underline{w} \) can be chosen to be sufficiently close to 0 and because \( b > 0 \), a sufficient condition for (5.7) is:

\[
\lim_{w \to 0} \left[ u(w) + u'(w)(a - w) \right] = \infty \text{ for all } a > 0.
\]
This sufficient condition is satisfied by the example \( u(w) = (w^{1 - \eta} - 1)/(1 - \eta) \) with \( \eta > 1 \).

Define

\[
\Omega = \{ w(V) : w(V) \in [\underline{w}, \bar{w}] \text{ for all } V; w(\bar{V}) = \bar{w}; \\
\text{and } w(V) \text{ is continuous and (weakly) increasing} \},
\]

\[
\Omega' = \{ w \in \Omega : w(V) \text{ is strictly increasing for all } V < \bar{V} \}.
\]

I establish that a fixed point of \( \psi \) exists in \( \Omega \) and then show that it lies in the subset \( \Omega' \).

First, the following lemma holds (see Appendix B for a proof):

**Lemma 5.1.** For any \( w \in \Omega \), \( J_w(V) \) defined by (5.1) is strictly positive, bounded, strictly decreasing and continuously differentiable for all \( V \). The function \( p_w(V) \) defined by (5.2) indeed has all the properties stated in Claim 1.

Because \( p_w(V) \) has all the properties in Claim 1, parts (i) - (iv) of Lemma 3.1 hold. In particular, there is a unique and interior solution to (3.1), \( F_w(V) \), which is continuous and strictly increasing for all \( V < \bar{V} \). Moreover, \( D'_w(V) = -p_w(F_w(V)) < 0 \).

**Theorem 5.2.** Maintain Assumptions 1, 2 and 3. Assume that the image of \( \psi \) is compact. Then, the mapping \( \psi \) has a fixed point \( w^* \in \Omega' \). The implied functions \( J_{w^*}(V) \) and \( p_{w^*}(V) \) are strictly concave, in addition to the properties stated in Lemma 5.1.

**Proof.** See Appendix C.

This theorem implies the central properties of wage-tenure and wage-quit relationships. First, wages and values increase with tenure. Second, because \( p(.) \) is decreasing and \( F(.) \) is increasing, the quit rate of a worker decreases with tenure and wages. Note that the theorem does not claim uniqueness of the equilibrium, because it is difficult to prove that the mapping \( \psi \) is a contraction. However, in the numerical example in section 7, the equilibrium is unique.

In the remainder of this paper, I will suppress the * on the fixed point and the subscript \( w^* \) on the equilibrium functions \( J, p, F \) and \( D \). Moreover, I will focus on wage profiles that are smooth over tenure. The following corollary describes the additional properties generated by smooth wage profiles (see Appendix D for a proof):

**Corollary 5.3.** If \( w(V(t)) \) is a smooth function, i.e., if \( |\dot{w}(V(t))| < \infty \) for all \( t \), then \( w(V) \) is differentiable, with \( 0 < w'(V) < \infty \) for all \( V \). Moreover, the following results hold for all \( V < \bar{V} \): (i) the derivatives \( J''(V) \), \( p''(V) \) and \( F'(V) \) exist and are finite; (ii) \( \bar{V} \) and \( \dot{J} \) both exist, with \( \bar{V} > 0 \) and \( \dot{J}(V) < 0 \).
To conclude this section, let me compare wages with unemployment benefits. To do so, let $B(V)$ be the benefit that achieves the value $V$ for an unemployed worker. For any given $V$, $B(V)$ is the solution for $b$ in (3.4) where $V_u$ is replaced with $V$. That is,

$$B(V) = u^{-1}(\delta V - \lambda_0 D(V)).$$

(5.8)

Since $D'(V) < 0$, then $B'(V) > 0$. Moreover, $B(V) = \bar{w}$. The following corollary follows from (3.3) and the feature $\dot{V} > 0$ for all $V < \bar{V}$ (the proof is omitted):

**Corollary 5.4.** $w(\bar{V}) = B(\bar{V})$. If $\lambda_0 \leq \lambda_1$, then $w(V) < B(V)$ for all $V \in [v_1, \bar{V})$.

The novel part of this corollary is the case $\lambda_0 = \lambda_1$. In this case, an unemployed worker has the same access to jobs as an employed worker, but the unemployment benefit must be higher than the wage in order for an unemployed worker to achieve the same value $V$ as an employed worker. If the unemployment benefit is the same as (or lower than) an employed worker’s wage, the present value for the unemployed worker is lower than that for the employed worker. The reason is that an employed worker enjoys the prospect of rising wages over tenure while an unemployed worker’s benefit does not change over time. If an unemployed worker wants to achieve the same value as an employed worker, this disadvantage of an unemployed worker must be compensated by a higher unemployment benefit. The result may hold even for some $\lambda_0 > \lambda_1$.\(^{15}\)

**6. Equilibrium Distributions of Workers and Firms**

Now let me turn to the distribution of workers. Recall that $n$ is the fraction of employed workers and $G$ is the cumulative distribution function of employed workers. Let $g$ be the density function corresponding to $G$. Note that this distribution is over values. Once this distribution is determined, the distribution of employed workers over wages can be deduced from $G_w(w(V)) = G(V)$, with a density function $g_w(w) = g(V)/w'(V)$.

To determine the distribution, let me examine the group of workers who are employed at values less than or equal to $V$, where $V \in [v_1, \bar{V}]$. The measure of this group is $nG(V)$. There is only one inflow into this group, which is the flow of workers from unemployment into employment at the value $v_1$. In a small interval of time, $(dt)$, this inflow is $\lambda_0 (1 - n) p(v_1) dt$. There are three flows out of the described group. First, the workers in the group may die, the flow of which is $\delta nG(V) dt$. Second, for the workers whose values lie

\(^{15}\)Of course, if $\lambda_0 < \lambda_1$, then an unemployed worker has a more difficult access to jobs than an employed worker. This additional reason for $B(V) > w(V)$ is discussed in Burdett and Mortensen (1998).
in \((V - \dot{V}dt, V]\), their contracts increase their values above \(V\) after the length of time \((dt)\). The measure of this outflow is \(n[G(V) - G(V - \dot{V}dt)]\). Third, some workers in the group quit for offers whose values are higher than \(V\). These quitters are currently employed in \(\left(F^{-1}(V), V\right]\) if \(F^{-1}(V) \geq v_1\), and in \((v_1, V]\) if \(F^{-1}(V) < v_1\). Thus, quitting generates the following measure of outflow from the described group:

\[
(dt) \lambda_1 n \int_{\max\{v_1, F^{-1}(V)\}}^{V} p(F(z)) dG(z) .
\]

For the distribution of employed workers to be stationary, the flow of workers into the described group must be equal to the sum of the outflows. Imposing this requirement, re-arranging terms and taking the limit \(dt \downarrow 0\), I have:

\[
\lim_{dt \downarrow 0} \frac{G(V) - G(V - \dot{V}dt)}{dt} = \lambda_0 \frac{1 - n}{n} p(v_1) - \delta G(V) - \lambda_1 \int_{\max\{v_1, F^{-1}(V)\}}^{V} p(F(z)) dG(z) .
\] (6.1)

To solve for the distribution, partition the support of \(G\) into subintervals \([v_j, v_{j+1})\), where \(v_j\) is defined by (4.1). Add a subscript \(j\) to \(g(V)\) and \(G(V)\) for \(V \in [v_j, v_{j+1})\). I prove the following theorem in Appendix E:

**Theorem 6.1.** The fraction of employed workers is:

\[
n = \frac{\lambda_0 p(v_1)}{\delta + \lambda_0 p(v_1)} .
\] (6.2)

The distribution of employed workers, \(G\), is continuous for all \(V\), with \(G(v_1) = 0\). The density function, \(g(V)\), is continuously differentiable for all \(V \neq v_2\) and it satisfies:

\[
g(V) \dot{V} = \delta [1 - G(V)] - \lambda_1 \int_{\max\{v_1, F^{-1}(V)\}}^{V} p(F(z)) dG(z) .
\] (6.3)

With \(V(0) = v_1\), \(T\) in (3.12) and \(\gamma\) in (3.7), \(g\) can be solved piece-wise as follows:

\[
g_1(V) \dot{V} = \delta \gamma(V, v_1) ,
\] (6.4)

\[
g_j(V) \dot{V} - g_j(v_j) \dot{v_j} \gamma(V, v_j) = \lambda_1 \int_{v_j}^{V} \gamma(V, z)p(z)g_{j-1}(F^{-1}(z))dF^{-1}(z)
\] (6.5)

where (6.5) holds for \(j \geq 2\). Moreover, \(g_j(v_j) = \lim_{V \uparrow v_j} g_{j-1}(V)\) for all \(j\).

