Dynamic Contracts with Worker Mobility
via Directed On-the-Job Search

Kunio Tsuyuhara

This Version: December, 2014

Abstract

This study proposes a model with dynamic incentive contracts and on-the-job search in a frictional labor market. The optimal long-term contract exhibits an increasing wage-tenure profile. With increasing wages, worker effort also increases with tenure. These two features imply that the probabilities of both voluntary and involuntary job separation decrease with both job-tenure and the duration of employment. Given these results, workers experience differing labor market transitions—between employment, unemployment, and across different employers—and the equilibrium generates endogenous heterogeneity among ex-ante homogeneous workers.

Keywords: Dynamic Contracts, Moral Hazard, Directed On-the-Job Search

JEL Codes: D8, E24, J3, J6.

1This paper is based on Chapter 2 of my Ph.D. thesis submitted to the University of Toronto. I am deeply indebted to Shouyong Shi for his extensive help and encouragement in writing this paper. I am grateful to the editor, Guido Menzio, and two anonymous referees for their insightful and detailed suggestions that helped me significantly improve the paper. I also thank Ettore Damiano, Jan Eeckhout, Li, Hao, Gueorgui Kambourov, Philipp Kircher, Nobu Kiyotaki, Matt Mitchell Robert Shimer, Henry Siu, Scott Taylor, Trevor Tombe, Randy Wright, and conference and seminar participants at the University of Calgary, Toronto, Osaka, and Vanderbilt University, Econometric Society (Tokyo and Barcelona), RMM, SAET (Faro), and SED (Istanbul and Cyprus) for helpful comments and suggestions at various stages of this paper. Financial support from Shouyong Shi’s Bank of Canada Fellowship and Canada Research Chair and from Ettore Damiano’s SSHRC Grant is acknowledged. Please address correspondence to: Kunio Tsuyuhara, Department of Economics, University of Calgary, 2500 University Dr. N.W. Calgary, Alberta, Canada E-mail: kunio.tsuyuhara@ucalgary.ca.
1 Introduction

This study proposes a model with dynamic incentive contracts and on-the-job search in a frictional labor market. Employment relationships are generally long-term (Farber, 2010). Nevertheless, workers frequently move between jobs as well (Nagypál, 2008). Given these two important characteristics of the labor market, the main purpose of this study is to provide new insights into how work incentive within a long-term employment relationship interacts with worker mobility in the labor market.

The model in this study has several important features. First, I assume the worker search is a directed process; that is, the firms offer a contract to attract workers, and the workers direct their search to a particular job offer. This assumption enables a clean characterization of the optimal contract. Second, I assume both employed and unemployed workers are allowed to search. The job discontinues if the worker leaves for another job, i.e., voluntary separation (quits). In addition, the job output depends on the employed worker’s unobservable effort, and the job discontinues if the output is “too low,” i.e., involuntary separation (layoff). This endogenous job separation provides an additional work incentive, which evolves over time and interacts with labor mobility.

There are three key results in this study. First, the firm offers a dynamic wage contract that is increasing with tenure on the job. The result shows there are two reasons to offer the increasing wage-tenure contract: one reason is to retain the worker, and the second reason is to provide an incentive to exert effort. The decomposition of wage variation in those two reasons is consistent with the back-loaded wage contract typically found in the on-the-job search models, such as Burdett and Coles (2003) and Shi (2009) and in efficiency wage models, such as Lazear (1981) and Lazear and Moore (1984). Second, the labor market can be divided into two segments. In the competitive segment the wage increases for both retention and incentive reasons, whereas in the monopsonistic segment in which no firms enter to compete for the employed worker, the wage still increases because of the incentive reason. Third, with increasing wages, worker effort also increases with tenure, which decreases the probability of layoffs. Because increasing wage itself decreases the probability of quits over time, the optimal contract implies that the probabilities of both quits and layoffs decrease with tenure. Consequently, workers experience differing labor market transitions between employment, unemployment, and across different employers.

\(^2\)Acemoglu and Shimer (1999a, 1999b) and Moen (1997) are early applications of directed search in the labor market. Delacroix and Shi (2006) develop a model of directed search on-the-job.
and endogenous heterogeneity among ex-ante homogeneous workers emerges.

The theory described in this study is based on two important studies. On the one hand, the dynamic contracting problem in this model is similar to the model of repeated moral hazard in Spear and Srivastava (1987). I follow their recursive contracting approach and solve the firm’s optimal contracting problem in an equilibrium framework. On the other hand, the analysis of dynamic contracts with directed on-the-job search in this study closely follows that of Menzio and Shi (2010), and I prove the existence of a block recursive equilibrium (BRE) where the individual firm’s contracting problem and worker’s job search problem do not depend on the distribution of workers in equilibrium. The main addition of this study to Menzio and Shi (2010) is the moral hazard problem of a worker’s effort. While this study does not consider aggregate and idiosyncratic shocks, the result shows how Menzio and Shi’s (2010) analysis can incorporate a richer contractual environment.

There is currently growing literature on asymmetric information in search theoretic labor market models. Moen and Rosén (2010) and Zaharieva (2010) analyze employed worker unobservable effort and contracting problem in a competitive search framework. This study is different from these studies as it considers dynamic wage contracts and on-the-job search. By neglecting to consider these aspects, aforementioned studies preclude the interaction between the dynamic incentive problem and worker mobility in the labor market. Other recent studies include that of Moen and Rosén (2006) on an efficiency wage contracting problem in a random search framework, and that of Guerrieri, Shimer, and Wright (2009) and Carrillo-Tudela and Kaas (2013) on adverse selection problems.

Relational contracting models show that a long-term relationship acts as a self-enforcing mechanism to maintain incentives. In a closely related study, Board and Meyer-Ter-Vehn (2011) analyze the effect of on-the-job search in a competitive labor market. However, even with on-the-job search, the lack of commitment on the firm’s side in their model implies the optimal contract that pays a constant wage for the duration of the match. Therefore, their model is unable to capture the dynamics of work incentive on the job and the resulting implications on wage, productivity, and job turnover.

Following the original contribution by Shi (2009), Menzio and Shi (2010) establish the notion of BRE with on-the-job search in a stochastic environment. Other applications of the BRE include Gonzalez and Shi (2010) and Menzio, Shi, and Sun (2013).

Burdett and Mortensen (1980) is one of the original theoretical contributions. They extend the theory of job search to include implicit contracts and develop a consistent theory of labor market equilibrium. Cao and Wang (2013) empirically investigate a contracting problem in a dynamic search model.

See especially MacLeod and Malcomson (1998) for related macroeconomic implications.
2 Labor Market with Search Friction and Moral Hazard

2.1 Preferences and Technology

There is a continuum of infinitely–lived workers with a unit measure and a continuum of firms whose measure is determined by competitive entry. All workers and firms are ex-ante homogeneous. Time is discrete and indexed by $t$. Workers cannot save or borrow against their future income, and in each period a worker consumes a wage income of $\omega_t = w_t$ if employed or an unemployment benefit of $\omega_t = b$ if unemployed, where $b$ is constant and independent of the duration of unemployment. If employed, each worker exerts effort, $e_t$, for a project of the firm during each period. Period utility of a worker is $u(\omega_t) - c(e_t)$. The consumption utility $u : \mathbb{R} \to \mathbb{R}$ is a twice continuously differentiable, strictly increasing, and strictly concave function, and its first derivative is bounded; that is, $u'(\omega) \in [u', u'']$ for all $\omega$. The disutility of effort $c : \mathbb{R} \to \mathbb{R}$ is a twice continuously differentiable, strictly increasing, and strictly convex function, and its first derivative is bounded; that is, $c'(e) \in (0, c']$ for all $e$. Each worker maximizes the expected lifetime sum of utilities that are discounted at rate $\beta \in (0, 1)$.

Each entering firm is endowed with a job and hires a worker for the job. A job consists of a series of projects, one of which is executed in each period. The duration of a job is stochastic for the following reason. The project output is either $y$, called success, with probability $r(e)$ or 0, called failure, with probability $1 - r(e)$, depending on the employed worker’s effort. The probability of success $r : \mathbb{R} \to [0, 1]$ is a twice continuously differentiable, strictly increasing, and strictly concave function, and its first derivative is bounded; that is, $r'(e) \in (0, r]$ for all $e$. If the project succeeds, the job continues in the next period, while the match separates and the worker becomes unemployed if the project fails.\footnote{This is a physical assumption, not the firm’s decision, in the same spirit as an exogenous separation typically assumed in the literature.} The expected period profit of a firm is $r(e)y - w$. Each firm maximizes the expected sum of profits that are discounted at the rate $\beta$.

