Equilibrium Wage-Tenure Contract with Unobserved Human Capital

Seung-Gyu Sim
University of Wisconsin at Madison

October 19, 2009

Abstract

This paper develops an equilibrium job search model in which the employed worker privately accumulates human capital and keeps searching for a better paying job. If he accumulates human capital, he can produce more at the same fixed cost. If he finds a better offer, he switches to the job at a fixed switching cost. The firm can observe neither the level of human capital of the worker, nor the job search outcome. To induce truthful revelation, the firm pays more when he produces more (productive promotion). To discourage job turnover, the firm back-loads some portion of wage payments and pays conditioning on his future stay (non-productive promotion). We show that there are multiple wage-ladders that a worker can climb up in equilibrium. When he receives a non-productive promotion, he moves up one rung in the current ladder, while a productive promotion moves the worker to a higher-valued ladder. We estimate the model using indirect inference to investigate the effect of human capital accumulation on individual wage growth. In particular, we capture the effect of productive promotion separately from that of non-productive promotion using reemployment wages. In the NLSY79 data, the average wage after 20 years of market experience is 188% higher than the average of the first full-time wages. Counterfactual analysis using the structural parameter estimates shows that if a typical worker were not able to accumulate human capital, his wage would grow by 42%.

1 Introduction

This paper develops an equilibrium job search model in which the employed worker privately accumulates human capital and keeps searching for a better paying job. He[1] privately learns how to reduce his cost (time) simply by repeating similar tasks on the job. Given right incentives, he can produce more at the same fixed cost. Without incentives, he enjoys private gains from leisure. To induce truthful revelation, the firm pays more when he produces more (productive promotion). In the mean time, he keeps searching for a better paying job as unemployed workers do. The search outcome is not observed by the firm. To retain the worker longer or to extract more surplus from the early leaver, the firm back-loads some portion of wage payments and pays conditioning on his future stay (non-productive promotion). We show that there are multiple wage-ladders that a worker can climb up in equilibrium. When he receives a non-productive promotion, he moves up one rung in the current ladder, while a productive promotion moves him to a higher-valued ladder. We estimate the model to investigate the effect of human capital accumulation on individual wage growth.

Topel and Wald (1992) point out that a typical male worker works for 40 years, and his wage doubles over his career (in the cross section). In our sample of white male high school graduates, we also find that the average wage after 20 years of market experience is 1.88 times larger than the the average of first full-time wages. It grows by 43% and 61% after 5 and 10 years of market experience, respectively. In general, this is interpreted as the outcome of sequential job turnover and human capital accumulation. However, to incorporate on-the-job search behavior and human capital accumulation into a search framework brings about two fundamental questions; 'what is the best strategy by the firm against the worker's on-the-job search behavior?' and 'why does the monopsony[2] firm pay for the additional human capital accumulated on the job?'.

In their seminal work, Burdett and Mortensen (1998) (hereafter BM) demonstrate that on-the-job search creates heterogeneity in workers' reservation wages, which generates wage dispersion among homogeneous workers with homogeneous jobs. In the BM framework, the individual wage rises through sequential job turnover, although workers receive a constant wage in one job. Stevens (2004) points out that firms can do better by offering a wage-tenure contract rather than a single wage. She shows that when workers are risk neutral, there are multiple equilibrium wage-tenure contracts which increase pay with job tenure. Motivated by this, Burdett and Coles (2003) (hereafter BC) build up an equilibrium job search model in which workers are risk averse so that firms gradually raise wage payments with job tenure. They show that there exists a

[1]Throughout the paper, we use the masculine pronoun for a worker because we deal with white male sample in the empirical work.

[2]due to search friction

unique wage-tenure schedule in equilibrium and different firms choose different starting points on it. We extend the BC framework by incorporating (unobserved) human capital accumulation. Workers are risk averse and ex ante homogenous, but independently accumulate human capital. In our model, wage rises not only due to job turnover and the back-loading strategy of the firm, but also due to productivity increases.

Recently, Burdett, Carrillo-Tudela, and Coles (2008) develop a job search model with human capital accumulation. They assume that workers are paid by a piece rate sharing rule, which is determined by the firm at the recruiting stage. Through numerical simulations, they provide a coherent variance decomposition of the wage paid. However, their work is vulnerable to the criticism that firms have no reason to stick to a single constant piece rate. In contrast, this paper is motivated by the fact that firms put substantial effort into designing and implementing incentive schemes. In our model, each firm offers an optimal menu of contracts which specifies the lifetime value of the worker contingent on the worker's output. Given the incentive scheme, the worker increases output to receive the information rent when he becomes more productive. The promised lifetime value is delivered in the form of a wage payment, a promotion rate, and a new contract after promotion (continuation value).