The above theorem documents a few features of the equilibrium distribution of employed workers. First, the distribution is non-degenerate and continuous, despite the facts that all
workers and all matches are identical. This feature is remarkable because it is not possessed by previous models of directed search. As mentioned in the introduction, previous models of directed search can only produce a small number of equilibrium wages among homogeneous workers, and this number is often one. On the other hand, the model of undirected search by Burdett and Mortensen (1998) can generate a continuous distribution of values or wages, but the distribution becomes degenerate once each worker is allowed to observe two or more offers before search. In my model, each searching worker observes all offers before application. Yet, the equilibrium features a continuous distribution of employed values or wages among homogeneous workers. The causes for this continuous distribution are on-the-job search and wage-tenure contracts. On-the-job search produces a wage ladder among homogeneous workers, because a worker who gets a job earlier will search for a better job than a worker who gets the same job later. On the other hand, wage-tenure contracts increase wages smoothly with tenure, thus filling in the gap between two adjacent rungs of the wage ladder and smoothing the wage distribution.

Second, there is no mass at the lowest value of employment, since \( G(v_1) = 0 \). This feature is remarkable because there is a positive mass of unemployed workers who apply only to \( v_1 \). Despite this concentration of applications, there is no build-up of workers at \( v_1 \). All the workers who are employed at \( v_1 \) will only stay at \( v_1 \) for a very short length of time. Some of them will quit for better jobs or die, while the rest of them will experience wage increases according to the contracts. Thus, the mass of workers at \( v_1 \) is zero in the stationary equilibrium. In fact, there is no mass point anywhere in the support of the distribution. Moreover, the density function is differentiable except for \( V = v_2 \).

Third, the density function of employed workers can be computed recursively. Starting with \( j = 1 \), one can compute \( g_1 \) from (6.4). Taking the limit \( V \uparrow v_2 \) in the formula yields \( g_2(v_2) \). Then, setting \( j = 2 \) in (6.5) yields \( g_2(V) \). Taking the limit \( V \uparrow v_3 \) in the result yields \( g_3(v_3) \). Continuing this process, one can obtain \( g_j \) for all \( j \).

The following corollary describes the upper tail of the density function (see Appendix E for a proof):

**Corollary 6.2.** \( g(\bar{V}) = 0 \) and \( g_w(\bar{w}) = 0 \) if and only if \( |\delta - \lambda_1 p(\bar{V})/F'(\bar{V})| \neq 0 \). Under this condition, the density function of employed workers is decreasing at values sufficiently close to \( \bar{V} \) or, equivalently, at wages close to \( \bar{w} \). Moreover, the CES matching function in

\[16\] I thank Guido Menzio for pointing out that the distribution is continuous at \( v_1 \). The density function may fail to be differentiable at \( v_2 \) because offers below \( v_2 \) in the interval \( (v_1, v_2) \) do not receive applications from employed workers, while offers above \( v_2 \) do.
Example 2.1 satisfies the condition, $|\delta - \lambda_1 p(\bar{V})/F'(\bar{V})| \neq 0$, while the urn-ball matching function satisfies the condition if and only if $\bar{q} \neq \delta / \lambda_1$.

The decreasing density function at high wages is a robust feature of the data (see Kiefer and Neumann, 1993). Note that the condition required for $g(\bar{V}) = 0$ in the corollary is satisfied easily by the two matching functions in Example 2. Also note that undirected search models with homogeneous matches produce an increasing density function (see Burdett and Mortensen, 1998, and BC).

Directed search is able to generating a decreasing density function at high values or wages because workers choose their applications optimally. To see why, consider a worker with a value $V$ with $F(V) < \bar{V}$. This worker also observes all other offers. For the worker to apply to the target value, $F(V)$, rather than higher offers, higher offers must be more difficult to be obtained than the target value. For this to be true, the measure of recruiting firms per applicant must be smaller at high values than at the target value. In particular, at values close to the upper bound $\bar{V}$, the measure of recruiting firms per applicant should be close to zero. In turn, as few workers apply to such high values, it is indeed optimal for only few firms to recruit at these values. The measure of workers who succeed in obtaining jobs at values near $\bar{V}$ is close to zero. This feature makes the density function of employed values decreasing near the upper end of the distribution.

In contrast, undirected search models produce an increasing density function. By the assumption of undirected search, all applicants apply to a high offer and to a low offer with the same probability. Relative to a low offer, a high offer has the advantage of increasing acceptance and retention, without the negative response on the application side. Thus, more firms recruit at high values than at low values. The increasing density of recruiting firms ensures that recruiting yields the same expected profit at all equilibrium values. It also results in more workers being employed at high values than at low values, i.e., an increasing density function of employed workers. To modify this unrealistic prediction, models of undirected search had to introduce particular distributions of heterogeneity across matches in workers’ or firms’ characteristics (e.g., van den Berg and Ridder, 1998).

Let me illustrate more formally why the two models generate different shapes of the upper tail of the density function. With competitive entry, both models require that a firm’s hiring rate at any given offer $V$ should satisfy the zero-profit condition: $q(V) = k/J(V)$. Moreover, because the value function of a firm, $J(V)$, is decreasing and concave in both models, $q(V)$ is increasing and convex. However, the link between $q(V)$ and the distribution of employed workers differs between the two models. With undirected search, this link is
tight because the acceptance of a firm’s offer $V$ can come from any applicant whose current value is below $V$. That is,

$$q(V) = \lambda_1 n G(V) + \lambda_0 (1 - n).$$

Because $q$ is convex, then $G$ must be convex, which implies that the density function must be increasing. Directed search breaks this link between $G$ and $q$. With directed search, because the firms offering $V$ only attract the applicants whose current value is $F^{-1}(V)$, these firms and applicants form a submarket that is separated from other firms and applicants. As a result, a firm’s hiring rate does not depend on the distribution of workers. Then, optimal applications require that the density function of employed workers should be decreasing at high values or wages, as explained above.$^{17}$

On the lower end of the distribution of employed workers, the density function in the current model may or may not be increasing. In particular, it can be verified that $g'(v_1) > 0$ if and only if $w'(v_1) > 2 [\delta + \lambda_1 p(v_2)]$. This condition is needed because there is a mass of unemployed workers who apply only for jobs at $v_1$, which generates a large flow into employment at $v_1$. For the density function to be increasing around $v_1$, wages must increase sufficiently with tenure to take these workers quickly out of $v_1$. However, this condition would not be necessary for the described result if unemployed workers drew benefits from a continuous distribution, rather than having the same benefit. I have examined this extension of the model in a previous version of the paper and showed that the model exhibits $g(v_1) = 0$ as well as $g(\bar{V}) = 0$.

7. Comparative Statics and a Numerical Example

In this section, I conduct comparative statics to further contrast the current model with undirected search. Moreover, I compute an example to illustrate the equilibrium.

The following corollary summarizes the effects of the unemployment benefit, $b$, and the probability with which an unemployed worker obtains the application opportunity, $\lambda_0$:

**Corollary 7.1.** The parameters $b$ and $\lambda_0$ do not affect the functions $w(\cdot)$, $F(\cdot)$, $p(\cdot)$, $q(\cdot)$ and $J(\cdot)$. However, an increase in $b$ increases the value for unemployed workers, $V_u$.

$^{17}$In related models of directed search, Delacroix and Shi (2006) and Galenianos and Kircher (2005) also establish the feature that the density function of employed wages is decreasing at high wages. However, they restrict firms’ offers to wage levels rather than wage-tenure contracts.

$^{18}$To verify, set $V = v_1$ in (6.4) and note $T(v_1) = 0$. Then $g(v_1) = \delta$. Substituting the expression for $V$ in (3.3) and differentiating with respect to $v_1$, one obtains the described condition for $g'(v_1) > 0$. 

25
increases the lowest value offered in the equilibrium, $v_1$, reduces the measure of employed workers, $n$, and affects the distribution of workers. An increase in $\lambda_0$ also increases $V_u$ and $v_1$, but its effect on $n$ is ambiguous.

**Proof.** The analysis in section 5 is independent of $b$, $\lambda_0$ and $G$. Thus, the functions, $w(\cdot)$, $p(\cdot)$, $q(\cdot)$, $J(\cdot)$, $F(\cdot)$, and $D(\cdot)$ do not change with $b$ or $\lambda_0$. However, the increase in $b$ or $\lambda_0$ increases $V_u$ (see (3.4)). Since $v_1 = F(V_u)$, then $v_1$ increases with $b$ and $\lambda_0$. By (6.2), an increase in $b$ increases $n$, but an increase in $\lambda_0$ has an ambiguous effect on $n$. These changes in $v_1$ and $n$ lead to changes in the distribution, $G$. QED

One effect of a higher $b$ or $\lambda_0$ is that it increases the lowest equilibrium value of employment. Let $\hat{v}_1$ be this new level of $v_1$. Then, the new baseline contract is the section of the original baseline contract that starts at $\hat{v}_1$. Thus, conditional on a worker’s current wage, the worker’s optimal application, the wage-tenure contract and the worker’s transition rate to another job are all independent of $b$ and $\lambda_0$.