2.2 Job Search and Employment Contracts

A firm and a worker find a match in the labor market. Both unemployed and employed workers have the opportunity to search for a job. The opportunity to search arises with probabilities $\lambda_u \in [0, 1]$ for unemployed workers and $\lambda_e \in [0, 1]$ for employed workers.
An entering firm posts a vacancy and offers a contract at a flow cost \( k > 0 \). An employment contract specifies the wage the firm pays and the level of effort the worker exerts in each period as long as the job continues. I assume firms commit to pre-specified sequences \( \{w_t, e_t\}_{t \geq 0} \) and that once the employment relationship begins, the firm cannot adjust these sequences depending on the project outcomes or make a counteroffer when its worker’s on-the-job search generates an outside offer.

A match discontinues in the following two cases. First, if the project succeeds and if the worker finds a new job via on-the-job search, the worker leaves the current job (quit), and the new match begins in the next period. Second, if the project fails, the worker loses the current job (layoff). Newly displaced workers cannot search and will stay unemployed for one period. In equilibrium, a contract endogenously determines probabilities of those voluntary and involuntary job separation. Given those implied probabilities, a worker calculates the \textit{value} of the contract; that is, the expected lifetime utility derived from the contract, denoted by \( x \). Similarly, a worker evaluates the value of unemployment, \( U \), based on the unemployment benefit \( b \) and the probability of finding a job.

The labor market consists of a continuum of submarkets, each of which offers contracts with same value and is indexed by \( x \). The entire labor market is denoted by \( X = [\underline{x}, \bar{x}] \), where \( \underline{x} < u(b)/(1 - \beta) \) and \( \bar{x} > u(y)/(1 - \beta) \). As unemployed workers find a job with positive probability, it is clear that in equilibrium \( U \in X \).

Let the value of contract \( x \) and the value of unemployment \( U \) represent the worker’s employment state in some period, and \( G_t \) denote the distribution of workers over \( X \) in period \( t \). Then, the fraction of unemployed workers in period \( t \) is identified by \( u_t = G_t(U) \).

I denote the evolution of \( G_t \) generically with an operator \( \Psi \) such that \( G_{t+1} = \Psi(G_t) \), where \( \Psi \) is endogenously determined according to worker job search and firm contracting decision.

I assume the worker job search is a directed process. That is, workers observe all the contracts in \( X \) and choose which submarket to visit. The ratio of vacant firms to searching workers in submarket \( x \) is denoted by \( \theta(x) \) and is referred to as the tightness of submarket \( x \). If a worker visits submarket \( x \), the employment probability is given by \( p(\theta(x)) \), where \( p : \mathbb{R}_+ \rightarrow [0, 1] \) is twice continuously differentiable, strictly increasing, and strictly concave, and satisfies \( p(0) = 0, p'(0) < \infty \). On the other hand, if a vacant firm enters a submarket \( x \), the hiring probability is given by \( q(\theta(x)) \), where \( q : \mathbb{R}_+ \rightarrow [0, 1] \) is twice continuously differentiable, strictly decreasing, and strictly convex, and satisfies \( \theta^{-1}p(\theta) = q(\theta), q(0) = 1 \). In addition, I assume that \( p(q^{-1}(\cdot)) \) is concave.
3 Equilibrium

The equilibrium of this economy consists of worker optimal job search decision, firm optimal contract design, and consistent market tightness. The description of the firm contracting problem follows the recursive contracting approach in Spear and Srivastava (1987), where history dependence can be summarized by the contract value as a state variable. That is, for a contract value $V$ promised to the worker at some time $t$, a function $J(V)$ yields the current payoff to the firm and a function $W(V)$ determines the value promised to the worker in the next period. The other equilibrium objects are described as a function directly of $V$ or indirectly through $W$. The construction of equilibrium closely follows Menzio and Shi (2010), and I define a BRE. The main addition to Menzio and Shi (2010) is the moral hazard problem of worker effort in a firm’s contracting problem, and this study extends the feature of block recursivity into a richer contractual environment.

3.1 Worker Job Search Problem

Given the opportunity to search, workers choose which submarket to visit, and as in standard labor search models, the optimal search strategy depends on the worker’s reservation value. For employed workers, because they can stay in the current job if the search is unsuccessful, the value of the remaining section of contract, i.e., $\{w_s, e_s\}_{s \geq t}$ for some $t$, is the reservation value for search. Following the terminology of the dynamic contracting literature, I call it continuation value of contract.\(^8\) For unemployed workers, the reservation value for search is the value of unemployment $U$, which is described in detail below.

Suppose, in general, the reservation value for search is $W \in X$. Given this value, if a worker visits submarket $x$, the worker succeeds in finding a job with probability $p(\theta(x))$. If unsuccessful, the worker stays in the current job. Therefore, the worker chooses $x$ to maximize the expected value of search: $p(\theta(x))x + (1 - p(\theta(x)))W$. I define the maximized expected value of job search given $W$ as follows:

$$D(W) = \max_{x \in X} p(\theta(x))(x - W).$$ (1)

For employed workers, $D(W)$ is the value of an outside option of an on-the-job search that is realized if they find an opportunity to search. The worker’s optimal search policy given

\(^8\)It is important to note that this value not only reflects the value of the future wages of the current job, but also reflects gain from future on-the-job search as well as loss from future job separation.
W is denoted by \( m(W) \). I also define the composite function \( \hat{p}(W) = p(\theta(m(W))) \), which is the probability that a worker finds a job in the optimally chosen submarket given \( W \).

Given this optimal search decision, the value of unemployment \( U \) is determined as follows. An unemployed worker receives and consumes the unemployment benefit \( b \) at the end of the period. In the next period, the worker finds an opportunity to search with probability \( \lambda_u \), and it gives the expected lifetime utility \( U + D(U) \). Otherwise, the worker stays unemployed with probability \( 1 - \lambda_u \) for another period, which gives \( U \). Because \( b \) is constant over time and does not depend on the length of the unemployment spell, the value of unemployment \( U \) satisfies the following recursive equation:

\[
U = u(b) + \beta(U + \lambda_u D(U)).
\] (2)

### 3.2 Firm Contracting Problem

Given the value \( V \) currently promised to the worker, the firm chooses the following: i) \( w \), how much wage to pay in this period, ii) \( e \), how much effort to induce, and iii) \( W \), how much continuation value to promise to the worker in the next period conditional on success of the project. In addition, firms are allowed to randomize these choices; that is, the firm offers a lottery of subcontracts \( (w_i, e_i, W_i)_{i=1,2} \).\footnote{As shown later, qualitative properties of the optimal contract do not depend on this randomization. The purpose of introducing the lottery is the potential nonconvexity of the constraint set and resulting nonconcavity of the firm’s objective function where the value can be improved by randomizing the control variables. Because of Lemma 1 and 2, I can apply the proof in Menzio and Shi (2010) that a two-point lottery is sufficient for guaranteeing the concavity.} Let \( \xi = ((w_i, e_i, W_i, \pi_i)_{i=1,2}) \), where \( \pi_i \) is the probability that the outcome of the lottery is \( (w_i, e_i, W_i) \). Therefore, the firm’s maximized value function \( J(V) \) satisfies the following functional equation:

\[
J(V) = \max_{\xi} \sum_{i=1,2} \pi_i \left( r(e_i)y - w_i + \beta r(e_i)(1 - \lambda_e \hat{p}(W_i))J(W_i) \right),
\] (3)

subject to

\[
V = \sum_{i=1,2} \pi_i \left\{ u(w_i) - c(e_i) + \beta [r(e_i)(W_i + \lambda_e D(W_i)) + (1 - r(e_i))U] \right\},
\] (4)

\[
e_i \in \arg\max_{e \in \mathbb{R}} \left( -c(e) + \beta (r(e)(W_i + \lambda_e D(W_i)) + (1 - r(e))U) \right), \ i = 1, 2, \text{ and } (5)
\]

\[
W_i \in \{ X : J(W_i) \geq 0 \}, \ i = 1, 2,
\] (6)

where \( \pi_i \in [0,1] \) for \( i = 1, 2 \), and \( \pi_1 + \pi_2 = 1 \).
The firm’s choice is subject to the following: (i) a promise-keeping constraint, (4), which requires $\xi$ to provide the worker with the lifetime utility $V$, (ii) an incentive compatibility constraint, (5), which requires the contract induces the worker to voluntarily exerts the desired level of effort, and (iii) an individual rationality constraint, (6), which requires the firm does not choose a continuation value that lead to a negative profit (value).