Interestingly, the equilibrium has multiple value-ladders that a worker can climb up or move from one to another. When he receives a non-productive promotion, he moves up one rung in the current ladder, while a productive promotion moves the worker to a higher-valued ladder. Previously, Delacroix and Shi (2006) show that the pure strategy equilibrium has a wage-ladder structure in the directed search framework. In their model, workers are homogeneous in terms of productivity, but have different reservation values due to on-the-job search. Workers simultaneously decide to which firm they apply. In this environment, they show that the equilibrium distributions of wage received and wage offered have discrete mass points. In our model, instead of workers, firms decide to which value they promote their workers, which, together with the fixed switching cost s, discretizes the support of the two distributions. The existence of the fixed switching cost is crucial in our model. However, we point out that there are no significant changes to the equilibrium outcome even though we send s to zero. In our numerical simulations with $s = 0.04, 0.02, 0.01$ and 0.005, the distributions of values offered and values earned remain unchanged except that they have more mass points and more refined values.[3]

In the empirical analysis, we investigate the effect of human capital accumulation on individual wage growth. For this, we construct a sample of white male high school graduates from the National Longitudinal Survey of Youth 1979. We keep track of individual workers in terms of 'non-employment', 'hours worked', 'job tenure', 'em-

[3] as long as we correspondingly increase the number of value grid points.

ployment tenure', 'work experience', 'market experience', 'wage' and 'reemployment wage'. We estimate the structural parameters of the model through indirect inference. The model implies that once the worker accumulates human capital, his productivity is permanently improved. In contrast, the accumulated offers through non-productive promotion or job turnover are reset when the worker is unemployed. Hence, we take advantage of the reemployment wage to extract the effect of human capital accumulation (productive promotion) on wage growth.

Using the estimates, we perform a counterfactual experiment on what would happen if a typical worker were not allowed to accumulate human capital. In the sample, the average wage after 20 years of market experience is 188% higher than the average of first full-time wages. The counterfactual analysis reports that if a typical worker were not able to accumulate human capital, his wage would grow by 41%. This result is somewhat surprising in the sense that the contribution of human capital accumulation on wage growth is not so large as in Altonji and Shakotko (1987) and Altonji and Williams (2005). They interpret 'return to experience' and 'return to tenure' as the outcome of general human capital accumulation and job specific human capital accumulation, respectively. Then, they show that return to experience (thus, general human capital accumulation) takes the lion's share in individual wage growth. Therefore, the effect of general human capital accumulation includes all factors which are transferred to the next job in job-to-job transition. But in our model, even the effect of non-productive promotion is also transferred to the next job through the reservation value. In that sense, their result includes the contribution of both productive promotion and non-productive promotion.

The paper proceeds as follows. In section 2, we build up the theoretical model and characterize the equilibrium that we are interested in. In section 3, we construct the sample and define relevant variables, and in section 4, we provide the estimation protocol and estimation result. Section 5 concludes.

2 The Model

2.1 Basic Framework

Consider a labor market populated by a unit measure of risk-neutral firms and a unit measure of risk-averse workers. Firms are infinitely-lived and homogeneous, while workers follow a stochastic birth-retirement process and independently accumulate human capital. In particular, a newly born worker starts his career with y_1 units of human capital, accumulates human capital during his lifetime, and retires at rate

$\rho \in (0, \infty)$. We call the worker with y units of human capital 'y-type worker'. His type is private information so that firms cannot observe it. The firm chooses a menu of contracts (or just contract interchangeably) in order to induce truthful revelation. We restrict our attention to a separating equilibrium in which firms offer the least cost incentive compatible menus of contracts[4], and do not screen out any types. The model is set in continuous time, and all workers and firms discount the future at rate $r \in (0, 1)$.