The reason for this independence is that the distribution plays no role in the determination of these functions, as discussed in section 4 and demonstrated in section 5. To recapitulate the mechanism, start with a worker’s application. For a worker’s application, $F(V)$, he only needs to know the employment rate function. In turn, the employment rate must satisfy the requirement that recruiting should yield zero expected profit at all equilibrium offers. This requirement pins down $p(V)$, given a firm’s value function, $J(V)$. Moreover, a firm value function depends only on what happens after the hiring, that is, on the wage function $w(V)$ offered and the worker’s quit rate, $p(F(V))$. Because $p(\cdot)$ and $F(\cdot)$ are only functions of optimal contracts by the proceeding argument, the only thing still to be determined is the function $w(V)$. The function $w(V)$ provides efficient sharing of the value between a firm and a worker, in the sense that $-\dot{J} = \dot{V}/u'(w)$. Again, $J$ and $V$ involve only the functions $F(V)$, $p(V)$, $J(V)$ and $w(V)$ (see (3.3) and (3.5)). The solution to this fixed-point problem is independent of the distribution of employed workers.

The above comparative statics have an obvious policy implication. They suggest that putting resources to change the aspects of the market related to unemployed workers is not an effective way to change wage contracts and wage mobility. The effective way is to directly change the aspects relevant for employed workers, such as $\lambda_1$. This is not the prediction of undirected search models. There, an increase in the unemployment benefit reduces the probability with which a given offer will be accepted by a worker, thereby increasing the equilibrium distribution of offers. As more firms offer high values than before, workers quit more quickly from low-value jobs. In order to mitigate this increase in
quits, firms offer contracts in which wages increase more quickly with tenure than before.\textsuperscript{19}

To conclude the analysis in this paper, let me compute one example. Let the matching function be the urn-ball matching function given in Example 2.1 and the utility function be $u(w) = (w^{1-\eta} - 1) / (1 - \eta)$. The parameter are given the following values:

$$
y = 100, k = 2, b = 40.5, \eta = 1.005, \bar{q} = 0.01, \\
\delta = 0.05, \lambda_0 = 1, \lambda_1 = 0.5.
$$

Because this example is only for illustration, the above parameter values have no particular significance. They imply that the highest equilibrium wage is $\bar{w} = 90$, which corresponds to the highest offer in the equilibrium, $\bar{V} \approx 89$. Moreover, $\underline{V} = u(b) / \delta \approx 73.35$. Set $w = b \times 10^{-5}$. Then, Assumption 3 is satisfied. To compute the model, I discretize the support of the function $w(V)$ into $\{V(1), V(2), ..., V(N)\}$, where $V(1) = \underline{V}$ and $V(N) = \bar{V}$, and use the mapping $T$ to iterate on the values of $w$ at these grid points. After obtaining the fixed point for $w$, I recover other functions ($q, p, F, D, T$).

I compute $dG$ recursively through the following discretized version of (6.3):

$$
\frac{dG(V(j))}{dT(V(j))} = \delta \left[1 - G(V(j+1))\right] - \sum_{i=\max\{v_1, F^{-1}(V(j+1))\}}^{j+1} \lambda_1 p(F(V(i-1)))dG(V(i-1)).
$$

The densify function is $g(V(j)) = dG(V(j)) / dV(j)$, where $dV(j) = V(j+1) - V(j)$.

I compute the equilibrium under these parameter values first and then contrast it with the equilibrium under $b = 27$. Only the diagrams for $b = 27$ will be exhibited but they are

\textsuperscript{19}One can verify these statements about undirected search models by introducing a continuous distribution of unemployment benefits into BC or Burdett and Mortensen (1998).
useful for discussing the results under both values of $b$.

![Graph](image)

Figure 3a. A worker’s wage and value over tenure

Figures 3a through 3d depict equilibrium contracts, the matching rate functions, and optimal application for $b = 27$. These figures confirm the properties established in previous sections. In particular, wages and workers’ values increase with tenure, and the employment rate is a strictly decreasing and concave function of the offer. For $b = 27$, the lowest equilibrium wage is $w_1 = 73.7$ and the lowest value offered is $v_1 = 85.75$. Starting at $w_1$, a worker needs to change jobs upward six times to get close to the highest wage, even if he succeeds in getting the offer every time he applies. If he always stays with a job, his wage will eventually get close to the highest level after a length of tenure equal to 77. For $b = 40.5$, in contrast, the lowest equilibrium wage is $w_1 = 76.43$ and the lowest value offered is $v_1 = 86.56$. In this case, it takes a length of tenure equal to 50 for a worker to get from $w_1$ to a level close to the highest wage, but the number of job changes required to get close to the highest wage is still six.
As analyzed above, an increase in \( b \) from 27 to 40.5 does not change the functions in Figures 3a through 3c. Instead, the increase in \( b \) only truncates these functions on the left, by increasing the lowest equilibrium offer from 85.75 to 86.56. In Figure 3a, the increase in \( b \) amounts to offering an unemployed worker the same contract as the one that a worker with a length of tenure 27 has in the equilibrium with \( b = 27 \). The contracts on the left of the vertical axis in Figure 3a are no longer offered in the equilibrium, but the new equilibrium contracts are given by the same function \( w(.) \), with \( w(t + 27) \) being the wage for a worker with a length of tenure \( t \). Similarly, in Figures 3b and 3c, the functions \( p(V), q(V) \) and \( F(V) \) under the higher \( b \) are the same as those under the lower \( b \), but the equilibrium offers start at \( v_1 = 86.56 \), where the vertical axes are drawn.
The distribution of wages depends on $b$. Figure 3d depicts the density function and the cumulative distribution of wages for $b = 27$. This figure provides additional information to...
the analysis in section 6. First, the density function has two main sections. One section is for \( w \in [w_1, w_2) \), where \( w_2 = 81.51 \), and the other section is for \( w > w_2 \). The density is decreasing for \( w < w_2 \). For \( w > w_2 \), the density first increases and then decreases.

Second, wages are distributed primarily in the section with \( w < w_2 \). As shown by the cumulative distribution, most workers (more than 90% of them) are employed below \( w_2 \). Similar features emerge when \( b \) is increased to 40.5. In this case, the lowest equilibrium wage is \( w_1 = 76.43 \). As a result, the measure of workers on the left hand side of the vertical axis in Figure 3d is redistributed to the right-hand side of the axis.

8. Conclusion

In this paper, I analyze the equilibrium in a labor market where firms offer wage-tenure contracts to direct the search of employed and unemployed workers. Each applicant observes all offers and there is no coordination among individuals. Because search is directed, workers’ applications (as well as firms’ recruiting decisions) must be optimal. This optimality requires the equilibrium to be formulated differently from the that in the large literature of undirected search. I provide such a formulation and show that the equilibrium exists. In the equilibrium, individuals explicitly tradeoff between an offer and the matching rate at that offer. This tradeoff yields a unique offer as the optimal target of a worker’s application. Despite this uniqueness and directed search, the stationary equilibrium has a non-degenerate and continuous distribution of wages.

One cause of wage dispersion is on-the-job search. Although all workers are identical in the model, some workers get jobs earlier than others and, hence, they will apply to high offers subsequently. This process creates a wage ladder. The other cause of the wage distribution is wage-tenure contracts, which fill in the gap between two rungs of the wage ladder. With risk-averse workers and imperfect capital markets, it is optimal for a firm to offer a wage profile that increases smoothly with tenure. Such a contract provides partial insurance to the worker and backloads wages to increase retention of the worker. The positive wage-tenure relationship implies that workers who are employed under the same contract but at different times may earn different wages. It also implies that the quit rate falls with tenure.

While preserving these realistic relationships between wages/quits and tenure, the current model generates several novel implications. First, because applicants separate themselves according to their current values, wage mobility is endogenously limited. Second, the density function of the wage distribution is decreasing at high wages, even when all
worker-firm pairs are equally productive. Finally, an increase in the unemployment benefit has no effect on an employed worker’s job-to-job transition rate and path, conditional on the worker’s current wage, although it affects the wage distribution.

The model constructed in this paper is tractable for studying business cycles with on-the-job search. A striking feature of the equilibrium is the dichotomy that individuals’ decisions and equilibrium contracts can all be characterized without any reference to the distribution of workers. With undirected search, instead, the distribution is a state variable in every individual’s decision problem. As the distribution evolves endogenously over business cycles, the large dimensionality of the state variables makes the task of determining the dynamic equilibrium analytically intractable and quantitatively challenging. The dichotomy eliminates this difficulty. Utilizing this feature, Menzio and Shi (2007) examine the implications of on-the-job search on business cycles, by incorporating aggregate and match-specific shocks into the model.
Appendix

A. Proof of Lemma 3.1

The result \( F(\bar{V}) = \bar{V} \) is evident. Let \( V < \bar{V} \) in the following proof. Temporarily denote \( K(f, V) = p(f)(f - V) \). Because \( p(.) \) is continuous and bounded, as stated in Claim 1, \( K(f, V) \) is continuous and bounded. Thus, the maximization problem in (3.1) has a solution. Since all interior values of \( f \) yield \( K(f, V) > 0 \), while the choices at the corners yield \( K(\bar{V}, V) = 0 = K(V, V) \) (because \( p(\bar{V}) = 0 \)), then the solution is interior. To show that the solution is unique, I show that \( K(f, V) \) is strictly concave in \( f \) for all \( f \in (V, \bar{V}) \).