Note that the wage does not affect the worker’s incentive in the current period, but the continuation value $W_i$ does. Because the outcome of the lottery is realized before the worker exerts effort, both subcontracts need to meet the incentive compatibility constraint. Moreover, firms cannot operate with negative profit in equilibrium. Therefore, the last constraint prevents the firm from making an empty promise for inducing higher current effort.

### 3.3 Free Entry Condition and Market Tightness

An entering firm chooses what contract to offer and, equivalently, which submarket to enter to find a worker. The firm’s expected profit of opening a vacancy in submarket $x$ is given by $q(\theta(x))J(x)$, the product between the hiring probability and the value of finding a worker. When choosing which submarket to enter, firms take the market tightness $\theta(x)$ for each submarket as given. For a given $\theta(x)$, if the expected profit is strictly smaller than the cost $k$ of opening a vacancy, then no firms enter $x$. If, on the other hand, the expected profit is strictly greater than $k$, firms create infinitely many vacancies in $x$. Therefore, in any submarket $x$ that is visited by a positive number of workers, $\theta(x)$ is consistent with the firm’s profit maximization if

$$q(\theta(x))J(x) - k \leq 0,$$

and $\theta(x) \geq 0$, with complementary slackness.

The above inequality, which holds with equality for any active submarkets, provides an equilibrium condition that relates two endogenous functions, $J(x)$ and $\theta(x)$. It is important to note that directed search enables the model to close this way and that the way equilibrium relates these functions is the key to characterizing the BRE.
3.4 Block Recursive Equilibrium

Definition 1. A Block Recursive Equilibrium (BRE) consists of the set of individual objects \( \{D^*, m^*, U^*, J^*, \xi^*, \theta^*\} \) and the operator \( \Psi^* \) such that

1. the value of job search \( D^* \) and the optimal search policy \( m^* \) satisfy equation (1),
2. the value of unemployment \( U^* \) satisfies equation (2),
3. the value of firm \( J^* \) and an optimal contract policy \( \xi^* \) satisfy equation (3),
4. the market tightness \( \theta^* \) satisfies condition (7),
5. \( \Psi^* \) is derived from the optimal policy functions, \( \xi^* \) and \( m^* \), and
6. the individual objects \( \{D^*, m^*, U^*, J^*, \xi^*, \theta^*\} \) are independent of the distribution of workers \( G_t^* \) for all \( t \).

The most important aspect of this equilibrium is that individual decisions are independent of the distribution of workers. This property is particularly useful for analyzing this model. First, I do not need to include the distribution of workers, which is a large dimensional object, as a state variable when solving each individual problem. That is, I can solve the optimal contracting problem as a representative firm’s problem without referring to the distribution even if the problem involves an on-the-job search. After solving each individual problem, I can compute worker transition over the contract values using the optimal policy functions \( m^* \) and \( \xi^* \). This transition process generates a stochastic process \( \Psi \). More importantly, the stochastic process is stationary because \( m^* \) and \( \xi^* \) do not depend on \( G_t \). This feature significantly reduces the computational burden.

The following proposition extends the existence of a BRE into an environment with moral hazard.

Proposition 1. A Block Recursive Equilibrium exists in this model.

See Appendix A for the detailed proof. Here I give an outline of the proof. The goal of this proof is to show that a solution to the functional equation (3) exists. To begin, I define a class of real-valued functions \( J(X) \) such that any function \( J \in J \) is (i) strictly decreasing and bi-Lipschitz continuous with respect to \( V \), (ii) bounded both from below and above, and (iii) concave.\(^{10}\)

Then, with an arbitrary function \( J \in J \), I compute each equilibrium object from the respective equilibrium condition. First, for any \( J \in J \), the competitive entry condition

\(^{10}\)Menzio and Shi (2010) show that this class of functions constitutes a non-empty, bounded, closed, and a convex subset of the space of bounded, continuous functions on \( X \) with the sup norm.
(7) with equality gives \( \theta \), the market tightness. Given the tightness function, worker job search problem gives the value of search \( D \) and the policy function \( m \), which in turn give \( \hat{p} \) and \( U \). These individual objects are all independent of \( G \) by construction.

Finally, I incorporate these objects into the right-hand side of the functional equation (3) to define an operator \( T \) on \( J(X) \):

\[
(TJ)(V) = \max_{\xi} \sum_{i=1,2} \pi_i (r(e_i)y - w_i + \beta r(e_i)(1 - \lambda \hat{p}(W_i)) J(W_i)). \tag{8}
\]

For the same reasons mentioned above, the constraints for the firm contracting problem are also independent of \( G \). Therefore, the updated value function does not depend on the distribution of workers. In the proof, I show that there exists a fixed point of this operator; that is, a function \( J^* \in J \) such that \( TJ^* = J^* \). Let \( \{D^*, m^*, U^*, \xi^*, \theta^*\} \) be the respective objects derived from such \( J^* \). Then, by construction, \( \{D^*, m^*, U^*, J^*, \xi^*, \theta^*\} \) are independent of the distribution of workers and constitute a BRE.

4 Characterization of the Optimal Contract

**Lemma 1.** Given a continuation value \( W \), there is a unique level of incentive compatible effort, described as a Lipschitz continuous function \( e(W) \). Moreover, \( e(W) \) is differentiable and increasing in \( W \) almost everywhere and its derivative is bounded.

**Proof:** The objective function (5) is a strictly concave function with respect to \( e \) under the assumptions of \( c(\cdot) \) and \( r(\cdot) \). Therefore, the first order condition of the problem (5) defines a generalized equation

\[
f(e, W) \equiv -c'(e) + \beta r'(e) \Omega(W) = 0,
\]

where \( \Omega(W) = W + \lambda_e D(W) - U \). Consider a pair \( (e^*, W^*) \) such that \( f(e^*, W^*) = 0 \). By the assumptions that \( c(\cdot) \) and \( r(\cdot) \) are twice continuously differentiable, (i) \( f \) is differentiable with respect to \( e \), and both \( f(e, W) \) and its derivative with respect to \( e \), \( D_e f(e, W) \) depend continuously on \( (e, W) \), and (ii) \( D_e f(e, W) \) is clearly invertible. Then, by Robinson’s implicit function theorem (Theorem 1.6 of Dontchev and Rochafellar, 2009), there is a unique solution \( e \) for \( e^* \) at \( W^* \) denoted by \( e^* = e(W^*) \).

Moreover, by Corollary 1.8 of Dontchev and Rochafellar (2009), the solution function \( e(\cdot) \) is Lipschitz continuous. Hence, \( e(\cdot) \) is differentiable almost everywhere around \( W^* \).
At any point of differentiability, the derivative of \(e(W)\) with respect to \(W\) is given by

\[
\frac{\partial e(W)}{\partial W} \bigg|_{e^*,W^*} = - \left( \frac{\partial f(e,W)}{\partial e} \bigg|_{e^*,W^*} \right)^{-1} \left( \frac{\partial f(e,W)}{\partial W} \bigg|_{e^*,W^*} \right) \\
= (e''(e^*) - \beta r''(e^*)\Omega(W^*))^{-1} \beta r'(e^*)(1 - \lambda e\hat{p}(W^*)) \\
= \Delta(e^*,W^*)(1 - \lambda e\hat{p}(W^*)),
\]

where \(\Delta(e,W) \equiv \frac{\beta r'(e)}{e''(e) - \beta r''(e)\Omega(W)}\). The second equality uses the fact that \(\Omega'(W) = 1 - \lambda e\hat{p}(W)\), which in turn applies the envelope theorem to the almost everywhere differentiable function \(D(W)\). Because \(c(\cdot)\) is strictly convex and \(r(\cdot)\) is strictly concave with bounded derivative, \(\Delta(e^*,W^*)\) is positive and bounded. This implies the desired results. \(\square\)

This lemma implies that by choosing the continuation value \(W\), the firm implicitly chooses the worker’s effort and the probability of project success in the current period. It also implies the firm can extract higher effort in the current period by promising a higher continuation value for the next period. This result highlights the source of work incentive even if the current wage does not depend on the project outcome in the current period. That is, if the job does not continue, the worker loses both the continuation value of the current job and the value of on-the-job search. That \(\Omega(W)\) is increasing in \(W\) implies both of these values are increasing in \(W\). Therefore, workers with a higher \(W\) have higher stakes in keeping the job and thus have a larger incentive to exert effort.