The worker maximizes his lifetime expected utility:

$$E\Big[\int_0^\infty e^{-(\rho+r)t}[u(w(t)) - c(t)]dt\Big],$$

where $u' > 0$ and $u'' < 0$. We assume that a worker does not have access to financial markets. An unemployed worker collects a flow benefits ($w(t) = b$) from the unemployment benefit at no cost ($c(t) = 0$). An employed worker receives wage payment w specified in the labor contract. The unit of output is normalized so that a y-type employed worker can produce y units of output. If he produces $\hat{y}$ units of output, his flow cost is given by

$$c(\hat{y}; y) = \begin{cases} c_0 - c_1(y - \hat{y}) & \text{if } \hat{y} \leq y \\ \infty & \text{otherwise} \end{cases}.$$

The cost function reflects time cost. Workers spend the same amount of time on working at the same fixed cost, which is captured by c_0. A worker with more human capital is able to produce more output. But since it is his private information, he produces more if the right incentive is provided. Otherwise, he produces less than his ability and enjoys private gains from shirking, which is captured by c_1. Note that in both the social planner problem and the market equilibrium, a y_j-type worker produces y_j units of output. The parameter c_1 only affects how the surplus is split. For simplicity, it is assumed that $c_0 = 0$ throughout the paper.[5]

The worker privately and stochastically accumulates human capital on the job at rate $\mu \in (0, \infty)$. When he gets a human capital accumulation shock, he accumulates $\Delta(> 0)$ units of human capital. It is assumed that

$$y \in \mathcal{Y} := \{y_1, y_2, \cdots, y_{n_j}\}, \quad \text{where} \quad y_j = y_1 + (j-1)\Delta.$$

One may argue that this is somewhat unrealistic since human capital accumula-

[4]Later, we will define it in a rigorous way.

[5]In the empirical study, c_0 and b cannot not be separately identified. For any pair of (b, c_0), there always exists multiple pairs of $(b', c_0') \neq (b, c_0)$ such that (b', c_0') generates the exactly same equilibrium outcome as (b, c_0).

tion occurs more frequently among less productive workers. Then, we can assign $(\mu_1, \mu_2, \cdots, \mu_{nj})$ instead of the unique μ. But, to keep the model simple, we assume it is constant.

In addition to the cost of working, we also assume a fixed cost of switching jobs, s. It reflects the fact that when a worker starts a new job, he must adjust himself to a new working condition and new environment. A worker will switch to a new job when the expected value of the new job covers his reservation value plus the switching cost. As a tie-breaking rule, it is assumed that if the net gains from switching is nonnegative, the worker prefers switching jobs.

An unemployed worker retires at rate $\rho \in (0, \infty)$, or finds a job offer at rate $\lambda_u \in (0, \infty)$. Let U_j denote the equilibrium asset value for a y_j-type unemployed worker. Also, let $F_j(\cdot)$ denote the equilibrium distribution of lifetime values offered to y_j-type workers by recruiting firms. The HJB equation for the y_j-type unemployed worker is given by

$$rU_j \quad = \quad u(b) + \lambda_u \int \max\{x - s - U_j, 0\} dF_j(x) - \rho U_j, \tag{1}$$

where x stands for a lifetime value offered to y_j-type workers on equilibrium path. It is examined more precisely in what follows.

Denote by m^f an arbitrary contract potentially offered by the firm. Let $E(\hat{y}; y_j, m^f)$ be the expected lifetime value of a y_j-type employed worker producing $\hat{y}$ units of output under contract m^f. Shortly, let $E_j(m^f) := E(y_j; y_j, m^f)$ when the y_j-type worker produces y_j units of output under contract m^f. The contract m^f specifies the action profile (or 'terms of trade') which consists of the levels of output produced by the worker and the corresponding lifetime values provided by the firm under the truthful revelation assumption. Then, m^f is characterized by $\{(y_j, E_j(m^f))\}_{j=1}^{n_j}$. In what follows, we use m^f and $\{(y_j, E_j(m^f))\}_{j=1}^{n_j}$ interchangeably to express a contract. The lifetime value of a y_j-type employed worker with contract m^f is determined by

$$\max \ \{E(\hat{y}; y_j, m^f) \mid \hat{y} \in \mathcal{Y}\}.$$

An employed worker retires at rate $\rho \in (0, \infty)$, separates from his job at rate $\delta \in (0, \infty)$, or finds another job offer at rate $\lambda \in (0, \lambda_u)$. Also, he accumulates human capital at exogenous rate $\mu \in (0, \infty)$, and gets promoted at the rate promised by the firm. For expositional convenience, we assume that when he gets promoted, he signs a new contract with higher contingent values. The promised lifetime value $\{E_j(m^f)\}_{j=1}^{n_j}$ is delivered through wage payments (as a function of output), $\{w^f(y_j)\}_{j=1}^{n_j}$, promotion rates, $\{\eta^f(y_j)\}_{j=1}^{n_j}$, and a new contract after promotion $\{(y_j, E_j(m^{f'}))\}_{j=1}^{n_j}$. For each $y_j (< y_{n_j})$, the lifetime value of the y_j-type worker who produces $\hat{y}$ units of output under contract m^f is given by