To do so, let \( \alpha \in (0, 1) \). Let \( f_1 \) and \( f_2 \) be two arbitrary interior values with \( f_2 > f_1 > V \). Denote \( f_\alpha = \alpha f_1 + (1 - \alpha)f_2 \). Then,

\[
K(f_\alpha, V) = p(f_\alpha)[\alpha(f_1 - V) + (1 - \alpha)(f_2 - V)] \\
&\geq [\alpha p(f_1) + (1 - \alpha)p(f_2)][\alpha(f_1 - V) + (1 - \alpha)(f_2 - V)] \\
= \alpha K(f_1, V) + (1 - \alpha)K(f_2, V) + \alpha(1 - \alpha)[p(f_1) - p(f_2)]|f_2 - f_1| \\
> \alpha K(f_1, V) + (1 - \alpha)K(f_2, V).
\]

The two equalities come from rewriting, the first inequality from concavity of \( p \), and the last inequality from the feature that \( p(f) \) is strictly decreasing. Thus, \( K(f, V) \) is strictly concave in \( f \), which establishes part (i) of the Lemma.

For part (ii), uniqueness of the solution implies that \( F(.) \) is continuous by the Theorem of the Maximum. To show that \( D(.) \) is differentiable, let \( V_1 \) and \( V_2 \) be two arbitrary values with \( V_1 < V_2 < \bar{V} \). Express \( F_i = F(V_i) \) for \( i = 1, 2 \). Uniqueness of the solution implies \( K(F_1, V_1) > K(F_2, V_1) \) and \( K(F_2, V_2) > K(F_1, V_2) \). Thus,

\[
D(V_2) - D(V_1) > K(F_1, V_2) - K(F_1, V_1) = -p(F_1)(V_2 - V_1); \\
D(V_2) - D(V_1) < K(F_2, V_2) - K(F_2, V_1) = -p(F_2)(V_2 - V_1).
\]

Divide the two inequalities by \( (V_2 - V_1) \) and take the limit \( V_2 \to V_1 \). Because \( F(.) \) is continuous, the limits show that \( D(V) \) is differentiable at \( V_1 \) and that \( D'(V_1) = -p(F_1) \).

Since \( V_1 \) is arbitrary, this argument establishes part (ii).

For part (iii), again take two arbitrary values \( V_1 \) and \( V_2 \), with \( V_1 < V_2 \). Then, \( p(F_j)(F_j - V_i) < p(F_i)(F_i - V_i) \) for \( j \neq i \). I have:

\[
0 > [p(F_2)(F_2 - V_1) - p(F_1)(F_1 - V_1)] + [p(F_1)(F_1 - V_2) - p(F_2)(F_2 - V_2)] \\
= p(F_2)(V_2 - V_1) + p(F_1)(V_1 - V_2) = [p(F_2) - p(F_1)](V_2 - V_1).
\]

This result implies \( p(F_2) < p(F_1) \). Because \( p(.) \) is strictly decreasing, \( F(V_2) > F(V_1) \).

For part (iv), note that differentiability of \( p \) implies that \( F(V) \) is given by the first-order condition, (3.2). Also, because \( p \) is concave and decreasing, the following inequalities hold for all \( V_1 \) and \( V_2 \) with \( V_2 > V_1 \):

\[
p(F_1) \geq p(F_2) - p'(F_1)(F_2 - F_1),
\]
Thus, verify that workers who are employed at \( \bar{w} = \delta \) continuously differ.

This implies \( (F_2 - F_1) / (V_2 - V_1) \leq 1/2 \) for all \( V_2 \neq V_1 \), and so \( F \) is Lipschitz.

Finally, for part (v), if \( p \) is twice differentiable, then differentiating the first-order condition generates the derivative \( F'(V) \), and the above Lipschitz property yields \( F'(V) \leq 1/2 \). In this case, \( D'(V) = -p'(F(V))F'(V) \). QED

**B. Proofs of Lemmas 3.2 and 5.1**

I first derive the optimality conditions of the problem \( \mathcal{P} \) and the condition (3.11). From the Hamiltonian of the problem \( \mathcal{P} \), the optimality condition for \( \gamma \) is \( H/\gamma = -d\Lambda/dV \) and the optimality condition for \( w(V) \) is \( H/\gamma = 1/u'(w(V)) \). Integrating the condition for \( \gamma \) yields \( \Lambda = J(V) \), which implies \( H/\gamma = -J'(V) \). Substituting into the condition for \( w(V) \) yields (3.9). To derive (3.11), use the definition of the Hamiltonian and the optimality condition for \( w \) to solve:

\[
\Lambda = \frac{1}{\delta + \lambda q p(F(V))} \left[ y - w(V) - \frac{\delta V - u(w(V)) - \lambda D(V)}{u'(w(V))} \right].
\]

Differentiating this equation with respect to \( V \) and substituting the result \( d\Lambda/dV = -1/u' \), one obtains an expression for \( u'(V) \). Substituting this expression and (3.3) into \( \dot{w} = u'(V)\dot{V} \) yields (3.11).

Next, I prove the rest of Lemma 3.2. By Claim 1 and Lemma 3.1, \( p'(F(V)) < 0 \) and \( F'(V) > 0 \) for all \( V < \tilde{V} \). Because \( J(V) > 0 \) for all \( V \), as shown later, then (3.11) implies \( \dot{w} = \dot{w}(V(t)) > 0 \) for all \( V(t) < \tilde{V} \). Because \( \tilde{V} \) is the highest value offered, then \( p(F(\tilde{V})) = 0 \) and \( \dot{V} = 0 \) at \( V = \tilde{V} \). Then \( D(\tilde{V}) = 0 \), and (3.3) implies \( \dot{V} = u(\dot{w})/\delta \). Similarly, because \( \dot{J}(\tilde{V}) = 0 \), (3.5) implies \( \dot{J}(\tilde{V}) = (y - \dot{w})/\delta \). Because recruiting at \( \dot{w} \) should yield zero net profit, \( q(\tilde{V})J(\tilde{V}) = k \); that is, \( \dot{w} = y - \delta k/q(\tilde{V}) \). If \( q(\tilde{V}) = q \), then the stated expressions for \( \dot{w} \) and \( \dot{J}(\tilde{V}) \) follow. Since \( q < \infty \) by Assumption 2, then \( \dot{w} < y \) and \( \dot{J}(\tilde{V}) > 0 \).

To show \( q(\tilde{V}) = q \), suppose that \( q(\tilde{V}) = q - a \) to the contrary, where \( a > 0 \). Because \( \dot{q}(\tilde{V})J(\tilde{V}) = k > 0 \) and \( J(\tilde{V}) = (y - \dot{w})/\delta \), then \( \dot{w} = y - \delta k/(q - a) \). Consider a firm that deviates from \( \dot{w} \) to \( \dot{w} + \epsilon \), where \( \epsilon > 0 \), which generates a value for a worker as \( \tilde{V} = u(\dot{w} + \epsilon)/\delta \). Because the firm is the only one that offers a wage higher than \( \dot{w} \), the workers who are employed at \( \dot{w} \) will all apply to this firm, which yields \( q(\tilde{V}) = q \). The firm’s expected value of recruiting is \( q(\tilde{V})J(\tilde{V}) = (y - \dot{w} - \epsilon)q/\delta \), which exceeds \( k \) for sufficiently small \( \epsilon > 0 \). This result contradicts the fact that \( \tilde{V} \) is an equilibrium value. Thus, \( q(\tilde{V}) = q \). This completes the proof of Lemma 3.2.

Now, turn to Lemma 5.1. Let \( w(V) \) be an arbitrary function in \( \Omega \). It is easy to verify that \( J_w(V) \) defined by (5.1) is strictly positive, bounded, strictly decreasing and continuously differentiable, with \( J'(V) = -1/u'(w(V)) < 0 \). Because \( w(V) \) is increasing, then \( J'(V) \) is decreasing and \( J(V) \) is (weakly) concave. Moreover, \( J_w(\tilde{V}) = k/q \). Similarly,
Lemma C.1. \( p_w(V) \) defined by (5.2) is bounded and continuous for all \( V \) (including \( V = \bar{V} \)), with \( p_w(\bar{V}) = M(\bar{q}) = 0 \). For all \( V < \bar{V} \), \( p_w(V) \) is differentiable and strictly decreasing because
\[
p'_w(V) = \left( M' \frac{k}{J^2_w} \right) \frac{1}{u'(w(V))} < 0,
\]
where the argument of \( M' \) is \( k/J_w(V) \) and where \( M' < 0 \) under Assumption 2. Moreover, for any given value \( V \),
\[
\frac{d}{dJ_w} \left( M' \frac{k}{J^2_w} \right) = \frac{k}{J^3_w} \left( -\frac{k}{J_w} M'' - 2M' \right) \geq 0,
\]
where the inequality follows from part (iii) of Assumption 2. Because \( J_w(V) \) is decreasing and \( M' < 0 \), the function \( M'k/J_w(V) \) is decreasing in \( V \). Because \( 1/u'(w(V)) \) is increasing in \( V \) and \( M' < 0 \), then \( p'_w(V) \) is decreasing. That is, \( p_w(V) \) is (weakly) concave. QED

C. Proof of Theorem 5.2

The sets \( \Omega \) and \( \Omega' \) are defined prior to Lemma 5.1 and the mapping \( \psi \) is defined by \( w_1(V) = (\psi w)(V) \), where \( w_1 \) is the solution to (5.3). It can be verified that \( \Omega \) is a closed and convex set. Lemmas C.1 and C.2 below state that \( \psi : \Omega \to \Omega' \) is a continuous mapping in the supnorm. Under the assumption that the image of \( \psi \) is compact, the Schauder fixed point theorem implies that \( \psi \) has a fixed point in \( \Omega \), denoted as \( w^* \). Because \( w^*(V) = (\psi w^*)(V) \in \Omega' \), then \( w^*(V) \) is strictly increasing for all \( V < \bar{V} \). This implies that \( J_{w^*}(V) \) and \( p_{w^*}(V) \) are strictly concave, in addition to the properties stated in Lemma 5.1.