**Lemma 2.** Under the optimal contract, the current period wage is independent of the realization of lottery, i.e., \(w_1 = w_2\) in every period.

*Proof:* Let \(\eta(V)\) be the Lagrange multiplier for the promise-keeping constraint. The first order condition of the maximization problem with respect to \(w_i\) is \(-1 + u'(w_i)\eta(V) = 0\). This implies that \(w_1 = w_2\). \(\square\)

The lottery is necessary for the optimality when the constraint set is not convex, and thus the objective function is not concave. Lemma 1 implies the choice of \(W_i\) uniquely determines \(e_i\). Therefore, the lottery, if used at all, essentially randomizes over \(\{w_i, W_i\}_{i=1,2}\). However, the current wage does not affect a worker’s incentive and thus the constraint set. Therefore, randomizing continuation values of the contract is sufficient for this purpose.

---

\(^{11}\)Lemma 4 in the appendix shows \(D(\cdot)\) is a Lipschitz continuous and almost differentiable function.
4.1 Optimal Wage-Tenure Contract

The following proposition characterizes how the continuation value of the contract evolves over time.

**Proposition 2.** Under the optimal contract, continuation values weakly increase with tenure, independent of the realization of lottery; that is, $W_i(V) \geq V$ for $i = 1, 2$. Moreover, the dynamics of the continuation value is characterized by the Euler equation

$$J'(V) - J'(W_i) = -\frac{\lambda e J(W_i)\hat{p}'(W_i)}{1 - \lambda e \hat{p}(W_i)} + \frac{r'(e(W_i))}{r(e(W_i))} \left( \frac{y}{\beta} + (1 - \lambda e \hat{p}(W_i))J(W_i) \right) \Delta(e(W_i), W_i) \tag{9}$$

almost everywhere on $X$.

**Proof:** Using the previous lemmas and strict concavity of $u$, the current period wage that is consistent with constraints (4) is expressed as

$$w(V, \phi) \equiv u^{-1} \left( V - \beta U - \sum_i \pi_i \left[ -c(e(W_i)) + \beta r(e(W_i))[W_i + \lambda e D(W_i) - U] \right] \right), \tag{10}$$

where

$$\phi \in \left\{ \{W_i, \pi_i\}_{i=1,2} : W_i \in \{X : J(W_i) \geq 0\}, \pi_i \in [0,1] \text{ for } i = 1, 2, \text{ and } \pi_1 + \pi_2 = 1 \right\}.$$  

Using Lemma 2 and the expression of (10), the functional equation (3) becomes

$$J(V) = \max_{\phi} \left\{ -u^{-1} \left( V - \beta U - \sum_i \pi_i \left[ -c(e(W_i)) + \beta r(e(W_i))[W_i + \lambda e D(W_i) - U] \right] \right) + \sum_i \pi_i r(e(W_i)) \left[ y + \beta (1 - \lambda e \hat{p}(W_i))J(W_i) \right] \right\}. \tag{11}$$

I ignore the feasibility constraints for $W_i$ and $\pi_i$ to characterize the interior solution. As shown in the Appendix, $J(\cdot)$, $D(\cdot)$, and $\hat{p}(\cdot)$ are almost everywhere differentiable functions, and Lemma 1 shows that $e(\cdot)$ is differentiable almost everywhere. Therefore, the objective function in (11) is also almost everywhere differentiable with respect to $W_i$. However, for the possible nondifferentiable points, I cannot use the standard calculus for characterizing the optimality. Hence, I apply the theory of nonsmooth analysis to characterize the solution.\(^{12}\) For my purpose, it is sufficient to invoke the general necessary condition that the right derivative of the objective function with respect to $W_i$ is negative at the maximal

\(^{12}\)See Clarke (1983) for a detailed exposition on the discussion in the following.
Then, the necessary condition for the optimality is
\[
\frac{1}{w'(w(V, \phi^*(V)) + \pi_i \beta r(e(W_i))(1 - \lambda_e \hat{p}(W_i))} + \pi_i r'(e(W_i))(y + \beta(1 - \lambda_e \hat{p}(W_i))J(W_i)) \Delta(e(W_i), W_i)(1 - \lambda_e \hat{p}(W_i)) + \pi_i \lambda_e J(W_i) \hat{p}'(W_i) \leq 0,
\]
where the optimal of the objective function in (11), if exists, is denoted by \( \phi^*(V) \), and the dependence of optimal \( W_i \) on \( V \) is understood throughout the argument. Dividing the inequality through by \( \pi_i, \beta, r(e(W_i)) \), and \( (1 - \lambda_e \hat{p}(W_i)) \) and rearranging terms yield
\[
-\frac{1}{w'(w(V, \phi^*(V)) + \pi_i \beta r(e(W_i))(1 - \lambda_e \hat{p}(W_i))} - J'(W_i) + \frac{\lambda_e J(W_i) \hat{p}'(W_i)}{1 - \lambda_e \hat{p}(W_i)} + \frac{r'(e(W_i))}{r(e(W_i))} \left( \frac{y}{\beta} + (1 - \lambda_e \hat{p}(W_i))J(W_i) \right) \Delta(e(W_i), W_i).
\]
Because \( \hat{p}'(W_i) \) is non-positive and a feasible contract restricts \( J(W_i) \) to be non-negative, the right-hand side of this inequality is positive. Hence, the above inequality implies
\[-w'(w(V, \phi^*(V)) + \pi_i \beta r(e(W_i))(1 - \lambda_e \hat{p}(W_i)) - J'(W_i) \geq 0.\]
Applying Theorem 1 in Milgrom and Segal (2002) to equation (11) implies \( J'(V) \geq -w'(w(V, \phi^*(V)) + \pi_i \beta r(e(W_i))(1 - \lambda_e \hat{p}(W_i)) \). These two inequalities imply \( J'(V) - J'(W_i) \geq 0 \), which implies \( W_i(V) \geq V \) for \( i = 1, 2 \).
If \( J(\cdot) \) is differentiable at both \( V \) and \( W_i(V) \) and if \( D(\cdot), \hat{p}(\cdot) \), and \( e(\cdot) \) are differentiable at \( W_i(V) \), then we have
\[
-\frac{1}{w'(w(V, \phi^*(V)))} - J'(W_i) = -\frac{\lambda_e J(W_i) \hat{p}'(W_i)}{1 - \lambda_e \hat{p}(W_i)} + \frac{r'(e(W_i))}{r(e(W_i))} \left( \frac{y}{\beta} + (1 - \lambda_e \hat{p}(W_i))J(W_i) \right) \Delta(e(W_i), W_i)
\]
and \( J'(V) = -\frac{1}{w'(w(V, \phi^*(V)))} \), which together imply the Euler equation (9) that holds almost everywhere.

The following proposition translates the novel implications of Proposition 2 in terms of wage dynamics.

**Proposition 3.** Under the optimal contract, wages weakly increase with tenure, independent of the realization of the lottery.
Proof: To simplify the expression, I write \( w(V) = w(V, \phi^*(V)) \) as the optimal wage offer given \( V \). Applying Theorem 1 in Milgrom and Segal (2002) again, I have \( -\frac{1}{w'(w(V))} \geq J'(V) \) and \( J'(W_i+) \geq -\frac{1}{w'(w(W_i)))} \). By the concavity of \( J'(\cdot) \), I also have \( J'(V) \geq J'(W_i+) \). These inequalities imply that \( \frac{1}{w'(W_i))} \geq \frac{1}{w'(w(V))} \), which, in turn, implies \( w(W_i) \geq w(V) \) for \( i = 1, 2 \) by the concavity of \( u(\cdot) \). □

From Propositions 2 and 3, equation (9) recursively characterizes the wage dynamics. Here, the right-hand side of the equation consists of two terms; the one with \( \hat{p}' \) and the other with \( r' \). These two terms imply that the variation in the wage can be decomposed into two forces: the one due to worker on-the-job search, and the other due to endogenous worker effort. Because these terms are positive, both forces contribute to an increasing wage profile.

The decomposition of wage variation in those two forces is consistent with the back-loaded wage contract typically found in on-the-job search models, such as Burdett and Coles (2003) and Shi (2009) as well as in efficiency wage models, such as Lazear (1981) and Lazear and Moore (1984). The optimal contract rewards those who do not quit and those who succeed in the project. The model here provides a simple framework to incorporate these two forces, which have been previously studied separately.

### 4.2 Optimal Worker Effort

Finally, in addition to the optimal wage-tenure profile, the result of Proposition 2 implies a novel property of the optimal incentive-compatible effort.

**Proposition 4.** Under the optimal contract, a worker’s optimal incentive-compatible effort weakly increases with tenure.

**Proof:** For a given contract value \( V \) at a time, define \( W^k(V) \) as a contract value in \( k \)-period ahead, so that \( W^1(V) = W(V) \), \( W^2(V) = W(W(V)) \), and so on. For expositional clarity, I suppress the index for the lottery. By Proposition 2, I have \( W^{k+1}(V) \geq W^k(V) \) for all \( k \) as long as the worker keeps the job. By Lemma 1, the worker’s optimal incentive-compatible effort is a unique increasing function of the continuation value. Therefore \( e(W^{k+1}(V)) \geq e(W^k(V)) \) for all \( k \), which implies the desired result. □

This result contrasts with the decreasing effort discussed in Holmstrom (1999). In his model, workers exert effort for the firm to learn about their productive ability. Because the productive ability is assumed to be constant, once the firm learns the ability, the
worker has no incentive to exert further effort. Therefore, in his model, the work incentive is highest for new workers, and it decreases as the firm learns about their abilities. In this model, however, as Lemma 1 implies, the work incentive is higher for workers with higher continuation value. Then, as the continuation value increases with tenure, the work incentive also increases with tenure.

5 Discussion

5.1 Labor Market Structure

The structure of wage dynamics described above implies that the labor market is divided into a competitive segment and a monopsonistic segment. Figure 1 shows the optimal search strategy and optimal continuation value for a given value of current job. If the current value is sufficiently low, the worker may find a better contract and leave for another employer if on-the-job search is successful. In this competitive segment, a worker’s on-the-job search and endogenous effort both contribute to increasing wages. Once the continuation value becomes sufficiently high, there are no other firms that offer a better contract. This is where the optimal search strategy and optimal continuation value intersect in the figure and \( \hat{p} \) becomes flat at zero, and \( \hat{p}'(W) = 0 \). At this point, the current employer does not have an incentive to compete with the other firms for the employee, and this point determines the highest starting wage offered in equilibrium.

However, equation (9) implies that the wage may still increase even if there is no worker retention motive. By increasing wages beyond the highest starting wage, a firm can still increase the probability of project success and thus increase the expected profit. Thus, an employed worker’s wage will continue to rise until the cost of additional raise exceeds the gain from additional effort. This property implies that the highest starting wage is strictly lower than the highest wage in the equilibrium. The difference between these two bounds on wages is caused by a firm’s monopsonistic power and the need to induce effort.

Note that this gap between the highest equilibrium wage and the highest starting wage is absent in related models of on-the-job search without the moral hazard problem, e.g., Shi (2009) and Menzio and Shi (2010). In those models, once the wage has reached the highest starting wage, increasing the wage further does not change the probability of job separation and the expected profit. In my model, however, there is an incentive motive as in efficiency wage models to further increase the wage. Thus, by introducing the
moral hazard problem regarding worker effort, the current model widens the dispersion of equilibrium wages and extends the wage-tenure profile.

5.2 Wage Mobility and Job Transitions

The optimal contract implies that wage and effort increase with tenure on a job. However, the result also implies that wage and effort increase with duration of employment. A worker’s wage increases in two cases: a raise within a firm or finding a higher-paying job through on-the-job search. In this environment, workers search for a job that provides a higher value than the continuation value promised by the current employer. Lemma 1 and Proposition 3 imply that wage and effort is monotone in the value of contract. Therefore, wages and effort increase if workers stay employed either with the current employer or with another employer after a job-to-job transition.

The optimal contract also implies that the probability of a job-to-job transition decreases with duration of employment. As in Shi (2009), increasing value of current job implies that the probability of finding a better job decreases over time. In addition, increasing effort raises the probability of success and thus lowers the probability of job loss. For both reasons, the probabilities of both voluntary and involuntary job separation decrease with duration of employment. This result highlights how contracts and employed workers’ work incentive interact with their job transitions in the labor market.

An increase in probability of success can also be interpreted as an increase in the expected labor productivity. In this model, newly employed workers are paid low and exert low effort and thus their productivity is low. As the duration of employment becomes longer, the worker effort increases and their productivity also increases. If the worker loses the job at some time and goes through a period of unemployment, the worker will again start with a low wage job, which results in lower productivity. The dynamics labor productivity is often attributed to worker human capital accumulation (Menzio, et al., 2012) or labor market sorting (Lise and Robin, 2013). In contrast, this study provides a new explanation driven by the worker’s incentive regarding the increase in the worker’s productivity due to duration of employment and the reduction in productivity due to unemployment.

Using U.S. Census data, Menzio, et al. (2012) find that tenure profiles of employment-to-unemployment and employment-to-employment transition rates are decreasing. They also find that age profiles of these rates show similar pattern. Even though my model does
not incorporate life-cycle structure of the workers, their findings are generally consistent with the qualitative implications of my model.

5.3 Endogenous Worker Heterogeneity

In this model, dynamic contracts and worker on-the-job search generate endogenous worker heterogeneity in equilibrium even though firms and workers are ex-ante homogeneous. The mechanism is similar to the ones in Burdett and Coles (2003, 2010) and Shi (2009); depending on the history of job transition, the workers are employed at different wages. In addition to this wage dispersion, different wages imply different levels of effort exerted by the workers. This in turn implies different separation rates among homogeneous workers, depending on the duration of employment; job separation rates (both quits and layoffs) are higher for workers with shorter duration of continuous employment and lower with longer duration of continuous employment. In addition to the tenure profile of transition rates discussed in the previous section, these cross-sectional results are also supported by the finding in Menzio, et al. (2012).

6 Conclusion

This study proposed a model with dynamic incentive contracts and on-the-job search in a frictional labor market. I proved the existence of a block recursive equilibrium in an environment with moral hazard of worker effort on the job. I provided a characterization of the optimal contract as an increasing wage-tenure profile. With increasing wages, worker effort also increases with tenure. These two features of the optimal contract imply that probabilities of both voluntary and involuntary job separation decrease with job-tenure as well as with the duration of employment. Given these results, workers experience differing labor market transitions—between employment, unemployment, and across different employers—and the equilibrium generates endogenous heterogeneity among ex-ante homogeneous workers.

The main contribution of this study is theoretical, providing a framework to incorporate a dynamic incentive problem into a labor search model. Nevertheless, it provides a novel insight into how employed worker incentives and their mobility interact and further affect the aggregate outcomes. I believe further work needs to be done to better understand the implications of incentive problems in a macro and labor market framework.
Appendix

A Existence of a Block Recursive Equilibrium

In this appendix, I provide a proof of the existence of a BRE in the environment of this study. The goal of this proof is to show that the operator $T$ defined by equation (8) has a fixed point. The procedure of the proof closely follows that of Menzio and Shi (2010).

First, I define a class of real-valued functions $J(X)$ such that any function $J \in J$ is (i) strictly decreasing and bi-Lipschitz continuous with respect to $V$, (ii) bounded both from below and above, i.e., $J(V) \in [\bar{J}, \check{J}]$ for all $V \in X$, and (iii) concave. Using these properties of $J \in J$, I present general properties of the equilibrium objects that I use to define the operator $T$, i.e., equation (8). Then, I prove the operator $T$ is a self-map (Lemma 8) and continuous (Lemma 9), which provide sufficient conditions for the existence of a fixed point.