**Lemma C.1.** \( \psi : \Omega \to \Omega' \subset \Omega \).

**Proof.** Temporarily denote the left-hand side of (5.3) as \( L(w_1) \) and the right-hand side as \( R(V) \). Recall that \( \bar{w} < y \). Because \( L(w) \) is continuous and strictly decreasing for all \( w < y \), it is invertible for all \( w \in [\bar{w}, \bar{y}] \). Then, \( w_1(V) = L^{-1}(R(V)) \). Pick an arbitrary \( w \in \Omega \). I show that \( w_1 \in \Omega' \). This is done in the following steps.

First, \( w_1(V) \) is continuous because \( J_w(\cdot) \), \( p_w(\cdot) \) and \( F_w(\cdot) \) are all continuous.

Second, \( w_1(V) \) is strictly increasing for all \( V < \bar{V} \); i.e., \( R(V) \) is strictly decreasing. To establish this result, pick arbitrary values \( V_1 \) and \( V_2 \), with \( \bar{V} \leq V_1 < V_2 < \bar{V} \), and let \( F_i = F(V_i) \) with \( i = 1, 2 \). I show that the following (stronger) property holds:
\[
0 < J_w(V_2)S \leq R(V_1) - R(V_2) \leq J_w(V_1)S, \tag{C.1}
\]
where
\[
S \equiv u'(w(V_1)) [\delta + \lambda_1 p_w(F_1)] - u'(w(V_2)) [\delta + \lambda_1 p_w(F_2)] > 0.
\]
Note that \( S > 0 \), indeed, because \( w(V) \) is increasing, \( u'(w) \) is strictly decreasing, \( p_w(F) \) is strictly decreasing and \( F_w(V) \) is strictly increasing. To establish (C.1), note that \( J_w(V) \) is decreasing and concave with derivative \( J'_w(V) = -1/u'(w(V)) < 0 \). Then,
\[
\frac{V_2 - V_1}{u'(w(V_1))} \leq J_w(V_1) - J_w(V_2) \leq \frac{V_2 - V_1}{u'(w(V_2))}.
\]
35
Similarly, because the function \([\delta V - \lambda_1 D_w(V)]\) is increasing and concave with derivative \([\delta + \lambda_1 p_w(F)]\), I have:

\[
\delta + \lambda_1 p_w(F_1) \geq \frac{[\delta V_2 - \lambda_1 D_w(V_2)] - [\delta V_1 - \lambda_1 D_w(V_1)]}{V_2 - V_1} \geq \delta + \lambda_1 p_w(F_2).
\]

Using the first part of the above two results to substitute \(J_w(V_1)\) and \([\delta V_1 - \lambda_1 D_w(V_1)]\) in \(R(V_1)\), I get \(R(V_1) - R(V_2) \geq J_w(V_2)S\). Using the second part of the above two results to substitute \(J_w(V_2)\) and \([\delta V_2 - \lambda_1 D_2(V_2)]\) in \(R(V_2)\), I get \(R(V_1) - R(V_2) \leq J_w(V_1)S\).

Third, \(w_1(V) \in [\bar{w}, \tilde{w}]\) for all \(V\), with \(w_1(\bar{V}) = \bar{w}\). Examine \(w_1(V)\). Because \(w(\bar{V}) = \bar{w}\), then (5.3) implies:

\[
L(w_1(\bar{V})) = R(\bar{V}) = u'(\bar{w})(y - \bar{w}) + u(\bar{w}) = L(\bar{w}).
\]

Because \(L(w)\) is strictly decreasing, the above equation implies \(w_1(\bar{V}) = \bar{w}\). Since \(w_1(V)\) is strictly increasing for \(V < \bar{V}\), then \(w_1(V) < \bar{w}\) for all \(V < \bar{V}\).

Finally, I show \(w_1(V) \geq w\). Since \(L'(w) < 0, w_1(V) \geq w\) if and only if \(L(w) \geq R(V)\). A sufficient condition is \(L(w) \geq R(V)\), because \(R(V)\) is a decreasing function. Note that the following holds:

\[
R(V) = u'(w(V)) [\delta + \lambda_1 p_w(F_w(V))] J_w(V) + \delta V - \lambda_1 D_w(V) < u'(w) [\delta + \lambda_1 p_w(F_w(V))] J_w(V) + u(b)
\]

\[
\leq u'(w) [\delta + \lambda_1 p_w(V)] J_w(V) + u(b)
\]

\[
\leq u'(\bar{w}) [\delta + \lambda_1 p_w(V)] J_w(V) + u(b).
\]

The first inequality follows from the facts that \(w(V) \geq w, V = u(b)/\delta\) and \(D_w(V) > 0\). The second inequality follows from the facts that \(F_w(V) \geq V\) and \(p_w(.)\) is decreasing. To obtain the third inequality, note that \(J_w(V) \leq J_w(V)\) and \(p_w(V) \leq p_w(V)\) for all \(V\). Therefore, a sufficient condition for \(w_1(V) \geq w\) is:

\[
L(w) \geq u'(w) [\delta + \lambda_1 p_w(V)] J_w(V) + u(b).
\]

This condition can be re-arranged as (5.7), which is assumed in the Theorem. This completes the proof of Lemma C.1.

**Lemma C.2.** \(\psi\) is continuous in the supnorm.

**Proof.** To show that the mapping \(\psi\) is continuous in the supnorm, I show that the following holds for all \(w_a, w_b \in \Omega\) and all \(V\):

\[
|(\psi w_a)(V) - (\psi w_b)(V)| \leq A \|w_a - w_b\|, \tag{C.2}
\]

where the norm is the supnorm and \(A > 0\) is a finite constant. Once this is done, then

\[
\|\psi w_a - \psi w_b\| = \sup |(\psi w_a)(V) - (\psi w_b)(V)| \leq A \|w_a - w_b\|,
\]

which implies that \(\psi\) is continuous in the supnorm.
To show (C.2), take arbitrarily \( w_a, w_b \in \Omega \) and \( V \in [\underline{V}, \bar{V}] \). Without loss of generality, assume \( w_a(V) \geq w_b(V) \) for the given value \( V \). Shorten the subscript \( w \) on \( J, p, F \), and \( D \) to \( i \), where \( i = a, b \). Also, denote the right-hand side of (5.3) with \( w = w_i(V) \) as \( R_i(V) \).

Because \( w \geq w > 0 \), Assumption 1 implies that there are positive and finite constants \( \omega_1 \) and \( \omega_2 \) such that \( \omega_1 \leq |u''(w)| \leq \omega_2 \) for all \( w \in [\underline{w}, \bar{w}] \). Then

\[
|L'(w)| = (y - w)|u''| \geq (y - \bar{w}) \omega_1 \equiv A_1.
\]

Note that \( A_1 \) is bounded above 0. Since \( L(w) \) is decreasing, then

\[
|R_a(V) - R_b(V)| = |L(\psi w_a(V)) - L(\psi w_b(V))| \geq A_1 |\psi w_a(V) - \psi w_b(V)|.
\]

I show that \( |R_a(V) - R_b(V)| \leq A_6 \|w_a - w_b\| \) for some positive and finite \( A_6 \). Then, (C.2) holds after defining \( A = A_6/A_1 \).

To establish the desired inequality for \( R \), suppress the given \( V \). I have:

\[
|R_a - R_b| = |\{[u'(w_a) - u'(w_b)] J_a + u'(w_b)(J_a - J_b)\} [\delta + \lambda_1 p_a(F_a)] + \lambda_1 u'(w_b) J_b [p_a(F_a) - p_b(F_b)] - \lambda_1 [D_a - D_b]|
\]

\[
\leq |[u'(w_a) - u'(w_b)] J_a + u'(w_b)[J_a - J_b]| [\delta + \lambda_1 p_a(F_a)] + \lambda_1 u'(w_b) J_b |p_a(F_a) - p_b(F_b)| + \lambda_1 |D_a - D_b|.
\]

I find the bound on each of the absolute values in the above expression.

Because \( u'' < 0 \), then

\[
|u'(w_a) - u'(w_b)| \leq |w_a - w_b| \max\{|u''(w_a)|, |u''(w_b)|\} \leq \omega_2 \|w_a - w_b\|.
\]

By the definition of \( J_w \),

\[
|J_a - J_b| = \left| \int_{\underline{V}}^{\bar{V}} u'(w_a(z)) - u'(w_b(z)) \frac{dz}{w_a(z) - w_b(z)} \right|
\]

\[
\leq \frac{1}{|u'(\bar{w})|} \int_{\underline{V}}^{\bar{V}} |u'(w_a(z)) - u'(w_b(z))| \frac{dz}{w_a(z) - w_b(z)} \leq \frac{\omega_2}{|u'(\bar{w})|} \|w_a - w_b\|.
\]

The coefficient of \( \|w_a - w_b\| \) is bounded because \( u''(\bar{w}) > 0 \) and \( 0 < \omega_2 < \infty \).

To develop bounds on \( |p_a(F_a) - p_b(F_b)| \) and \( |D_a - D_b| \), let \( \varepsilon = \|w_a - w_b\| > 0 \) with loss of generality. (If \( \|w_a - w_b\| = 0 \), then \( w_a = w_b \) for all \( V \), in which case \( |p_a(F_a) - p_b(F_b)| = |D_a - D_b| = \|w_a - w_b\| \); these provide the required bounds.) I examine two cases separately: the case where \( V \) is close to \( \bar{V} \) and the case where \( V \) is away from \( \bar{V} \). The separation is necessary because \( M'(q) \) and \( M''(q) \) might be unbounded at \( q = \tilde{q} \) (i.e., at \( V = \bar{V} \)).

Consider first the case where \( V \) is close to \( \bar{V} \). In this case, \( F_a(V) \) and \( F_b(V) \) are close to \( \bar{V} \). Because \( p_w(V) \) is continuous at \( V = \bar{V} \), and because \( F(V) \) is continuous, then for given \( \varepsilon > 0 \), there exists \( c > 0 \) such that

\[
\bar{V} - V < c \implies |p_i(F_i) - p_i(F_i(\bar{V}))| < \varepsilon/2, \quad \text{for } i \in \{a, b\}.
\]

Because \( F_i(\bar{V}) = \bar{V} \) and \( p_i(\bar{V}) = 0 \), the following holds for \( V > \bar{V} - c \):

\[
|p_a(F_a) - p_b(F_b)| \leq |p_a(F_a)| + |p_b(F_b)| < \varepsilon = \|w_a - w_b\|, \tag{C.5}
\]
\[ |D_a - D_b| \leq |p_a(F_a)| (F_a - V) + |p_b(F_b)| (F_b - V) < (\bar{V} - V) \|w_a - w_b\|. \] (C.6)

For the last inequality, I used the facts that \( |p_i(F_i)| < \varepsilon/2 \) and that \( F_i - V_i \leq \bar{V} - V \). (C.5) and (C.6) provide the required bounds when \( V > \bar{V} - c \).

Now consider the case where \( V \leq \bar{V} - c \), where \( c > 0 \) is constructed above. In this case, \( q < \bar{q} \) and, hence, Assumption 2 implies that \( |M'(q)| \) and \( |M''(q)| \) are bounded for \( q \in [q, \bar{q}] \). Because \( p(V) = M \left( \frac{k}{J(V)} \right) \), then

\[
\frac{dM(k/J)}{dJ} = \left( -\frac{k}{J^2} \right) M' \left( \frac{k}{J} \right),
\]

\[
\frac{d^2M(k/J)}{dJ^2} = \left( \frac{k}{J^3} \right) \left( -\frac{k}{J} M'' - 2M' \right).
\]

These absolute values are bounded above in the current case. Let \( A_2 \) and \( A_3 \) be the upper bounds. Define

\[ A_4 = A_2 \frac{\omega_2(V-V)}{[u'(w)]^2} < \infty. \]

For any \( x \in [V, \bar{V} - c] \),

\[ |p_a(x) - p_b(x)| \leq A_2 |J_a(x) - J_b(x)| \leq A_4 \|w_a - w_b\|, \]

\[
\left| \frac{dM_a}{dJ_a} - \frac{dM_b}{dJ_b} \right| \leq A_3 |J_a - J_b|.
\]

These results lead to the following result:

\[ |p'_a(x) - p'_b(x)| \leq \frac{dM_a/dJ_a - dM_b/dJ_b}{u'(w_a) - u'(w_b)} \]

\[
\leq \frac{dM_a}{dJ_a} \left| \frac{1}{u'(w_a)} - \frac{1}{u'(w_b)} \right| + \frac{1}{u'(w_b)} \left| \frac{dM_a}{dJ_a} - \frac{dM_b}{dJ_b} \right|
\leq \frac{A_2}{w'(w)^2} |u'(w_a) - u'(w_b)| + \frac{A_3}{w'(w)} |J_a - J_b|
\leq \frac{A_2}{w'(w)^2} \|w_a - w_b\| + \frac{A_4 A_3 A_2}{w'(w)} \|w_a - w_b\|.
\]

Suppose first that \( F_a \geq F_b \). If \( p_a(F_a) \geq p_b(F_b) \), then

\[ 0 \leq p_a(F_a) - p_b(F_b) \leq p_a(F_a) - p_b(F_a) \leq A_4 \|w_a - w_b\|. \]

The second inequality comes from the fact that \( p \) is decreasing and the last inequality from the bound on \( |p_a - p_b| \) just derived. If \( p_a(F_a) < p_b(F_b) \), then

\[ 0 < p_b(F_b) - p_a(F_a) = -p'_b(F_b)(F_b - V) + p'_a(F_a)(F_a - V) \]

\[
\leq (F_a - V) [p'_a(F_a) - p'_b(F_b)] \leq (\bar{V} - V) [p'_a(F_a) - p'_b(F_b)]
\leq \left[ 1 + \frac{A_2(V-V)}{A_2 u'(w)} \right] A_4 \|w_a - w_b\|.
\]

The equality follows from the first-order condition for \( F \), the second inequality from the supposition \( F_a \geq F_b \), the third inequality from the facts that \( p' \) is a decreasing function
and that \( F_a - V \leq \bar{V} - V \), and the last inequality from the bound on \(|p'_a - p'_b|\). Thus, if \( F_a \geq F_b \), then
\[
|p_a(F_a) - p_b(F_b)| \leq \left[ 1 + \frac{A_3(\bar{V} - V)}{A_2u'(\bar{w})} \right] A_4 \|w_a - w_b\|. \tag{C.7}
\]

Suppose now that \( F_a < F_b \). By switching the roles of \( F_a \) and \( F_b \), it can be shown that (C.7) continues to hold. Thus, (C.7) holds for arbitrary \( F_a(V) \) and \( F_b(V) \) with \( V \leq \bar{V} - c \).

Now let us examine \(|D_a - D_b|\) for the case \( V \leq \bar{V} - c \). If \( D_a \geq D_b \), then
\[
0 \leq D_a - D_b = p_a(F_a)(F_a - V) - p_b(F_b)(F_b - V)
\leq p_a(F_a)(F_a - V) - p_b(F_a)(F_a - V)
= (F_a - V) [p_a(F_a) - p_b(F_a)] \leq (\bar{V} - V)A_4 \|w_a - w_b\|.
\]

The first equality comes from the definition of \( D(V) \), the second inequality from the fact that \( p_b(f)(f - V) \) is maximized at \( f = F_b \), the last inequality from the bound on \(|p_a - p_b|\) derived above and the fact \( F_a - V \leq \bar{V} - V \). The same result holds if \( D_a < D_b \). Thus,
\[
|D_a - D_b| \leq (\bar{V} - V)A_4 \|w_a - w_b\|. \tag{C.8}
\]

Defining \( A_5 = \max\{A_4, 1\} \) and replace \( A_4 \) in (C.7) and (C.8) with \( A_5 \). The resulting bounds on \(|p_a - p_b|\) and \(|D_a - D_b|\) apply for both \( V > \bar{V} - c \) and \( V \leq \bar{V} - c \). Substituting these bounds, (C.3) and (C.4), I have:
\[
|R_a - R_b| \leq \left\{ \omega_2 J_a + u'(w_b) \frac{A_4}{A_2} [\delta + \lambda_1 p_a(F_a)] \right\} \|w_a - w_b\|.
\]

Let \( A_6 \) be the maximum value of the coefficient of \( \|w_a - w_b\| \) in the above expression, taken over \( V \in [\bar{V}, \bar{V}] \). Then, \( A_6 \) is bounded above. Setting \( A = A_6/A_1 \) establishes the inequality (C.2), which shows that \( \psi \) is continuous in the supnorm. This completes the proof of Lemma C.2 and, hence, of Theorem 5.2. QED

**D. Proof of Corollary 5.3**

To prove Corollary 5.3, suppose that \(|\dot{w}(V(t))| < \infty \) for all \( t \). If \( \bar{V} \neq 0 \), then \( w'(V) = \dot{w}/\bar{V} \) exists and is finite. If \( V = 0 \) at some value \( V_c \), such as \( \bar{V} \), then \( \delta V_c - u(w(V_c)) - \lambda_1 D(V_c) = 0 \). Differentiating this equation with respect to \( V_c \) yields:
\[
w'(V_c) = \frac{\delta + \lambda_1 p(F(V_c))}{u'(w(V_c))} \in (0, \infty). \tag{D.1}
\]

That is, \( w(V) \) is differentiable at \( V_c \). Thus, \( w'(V) \) exists and is finite for all \( V \). From (5.1), (5.2) and Lemma 3.1, one can then verify that \( J''(V) \), \( p''(V) \) and \( F''(V) \) all exist and are finite for all \( V < \bar{V} \).