A.1 General Properties of the Equilibrium Objects

First, for a function $J \in J(X)$, solving the competitive-entry condition (7) for $\theta$ yields

$$
\theta(x) = \begin{cases} 
q^{-1}(k/J(x)) & \text{if } J(x) \geq k \\
0 & \text{otherwise.}
\end{cases}
$$

(12)

Because $q$ is a probability, $q^{-1}$ is defined on $[0, 1]$; that is, $\theta$ is well defined only when $J(x) \geq k$. Because $J(x)$ is strictly decreasing, there exists a unique $\check{x} \in \mathbb{R}$ such that $J(x) < k$ for all $x > \check{x}$. In equilibrium, no firms enter a submarket with $x > \check{x}$ because offering a contract with a value more than $\check{x}$ provides negative expected profit. Therefore, the market tightness takes a nonnegative value only if $x \leq \check{x}$. This threshold value determines the highest starting value offered in equilibrium given $J$. Because $J$ is bounded from above, $\theta$ is also bounded from above. In addition, $\theta(x)$ exhibits the following properties.

Lemma 3. If $x < \check{x}$, $\theta(x)$ is strictly positive-valued, strictly decreasing, and bi-Lipschitz continuous. $\theta(x) = 0$ for all $x \geq \check{x}$.

See the proof of Lemma 4.1 in Menzio and Shi (2010). Strictly decreasing $\theta(x)$ implies that the market is tighter for entering firms in submarkets that offer lower values of contract, i.e., relatively more vacancies to compete for a worker than searching workers.
The worker optimal search problem exhibits the following properties.

**Lemma 4.**

1. For all \( W < \tilde{x} \), the objective function, \( f(x;W) = p(\theta(x))(x - W) \), is strictly concave with respect to \( x \).
2. A worker’s optimal search strategy, \( m(W) \in \arg\max_x p(\theta(x))(x - W) \), is unique, weakly increasing and bi-Lipschitz continuous.
3. For all \( W < \tilde{x} \), a worker’s value of searching, \( D(W) \), is strictly positive, weakly decreasing and Lipschitz continuous. If \( W \geq \tilde{x} \), \( D(W) = 0 \)

See the proof of Lemma 3.1 in Shi (2009) and the proof of Lemma 4.3 in Menzio and Shi (2010). The second statement is an immediate consequence of the first one. It implies that, for a given continuation value of the current contract, the worker finds a unique submarket to visit. Moreover, the optimal search policy is monotone in the current value of contract. A worker with a higher continuation value searches in a submarket that offers a higher value. However, once a value of the current contract reaches \( \tilde{x} \), there are no other firms that offer a better contract, and the value of search becomes zero.

Combining these properties of the market tightness function and the workers’ optimal search policy, the optimal employment probability \( \hat{p}(W) \) shows the following properties.

**Lemma 5.** \( \hat{p}(W) = p(\theta(m(W))) \) is weakly decreasing and bi-Lipschitz continuous. \( \hat{p}(W) = 0 \) for all \( W \geq \tilde{x} \). Moreover, it is almost everywhere differentiable with respect to \( W \).

See the proof of Corollary 4.4 in Menzio and Shi (2010). The almost-everywhere differentiability is a technical result due to Lipschitz continuity (Folland, 1999). Weakly decreasing \( \hat{p}(W) \) implies that the job-to-job transition probability decreases with the continuation value of the contract.

Next, I show that the worker’s incentive compatible optimal effort is continuous with respect to \( J \in \mathcal{J}(X) \). This is a new element in this study, and formal proofs are provided. Here, I denote \( e(W) \) as \( \tilde{e}(\Omega(W)) \) to explicitly show its dependence on \( W \) through \( \Omega(W) = W + \lambda e D(W) - U \). First, Lemma 6 shows that \( \Omega(\cdot) \) is continuous in \( J \). Then, Lemma 7 shows that \( e(\cdot) \) is continuous in \( J \) using the continuity of \( \Omega \).

**Lemma 6.** Consider \( J_m, J_n \in \mathcal{J}(X) \). Let \( \Omega_j(W) = W + \lambda D_j(W) - U_j \) be the function implied by \( J_j \) for \( j = m,n \). There exists an \( \varepsilon_\Omega > 0 \) such that if \( ||J_m - J_n|| < \rho \), then \( ||\Omega_m - \Omega_n|| < \varepsilon_\Omega \rho \).
Proof: Lemma 4.5 and 4.7 in Menzio and Shi (2010) show that there exist an \( \varepsilon_D > 0 \) and an \( \varepsilon_U > 0 \) such that if \(||J_m - J_n|| < \rho\), then \(||D_m - D_n|| < \varepsilon_D \rho\) and \(||U_m - U_n|| < \varepsilon_U \rho\). Applying these results to the definition of \( \Omega \), I have:

\[
\| \Omega_m(W) - \Omega_n(W) \|
\]

\[
= |\lambda(D_m(W) - D_n(W)) - (U_m - U_n)|
\]

\[
\leq \lambda|D_m(W) - D_n(W)| + |U_m - U_n| < (\lambda \varepsilon_D + \varepsilon_U) \rho \equiv \varepsilon_{\Omega} \rho,
\]

where \( \lambda \varepsilon_D + \varepsilon_U \equiv \varepsilon_{\Omega} \). This result holds for all \( W \in X \). Hence, \(||\Omega_m - \Omega_n|| < \varepsilon_{\Omega} \rho\) and \( \Omega(\cdot) \) is continuous in \( J \). \( \square \)

**Lemma 7.** Consider \( J_m, J_n \in J(X) \). Let \( e_j(W) = \tilde{e}(\Omega_j(W)) \) be the worker’s optimal effort function implied by \( J_j \) for \( j = m, n \). There exists an \( \varepsilon_e > 0 \) such that if \(||J_m - J_n|| < \rho\), then \(||e_m - e_n|| < \varepsilon_e \rho\).

*Proof:* The result of Lemma 1, i.e., Robinson’s implicit function theorem, holds even if I replace \( W \) with \( \Omega(W) \) as an argument of the worker effort function. Therefore, there exists a finite Lipschitz bound \( B_e \) such that

\[
|e_m(W) - e_n(W)| = |\tilde{e}(\Omega_m(W)) - \tilde{e}(\Omega_n(W))|
\]

\[
\leq B_e|\Omega_m(W) - \Omega_n(W)|.
\]

Then, using Lemma 6, \(|e_m(W) - e_n(W)| < B_e \varepsilon_{\Omega} \rho \equiv \varepsilon_{e} \rho\). This result holds for all \( W \in X \). Hence, \(||e_m - e_n|| < \varepsilon_{e} \rho\) and \( e(\cdot) \) is continuous in \( J \).

### A.2 Properties of the Operator \( TJ \)

The following two lemmas give the sufficient conditions for the operator \( TJ \) to admit a fixed point in \( J(X) \).

**Lemma 8.** Let \( \hat{J}(V) = (TJ)(V) \) for any \( J \in J(X) \). Then \( \hat{J}(V) \) belongs to the set \( J(X) \). That is, the operator \( TJ \) is self-mapping.

*Proof:* The goal of this proof is to show that the operator satisfies properties (i)–(iii).
To show property (i) is satisfied, let $F$ be the objective function of (11):

$$F(V, \phi) = \left\{ -u^{-1}\left( V - \beta U - \sum_i \pi_i \left[ -c(e(W_i)) + \beta r(e(W_i)) [W_i + \lambda D(W_i) - U] \right] \right) + \sum_i \pi_i r(e(W_i)) \left[ y + \beta (1 - \lambda \hat{p}(W_i) J(W_i)) \right] \right\}.$$

Note that the function $F(V, \phi)$ is differentiable in $V$. Then, by the Inverse Function Theorem,

$$F'(V, \phi) \equiv \frac{\partial F(V, \phi)}{\partial V} = -\frac{1}{u'(w)} \in \left[ -\frac{1}{u'}, -\frac{1}{u''} \right],$$

because of the assumption that the derivative of the utility function is bounded. Now, for any $V_a, V_b \in X$, such that $V_a \leq V_b$, I have

$$|\hat{J}(V_b) - \hat{J}(V_a)| \leq \max_{\phi} |F(V_b, \phi) - F(V_a, \phi)|$$

$$= \max_{\phi} \left| \int_{V_a}^{V_b} F'(V, \phi) dV \right|$$

$$\leq \max_{\phi} \int_{V_a}^{V_b} \left| F'(V, \phi) \right| dV$$

$$\leq \frac{1}{u'} |V_b - V_a|.$$

Therefore $\hat{J}(V)$ is Lipschitz continuous in $V$. From this result, $\hat{J}(V)$ is absolutely continuous and thus almost everywhere differentiable (Folland, 1999). Then, by Theorem 2 of Milgrom and Segal (2002), the difference $\hat{J}(V_b) - \hat{J}(V_a)$ is such that

$$\hat{J}(V_b) - \hat{J}(V_a) = \int_{V_a}^{V_b} F'(V, \phi(V)) dV,$$

where $\phi(V)$ is the optimal contract given $V$. This equation and (13) together imply that

$$B_J(V_b - V_a) \leq \hat{J}(V_b) - \hat{J}(V_a) \leq B_J(V_b - V_a)$$

where $B_J \equiv -\frac{1}{u'}$ and $B_J \equiv -\frac{1}{u''}$. Hence, $\hat{J}(V)$ is strictly decreasing with bounded difference. The last expression also implies that $\hat{J}(V)$ is bi-Lipschitz continuous in $V$.

To show property (ii) is satisfied, let $e = e(x)$ and $\bar{e} = e(\bar{x})$ be the lower and upper bounds of the level of effort given $x$ and $\bar{x}$, respectively. In addition, let $J = \frac{r(e)y - u^{-1}(x + e(c) - \beta x)}{1-\beta}$ and $\bar{J} = \frac{r(\bar{e})y - u^{-1}(x + \bar{c}(e) - \beta \bar{x})}{1-\beta}$ be the estimates of the bounds for $J \in \mathcal{J}$. Given these bounds, I show that the updated function $\hat{J}(V)$ is bounded.
Let $w(V)$ be the lowest possible wage for a feasible contract given $V$. That is

$$w(V) = \min_{\phi} u^{-1}\left(V - \sum_{i=1,2} \pi_i(-c(e(W_i)) + \beta[r(e_i)(W_i + \lambda D(W_i)) + (1 - r(e_i))U])\right).$$

As $u^{-1}$ is an increasing function, I have $w(V) \geq u^{-1}(x + c(e) - \beta \bar{x})$ for all $V \in X$. Then, the firm’s value for a given $V$ is

$$\hat{J}(V) \leq r(\bar{e})y - w(V) + \beta \bar{J}$$

$$\leq r(\bar{e})y - u^{-1}(x + c(e) - \beta \bar{x}) + \beta \bar{J} \leq \bar{J}.$$

Similarly, if $\bar{w}(V)$ is the highest possible wage given $V$, it can be shown that $\bar{w}(V) \leq u^{-1}(\bar{x} + c(\bar{e}) - \beta \bar{x})$. Then, the firm’s value for a given $V$ is

$$\hat{J}(V) \geq r(e)y - \bar{w}(V) + \beta J$$

$$\geq r(e)y - u^{-1}(\bar{x} + c(\bar{e}) - \beta \bar{x}) + \beta J \geq J.$$

Hence, $\hat{J}(V)$ is bounded both below and above by the respective estimates of $J$ and $\bar{J}$.

Finally, to show property (iii) is satisfied, I can apply the two-point convexification result in Appendix F in Menzio and Shi (2010). They show that $\hat{J}(V)$ is concave in a more general setting, and the presence of moral hazard problem in this study does not alter the proof. Therefore, the proof is omitted in this study.

□

**Lemma 9.** Consider $J_m, J_n \in J(X)$. Let $\hat{J}_j(W) = (TJ_j)(W)$, the updated firm’s value function implied by $J_j$, for $j = m, n$. There exists an $\varepsilon_T > 0$ such that if $||J_m - J_n|| < \rho$, then $||\hat{J}_m - \hat{J}_n|| < \varepsilon_T \rho$.

**Proof:** The goal of this proof is to find a bound for $|\hat{J}_m(V) - \hat{J}_n(V)|$ when $||J_m - J_n|| < \rho$ for some $\rho > 0$. Let $F_j$ be the objective function in (11) implied by $J_j$. Take $V \in X$ such that $\hat{J}_m(V) - \hat{J}_n(V) > 0$. Let $\phi_j$ be the maximizer of $F_j$ such that $\hat{J}_j(V) = F_j(V, \phi_j)$ and $w_j(\phi_j)$ the optimal wage function given by $J_j$.

Using the definition of $F$, the distance between $\hat{J}_m(V)$ and $\hat{J}_n(V)$ is such that

$$|\hat{J}_m(V) - \hat{J}_n(V)| \leq |F_m(V, \phi_m) - F_n(V, \phi_m)|$$

$$= \left| -w_m(\phi_m) + \sum_i \pi_{i,m} r(e_m(W_{i,m})) [y + \beta (1 - \lambda \hat{p}_m(W_{i,m})) J_m(W_{i,m})] \\
+ w_n(\phi_m) - \sum_i \pi_{i,m} r(e_n(W_{i,m})) [y + \beta (1 - \lambda \hat{p}_n(W_{i,m})) J_n(W_{i,m})]\right|$$

22
\[
\leq |w_m(\phi_m) - w_n(\phi_m)|
+ \sum_i \pi_{i,m} \left| r(e_m(W_{i,m})) [y + \beta(1 - \lambda \hat{m}(W_{i,m})) J_m(W_{i,m})] - r(e_n(W_{i,m})) [y + \beta(1 - \lambda \hat{n}(W_{i,m})) J_n(W_{i,m})] \right|.
\]

(14)

The first inequality uses optimality of \(\phi_n\) for \(F_n\), and \(F_n(V, \phi_m) \leq \hat{J}_n(V)\), and the second inequality uses the triangle inequality.

First, I consider the first term in (14). Using the promise-keeping condition, the distance between \(u(w_m(\phi_m))\) and \(u(w_n(\phi_m))\) is bounded by

\[
|u(w_m(\phi_m)) - u(w_n(\phi_m))| \leq \beta |U_m - U_n| + \sum_i \pi_{i,m} \left\{ |c(e_m(W_{i,m})) - c(e_n(W_{i,m}))| + \beta |r(e_m(W_{i,m})) \Omega_m(W_{i,m}) - r(e_n(W_{i,m})) \Omega_n(W_{i,m})| \right\}.
\]

As in the proof of Lemma 6, \(|U_m - U_n|\) is bounded by \(\varepsilon_U \rho\). Also, the distance between \(c(e_m(W_{i,m}))\) and \(c(e_n(W_{i,m}))\) is such that

\[
|c(e_m(W_{i,m})) - c(e_n(W_{i,m}))| \leq c'(\bar{e}) |e_m(W_{i,m}) - e_n(W_{i,m})| \leq c'(\bar{e}) \varepsilon e \rho,
\]

where the first inequality uses the standard result of calculus for the differentiable function \(c(\cdot)\) and the bound of \(e\) and the second inequality uses Lemma 7.