I still need to show that \( w'(V) > 0 \), \( \bar{V} > 0 \) and \( \dot{J}(V) < 0 \) in the case \( V < \bar{V} \). In this case, \( F(V) < \bar{V} \). Lemma 3.1 implies \( dp(F(V))/dV < 0 \). The right-hand side of (3.11)
is positive and finite, which implies $\dot{w}(V) > 0$. Thus, $w'(V)\dot{V} \in (0, \infty)$ for all $V < \bar{V}$. Because $w(V)$ is strictly increasing for all $V < \bar{V}$ and $\dot{V}$ is bounded (see (3.3)), then $w'(V) \in (0, \infty)$ and $\dot{V} \in (0, \infty)$ for all $V < \bar{V}$. Finally, $\dot{J}(V) = J'(V)\dot{V} \in (0, \infty)$ for all $V < \bar{V}$. This completes the proof of Corollary 5.3. QED

E. Proofs of Theorem 6.1 and Corollary 6.2

First, I derive (6.2). Set $V = \bar{V}$ in (6.1). Because $\dot{V} = 0$ at $V = \bar{V}$, the left-hand side of (6.1) is equal to 0 at $V = \bar{V}$. Moreover, the integral in (6.1) is equal to zero, because $F^{-1}(\bar{V}) = \bar{V}$. Thus, at $V = \bar{V}$, (6.1) yields (6.2).

Second, I show that $G$ is continuous; i.e., $G$ does not have a mass point. Suppose, to the contrary, that $G$ has a mass $m > 0$ at some value $V \in [v_1, \bar{V}]$. Then, $G(V) - G(V - \dot{V}dt) \geq m$ for all $dt > 0$, and so the left-hand side of (6.1) is equal to $\infty$. This is a contradiction, because the right-hand side of (6.1) is bounded.

Third, the density function, $g$, is continuous for all $V$ and obeys (6.3). To establish continuity of $g$, denote the derivative of $G$ from the left-hand side of $V$ as $g(V_-)$. The left-hand side of (6.1) is equal to $g(V_-)\dot{V}$. Because $G$, $F$, $F^{-1}$ and $p(.)$ are continuous, the right-hand side of (6.1) is continuous in $V$. Thus, $g(V_-)\dot{V}$ must be continuous. Because $\dot{V}$ is continuous, $g$ must be continuous. Then, I can express the left-hand side of (6.1) as $g(V)\dot{V}$. After substituting $p(v_1)$ from (6.2), (6.1) becomes (6.3).

Fourth, $g$ is continuously differentiable for all $V \neq v_2$. To see this, note that $F$, $F^{-1}$ and $p(.)$ are continuously differentiable. Since $g$ is continuous, $G$ is continuously differentiable, and so the right-hand side of (6.3) is continuously differentiable for all $V \neq v_2$. Thus, the left-hand side of the equation, $g(V)\dot{V}$, must be continuously differentiable for all $V \neq v_2$. Because $\dot{V}$ is continuously differentiable, $g(V)$ is continuously differentiable for all $V \neq v$.

Fifth, I derive (6.4). For $V \in (v_1, v_2)$, $F^{-1}(V) < v_1$, and so (6.3) becomes:

$$g_1(V)\dot{V} = \delta [1 - G(V)] - \lambda_1 \int_{v_1}^{V} p(F(z))g_1(z)dz. \quad (E.1)$$

Because $G(v_1) = 0$ by continuity of $G$, taking the limit $V \downarrow v_1$ in (E.1) leads to $g_1(v_1)\dot{v}_1 = \delta$. Differentiating (E.1) with respect to $V$ and using (3.8), I get:

$$\frac{d}{dV} \left[ \frac{\dot{V}g_1(V)}{\gamma(V, v_1)} \right] = 0. \quad (E.2)$$

Note that $g_1(v_1)\dot{v}_1 = \delta$ and $\gamma(V, v_1)/\gamma(z, v_1) = \gamma(V, z)$. Integrating (E.2) from $v_1$ to $V$ yields (6.4). Since $g$ is continuous, taking the limit $V \uparrow v_2$ in (6.4) gives $g(v_2)$.

Finally, I derive (6.5) by examining the case $V \in [v_j, v_{j+1}]$, where $j \geq 2$. In this case, $F^{-1}(V) \geq v_1$, and so (6.3) becomes:

$$g_j(V)\dot{V} = \delta [1 - G(V)] - \lambda_1 \int_{F^{-1}(V)}^{v_j} p(F(z))g_{j-1}(z)dz - \lambda_1 \int_{v_j}^{V} p(F(z))g_j(z)dz. \quad (E.3)$$
On the right-hand side of the equation, I have separated the two groups of applicants who successfully obtained jobs with values above \( V \): one coming from \((F^{-1}(V), v_j)\) and the other from \([v_j, V]\). Differentiating (6.3) with respect to \( V \) and rewriting the result, I obtain the following equation similar to (E.2):

\[
\frac{d}{dV} \left[ \dot{V} g_j(V) \right] = \frac{\lambda_1 p(V)}{\gamma(V,v_1)} g_{j-1} \left( F^{-1}(V) \right) \frac{dF^{-1}(V)}{dV}. \tag{E.4}
\]

Integrating this equation from \( v_j \) to \( V \) yields (6.5). Because \( g \) is continuous, then \( g_j(v_j) = \lim_{V \uparrow v_j} g_{j-1}(V) \), all \( j \). This completes the proof of Theorem 6.1.

Now, turn to Corollary 6.2. Because \( g_w(w) = g(V) / w'(V) \) and \( 0 < w'(V) < \infty \) (see Lemma 5.3), the property of \( g(V) \) stated in the corollary implies the property of \( g_w(\dot{w}) \). Thus, it suffices to establish the property of \( g(V) \). For this purpose, examine (6.3) at \( \dot{V} = V - \varepsilon \), where \( \varepsilon > 0 \) is a sufficiently small number. Setting \( \dot{V} = \dot{V} \) in (6.3), dividing the equation by \( \varepsilon \) and taking the limit \( \varepsilon \to 0 \), I get:

\[
g(V) \lim_{\varepsilon \to 0} \left( \frac{1}{\varepsilon} \dot{V}|_{V=\dot{V}} \right) = \delta g(V) - \lambda_1 p(V) g(V) \lim_{\varepsilon \to 0} \left\{ \frac{1}{\varepsilon} \left[ \dot{V} - F^{-1}(\dot{V}) \right] \right\}. \tag{E.5}
\]

To obtain the last term in the above equation, I used the Intermediate Value theorem to compute the integral in (6.3). Compute:

\[
\frac{1}{\varepsilon} \dot{V}|_{V=\dot{V}} = \frac{1}{\varepsilon} \left[ \dot{V}|_{V=\dot{V}} - \dot{V}|_{V=\dot{V}} \right] \to - \left( \frac{d\dot{V}}{dV} \right)_{V=\dot{V}} = - \left[ \delta - u'(\dot{w}) \right] = 0.
\]

The first equality comes from the fact that \( \dot{V} = 0 \) at \( V \), the second equality from substituting (3.3) and \( p(\dot{V}) = 0 \), and the last equality from substituting \( u'(\dot{w}) \) from (D.1). Moreover, compute:

\[
\lim_{\varepsilon \to 0} \left\{ \frac{1}{\varepsilon} \left[ \dot{V} - F^{-1}(\dot{V}) \right] \right\} = \lim_{\varepsilon \to 0} \left\{ \frac{1}{\varepsilon} \left[ F^{-1}(\dot{V}) - F^{-1}(\dot{V}) \right] - 1 \right\} = \frac{1}{F'(V)} - 1.
\]

The first equality comes from substituting \( \dot{V} = V - \varepsilon \) and \( \dot{V} = F(\dot{V}) \), while the second equality comes from the fact that \( dF^{-1} / dV = 1 / F' \). Substituting these results and the fact \( p(\dot{V}) = 0 \) into (E.5), I obtain: \( g(\dot{V}) \left[ \delta - \lambda_1 p(\dot{V}) / F'(\dot{V}) \right] = 0 \). Thus, \( g(\dot{V}) = 0 \) if and only if \( \left| \delta - \lambda_1 p(\dot{V}) / F'(\dot{V}) \right| \neq 0 \). Under this condition, continuity of \( g(V) \) implies that \( g'(V) < 0 \) when \( V \) is close to \( \dot{V} \).

To check when the two matching functions in Example 2.1 satisfy the required condition for \( g(\dot{V}) = 0 \), rewrite the condition as \( \left| \delta - \lambda_1 p(\dot{V}) \frac{dF^{-1}(V)}{dV}|_{V=\dot{V}} \right| \neq 0 \). Rewriting (3.2) as \( F^{-1}(V) = V + p(V)/p'(V) \) and differentiating it to obtain \( dF^{-1}(V) / dV \). Using \( p(V) = M(q(V)) \) and \( q(V) = k / J(V) \) to compute \( p'(V) \) and \( p''(V) \), I derive:

\[
p(V) \frac{dF^{-1}(V)}{dV} = 2M + \frac{M^2}{qM'/J^2} - \frac{M^2}{qM'} \left( \frac{qM''}{M'} \right) + 2.
\]
With the CES matching function,

\[ \frac{M}{qM'} = 1 - \frac{q^\rho}{1 - \alpha}, \quad \frac{M}{qM'} \left( \frac{qM''}{M'} + 2 \right) = 2 - \frac{1 - \rho}{1 - \alpha} q^\rho. \]

Because Assumption 2 requires \( \rho < 0 \), then \( \bar{q} = (1 - \alpha)^{1/\rho} \). In this case, the limit \( V \to \bar{V} \)
(i.e., \( q \to \bar{q} \)) implies: \( M \to 0, M' \to -\infty, \frac{M}{qM'} \to 0 \), and \( -\frac{M}{qM'} \left( \frac{qM''}{M'} + 2 \right) \to 1 + \rho \). Because \( J, J' \) and \( J'' \) are all bounded, \( p(V) dF^{-1}(V) / dV \to 0 \) as \( V \to \bar{V} \). Then, the condition required for \( g(\bar{V}) = 0 \) is \( \delta \neq 0 \), which is satisfied.

With the urn-ball matching function,

\[ \frac{M^2}{qM'} = \frac{(\bar{q} - q) \ln (1 - q/\bar{q})}{(\bar{q} - q) \ln (1 - q/\bar{q}) + q} M, \]

\[ \frac{M^2}{qM'} \left( \frac{qM''}{M'} + 2 \right) = -q \left\{ \left[ \frac{q}{(\bar{q} - q) \ln (1 - q/\bar{q}) + q} \right]^2 + \frac{2}{(\bar{q} - q) \ln (1 - q/\bar{q})} \right\}. \]

The limit \( V \to \bar{V} \) implies \( q \to \bar{q} \) and \( M \to 0 \). Then, the above two expressions approach 0 and \(-\bar{q}\), respectively, and so \( p(V) dF^{-1}(V) / dV \to \lambda_1 \bar{q} \). In this example, the condition required for \( g(\bar{V}) = 0 \) is \( \bar{q} \neq \delta / \lambda_1 \). QED
References


F. Supplementary Appendix

Concavity of the Firm’s Objective Function

I show that the firm’s program, \( (P) \), is concave in a neighborhood of the optimal wage contract. For any given value \( V(0) \), the firm’s problem is to choose a sequence, \( \{w(m)\}_{m=V(0)}^{V} \) to maximize:

\[
J(V(0)) = \int_{V(0)}^{V} \frac{[y - w(m)] \gamma(m, v_{1})}{\delta m - u(w(m)) - \lambda_{1}D(m)} dm.
\]

Note that the function \( J \) satisfies:

\[
\frac{dJ(V)}{dV} = \frac{[\delta + \lambda_{1}p(F(V))] J(V) - y + w(V)}{\delta V - u(w(V)) - \lambda_{1}D(V)}.
\]

To show that the firm’s objective function is concave, I compute Gâteaux derivatives of the objective function in a neighborhood of the optimal contract and show that the second-order derivative is negative. Let the optimal contract be \( \{w(m)\}_{m=V(0)}^{V} \). Take an arbitrary path \( \{\hat{w}(m)\}_{m=V(0)}^{V} \) that is continuous and bounded, with \( \hat{w}(V) = 0 \). I can express the contracts in the neighborhood of the optimal contract as

\[
w(m, c) = w(m) + c\hat{w}(m), \text{ for } m \in [V(0), V],
\]

where \( c \) is the size of the neighborhood. This alternative contract generates a new path of \( \gamma \), denoted as \( \gamma(V, V(0), c) \), and a new value of the firm, denoted as \( J(V(0), c) \). The restriction \( \hat{w}(V) = 0 \) implies that the alternative path continues to satisfy \( \dot{V} = 0 \) at \( V = \bar{V} \), which implies \( \gamma(V, V(0), c) = 0 \). The first- and second-order Gâteaux derivatives of \( J \) are:

\[
J_{c}(V, c) \equiv \frac{dJ(V, c)}{dc}, \quad J_{cc}(V, c) \equiv \frac{d^{2}J(V, c)}{dc^{2}}.
\]

I show that \( J_{cc}(V, 0) < 0 \) for all \( V \in [V(0), \bar{V}] \). Thus, there exists \( c_{0} > 0 \) such that \( J_{cc}(V, c) < 0 \) for all \( c \leq c_{0} \). In this sense, the firm’s problem is locally concave.

To compute \( J_{c} \) and \( J_{cc} \), note that \( J(V, c) \) satisfies the following equation:

\[
\frac{dJ(V, c)}{dV} = \frac{[\delta + \lambda_{1}p(F(V))] J(V, c) - y + w(V, c)}{\delta V - u(w(V, c)) - \lambda_{1}D(V)}, \text{ all } V \geq V(0).
\]

Differentiating this equation with respect to \( c \), I get:

\[
\frac{dJ_{c}(V, c)}{dV} = \frac{[\delta + \lambda_{1}p(F(V))] J_{c}(V, c) + \hat{w}(V) [1 + u'(w(V, c))dJ(V, c)/dV]}{\delta V - u(w(V, c)) - \lambda_{1}D(V)}.
\]

Substitute the term \( [\delta + \lambda_{1}p(F(V))] \) using the following fact:

\[
\frac{d\gamma(V, V(0), c)/dV}{\gamma(V, V(0), c)} = -\frac{\delta + \lambda_{1}p(F(V))}{\delta V - u(w(V, c)) - \lambda_{1}D(V)}.
\]
I can rewrite (F.1) as follows:

$$\frac{d}{dV} \left[ \gamma(V, V(0), c) J_{c}(V, c) \right] = \gamma(V, V(0), c) \dot{w}(V) \left[ 1 + u'(w(V, c)) \frac{dJ(V, c)}{dV} \right] \frac{\delta m - u(w(m, c)) - \lambda_1 D(m)}{\delta V - u(w(V, c)) - \lambda_1 D(V)}.$$

Integrate this equation over $V$. Using the facts $\gamma(\bar{V}, V(0), c) = 0$ and $\frac{\gamma(m, V(0), c)}{\gamma(V, V(0), c)} = \gamma(m, V, c)$, I derive the following result for all $V$:

$$J_{c}(V, c) = - \int_{V}^{\bar{V}} \dot{w}(m) \gamma(m, V, c) \frac{1 + u'(w(m, c)) \frac{dJ(m, c)}{dm}}{\delta m - u(w(m, c)) - \lambda_1 D(m)} dm.$$

Because the contract $\{w(V)\}$ is supposed to be optimal, $J_{c}(V, 0) = 0$. This requirement is met for an arbitrary $\dot{w}$ if and only if the above integrand at $c = 0$ is zero for all $m \in [V(0), \bar{V})$. That is,

$$- \frac{dJ(V, 0)}{dV} = \frac{1}{u'(w(V))}, \text{ all } V \in [V(0), \bar{V}). \quad \text{(F.3)}$$

This result reproduces the optimality condition, (3.9). Evaluate (F.1) at $c = 0$ and substituting (F.3), I get $dJ_{c}(V, 0)/dV = 0$ for all $V \in [V(0), \bar{V})$.

To compute $J_{cc}$, differentiate (F.1) with respect to $c$. Using (F.2) to substitute $[\delta + \lambda_1 p(F(V))])$ in the result yields:

$$\frac{d}{dV} \left[ \gamma(V, V(0), c) J_{cc}(V, c) \right] = \gamma(V, V(0), c) \dot{w}(V) \times \frac{u''(w(V, c)) \dot{w}(V) \frac{dJ(V, c)}{dV} + 2u'(w(V, c)) \frac{dJ(V, c)}{dV}}{\delta m - u(w(m, c)) - \lambda_1 D(m)}.$$

Evaluate this equation at $c = 0$. Using (F.3) and the result $dJ_{c}(V, 0)/dV = 0$, I obtain:

$$\frac{d}{dV} \left[ \gamma(V, V(0)) J_{cc}(V, 0) \right] = - \frac{\gamma(V, V(0)) [\dot{w}(V)]^2}{\delta V - u(w(V)) - \lambda_1 D(V)} \left[ \frac{u''(w(V))}{u'(w(V))} \right].$$

where I abbreviated $\gamma(V, V(0), 0)$ as $\gamma(V, V(0))$. Integrating the above equation over $V$ yields the following result for all $V \in [V(0), \bar{V}]$:

$$J_{cc}(V, 0) = \int_{V}^{\bar{V}} \frac{\gamma(m, V) [\dot{w}(m)]^2}{\delta m - u(w(m)) - \lambda_1 D(m)} \left[ \frac{u''(w(m))}{u'(w(m))} \right] dm.$$

Because the optimal path satisfies $\delta m - u(w(m)) - \lambda_1 D(m) > 0$ for all $m < \bar{V}$, then the assumptions on $u$ imply the desired result, $J_{cc}(V, 0) < 0$, for all $V \in [V(0), \bar{V})$. 

2