The distance between \(r(e_m(W_{i,m})) \Omega_m(W_{i,m})\) and \(r(e_n(W_{i,m})) \Omega_n(W_{i,m})\) is such that

\[
|r(e_m(W_{i,m})) \Omega_m(W_{i,m}) - r(e_n(W_{i,m})) \Omega_n(W_{i,m})| \\
\leq |r(e_m(W_{i,m})) \Omega_m(W_{i,m}) - r(e_n(W_{i,m})) \Omega_m(W_{i,m})| + |r(e_n(W_{i,m})) \Omega_m(W_{i,m}) - r(e_n(W_{i,m})) \Omega_n(W_{i,m})| \\
\leq |r(e_m(W_{i,m})) - r(e_n(W_{i,m}))| \bar{x} + |r(e_n(W_{i,m})) \Omega_m(W_{i,m}) - \Omega_n(W_{i,m})| \\
\leq r'(\bar{e}) |e_m(W_{i,m}) - e_n(W_{i,m})| \bar{x} + |\Omega_m(W_{i,m}) - \Omega_n(W_{i,m})| \\
\leq (r'(\bar{e}) \varepsilon e \bar{x} + \varepsilon \Omega) \rho,
\]

where the first inequality uses the triangle inequality, the second inequality uses the fact that \(\Omega(\cdot)\) is bounded by \(\bar{x}\), and the third inequality uses the standard result of calculus for the differentiable function \(r(\cdot)\), Lemma 7, and the bound of \(r \leq 1\). Therefore, the distance between \(u(w_m(\phi_m))\) and \(u(w_n(\phi_m))\) is bounded by

\[
|u(w_m(\phi_m)) - u(w_n(\phi_m))| \leq (\beta \varepsilon_U + c'(\bar{e}) \varepsilon e + \beta (r'(\bar{e}) \varepsilon e \bar{x} + \varepsilon \Omega)) \rho.
\]
As $u$ is a concave function, for any $w_1$ and $w_2$, $|w_1 - w_2| u' < |u(w_1) - u(w_2)|$. Hence, the first term of (14) is bounded as

$$|w_m(\phi_m) - w_n(\phi_n)| \leq \left( \frac{\beta \varepsilon_U + c'(e) e_e + \beta (r'(e) e_x + \varepsilon)}{u'} \right) \rho. \quad (15)$$

To bound the second term in (14), first, the distance between $r(e_m(W_{i,m})) J_m(W_{i,m})$ and $r(e_n(W_{i,m})) J_n(W_{i,m})$ is such that

$$|r(e_m(W_{i,m})) J_m(W_{i,m}) - r(e_n(W_{i,m})) J_n(W_{i,m})|$$

$$\leq |r(e_m(W_{i,m})) J_m(W_{i,m}) - r(e_n(W_{i,m})) J_m(W_{i,m})| + |r(e_n(W_{i,m})) J_m(W_{i,m}) - r(e_n(W_{i,m})) J_n(W_{i,m})|$$

$$\leq r'(e)|e_m(W_{i,m}) - e_n(W_{i,m})| J + |J_m(W_{i,m}) - J_n(W_{i,m})|$$

$$\leq (r'(e) e_x J + 1) \rho,$$

where the first inequality uses the triangle inequality, the second inequality uses the bound of $J$ and $r$ and the third inequality uses Lemma 7 and the assumption that $||J_m - J_n|| < \rho$.

Second, the distance between $p_m(W_{i,m}) J_m(W_{i,m})$ and $p_n(W_{i,m}) J_n(W_{i,m})$ is such that

$$|p_m(W_{i,m}) J_m(W_{i,m}) - p_n(W_{i,m}) J_n(W_{i,m})|$$

$$\leq |p_m(W_{i,m}) J_m(W_{i,m}) - p_n(W_{i,m}) J_m(W_{i,m})| + |p_n(W_{i,m}) J_m(W_{i,m}) - p_n(W_{i,m}) J_n(W_{i,m})|$$

$$= |p_m(W_{i,m}) - p_n(W_{i,m})| J_m(W_{i,m}) + p_n(W_{i,m})|J_m(W_{i,m}) - J_n(W_{i,m})|$$

$$\leq (\varepsilon p J + 1) \rho, \quad (16)$$

where the first inequality uses the triangle inequality. For the second inequality, the distance between $p_m(W_{i,m})$ and $p_n(W_{i,m})$ is $|p_m(W_{i,m}) - p_n(W_{i,m})| < \varepsilon p \rho$ by Lemma 4.5 in Menzio and Shi (2010). The second inequality also uses the bound of $J$ and $p_n \leq 1$, and the assumption that $||J_m - J_n|| < \rho$.

Last, the distance between $r(e_m(W_{i,m})) p_m(W_{i,m}) J_m(W_{i,m})$ and $r(e_n(W_{i,m})) p_n(W_{i,m}) J_n(W_{i,m})$ is such that

$$|r(e_m(W_{i,m})) p_m(W_{i,m}) J_m(W_{i,m}) - r(e_n(W_{i,m})) p_n(W_{i,m}) J_n(W_{i,m})|$$

$$\leq |r(e_m(W_{i,m})) p_m(W_{i,m}) J_m(W_{i,m}) - r(e_n(W_{i,m})) p_m(W_{i,m}) J_m(W_{i,m})|$$

$$+ |r(e_n(W_{i,m})) p_m(W_{i,m}) J_m(W_{i,m}) - r(e_n(W_{i,m})) p_n(W_{i,m}) J_n(W_{i,m})|$$

$$= |r(e_m(W_{i,m})) - r(e_n(W_{i,m}))| p_m(W_{i,m}) J_m(W_{i,m})$$

$$+ r(e_n(W_{i,m}))|p_m(W_{i,m}) J_m(W_{i,m}) - p_n(W_{i,m}) J_n(W_{i,m})|$$
\[
\leq ((r'(\epsilon)e_\epsilon \bar{J} + 1)\bar{J} + (\epsilon_p \bar{J} + 1))\rho,
\]

where the first inequality uses the triangle inequality, and the second inequality uses the standard result of calculus for the differentiable function \( r(\cdot) \), Lemma 7, the bound of \( J \) and \( r \), and the bound in (16).

Using these inequalities, the second term in (14) is bounded by

\[
\sum_i \pi_{i,m} |r(e_m(W_{i,m})) [y + \beta(1 - \lambda \hat{p}_m(W_{i,m}) J_m(W_{i,m})] - r(e_n(W_{i,m})) [y + \beta(1 - \lambda \hat{p}_n(W_{i,m}) J_n(W_{i,m})] |
\leq ((y + \beta + \beta \lambda \bar{J})(r'(\epsilon)e_\epsilon \bar{J} + 1) + r(\epsilon_p \bar{J} + 1)) \rho. \tag{17}
\]

Finally, using the bounds in (15) and (17), the inequality (14) becomes

\[
|\hat{J}_m(V) - \hat{J}_n(V)|
\leq \frac{(\beta \varepsilon U + c'(\bar{\epsilon})e_\epsilon \bar{\epsilon} + \beta (r'(\epsilon)e_\epsilon \bar{\epsilon} + \epsilon \Omega)) \rho + ((y + \beta + \beta \lambda \bar{J})(r'(\epsilon)e_\epsilon \bar{J} + 1) + r(\epsilon_p \bar{J} + 1)) \rho}{y'}
\equiv \varepsilon_T \rho.
\]

This inequality holds for all \( V \in X \). Hence, \( ||\hat{J}_m - \hat{J}_n|| < \varepsilon_T \rho \), and the operator \( T \) is continuous in \( J \).

\[\square\]

A.3 Existence of a BRE

Proof of Proposition 1: First, fix an arbitrary positive real number \( \varepsilon \). Let \( \rho_\varepsilon = \frac{\varepsilon}{\varepsilon_T} \). Then, if \( ||J_m - J_n|| < \rho_\varepsilon \), then \( ||\hat{J}_m - \hat{J}_n|| < \varepsilon \) for all \( \hat{J}_m, \hat{J}_n \in T(J) \). It implies that the family of functions \( T(J) \) is equicontinuous. Hence, the operator \( T \) satisfies the assumptions of Schauder’s Fixed Point Theorem (Stokey and Lucas with Prescott, 1989, Theorem 17.4); that is, (i) \( T \) maps the set \( J(X) \) into itself, (ii) \( T \) is continuous, and (iii) the family of functions \( T(J) \) is equicontinuous. Therefore, there exists a function \( J^* \in J(X) \) such that \( TJ^* = J^* \). Denote with \( \{D^*, m^*, U^*, \xi^*, \theta^*\} \) the respective functions derived from \( J^* \). By construction, the functions \( \{D^*, m^*, U^*, J^*, \xi^*, \theta^*\} \) do not depend on the distribution of workers and thus constitute a Block Recursive Equilibrium. \[\square\]
Figure 1: Search Strategy and Continuation Value

Note: The figure is based on the computed model with the following functional forms and parameter values: $u(w) = w^{1-\sigma}/(1-\sigma)$ with $\sigma = 2$, $c(e) = \frac{1}{2}e^2$, $p(\theta) = \theta(1 + \theta\gamma)^{-1/\gamma}$ with $\gamma = 0.5$, $r(e) = \exp(-\frac{\rho}{e})$ with $\rho = 0.008$, $y = 1$, $\beta = 0.996$, $b = 0.385$, $k = 0.073$, $\lambda_u = 1$, and $\lambda_e = 0.85$. 
References