$$\begin{aligned} rE(\hat{y};y_j,m^f) &= u(w^f(\hat{y})) - c(\hat{y};y_j) + \lambda \int \max\{x - s - E(\hat{y};y_j,m^f),0\}dF_j(x) - \rho E(\hat{y};y_j,m^f) \\ &+\delta(U_j - E(\hat{y};y_j,m^f)) + \mu(E(y_{j+1};y_{j+1},m^f) - E(\hat{y};y_j,m^f)) \\ &+\eta^i(\hat{y})(E(y_j;y_j,m^{f'}) - E(\hat{y};y_j,m^f)). \end{aligned} \tag{2}$$

When $y_j = y_{n_j}$, we drop $\mu(E(y_{j+1};y_{j+1},m^f) - E(\hat{y};y_j,m^f))$ in (2), because a y_{n_j}-type worker cannot accumulate human capital any more, having reached the maximum level of human capital. Then, incentive compatibility implies that

$$E_j(m^f) = \max\ \{E(\hat{y};y_j,m^f) \mid \hat{y} \in \mathcal{Y}\}. \tag{3}$$

As the second tie-breaking rule, we assume that when a y_j-type worker is indifferent, he truthfully produces y_j. The set of promise-keeping constraints requires that for each y_j,

$$\begin{aligned} rE_j(m^f) &= u(w^f(y_j)) - c_0 + \lambda \int \max\{x - s - E_j(m^f),0\}dF_j(x) - \rho E_j(m^f) \\ &+\delta(U_j - E_j(m^f)) + \mu(E_{j+1}(m^f) - E_j(m^f)) + \eta^i(y_j)(E_j(m^{f'}) - E_j(m^f)). \end{aligned} \tag{4}$$

As before, when $y_j = y_{n_j}$, we drop $\mu(E_{j+1}(m^f) - E_j(m^f))$ in (4).

A firm can create and maintain one vacant job. The firm with a vacancy posts a menu of contracts to attract workers and induce truthful revelation. It is assumed that a firm offers all new hires the same contract.[6] Once a worker accepts the contract, he immediately starts working. If he produces y_j units of output, the operating firm receives revenue y_j and makes wage payments from the revenue. The match is destroyed when he leaves the job either voluntarily or involuntarily. In equilibrium, since all jobs yield positive expected profit to firms, there is no endogenous firing.[7]

The operating firm with contract m^f optimally determines wage payments (as a function of output), $\{w^f(y_j)\}_{j=1}^{n_j}$, promotion rates, $\{\eta^f(y_j)\}_{j=1}^{n_j}$, and the new contract after promotion $\{(y_j, E_j(m^{f'}))\}_{j=1}^{n_j}$, given that the contingent schedules deliver the committed lifetime values $\{E_j(m^f)\}_{j=1}^{n_j}$, and the new contract $m^{f'}$ is also incentive compatible. Let $J_j(m^f)$ (or briefly J_j^f) be the equilibrium value of the job with a

[6] Anti-discrimination legislation might require equal treatment of all new hires. If not, the firm incurs a sufficiently large cost. We borrow this assumption from Burdett and Coles (2003).

[7] Firm cannot even threaten workers with firing, because workers already know that the threat is an empty threat.

y_j-type worker under contract m^f. The value of the job is given by

$$\begin{aligned} rJ_j(m^f) \;=\; \max_{w^f(y_j),\ \eta^f(y_j),\ m^{f\prime}\geq 0} \; & y_j - w^f(y_j) - (\rho + \delta + \lambda(1 - F_j(E_j(m^f)))J_j(m^f) \\ & + \mu(J_{j+1}(m^f) - J_j(m^f)) + \eta^f(y_j)(J_j(m^{f\prime}) - J_j(m^f)), \quad (5) \end{aligned}$$

subject to the set of incentive compatibility constraints (3) and the set of promise-keeping constraints (4).

We focus on the equilibrium in which the recruiting firm plays the following 'empty threat' strategy:

- The recruiting firm chooses m^f to maximize the expected profit.
- If the contract is accepted by a worker, they immediately execute the contract.
- If the contract is rejected due to the switching cost, the firm makes a revised offer compensating the switching cost.

The revision never occurs in equilibrium. It is introduced to rule out deviation by the operating firms. The existence of the switching cost discretizes the equilibrium support of $\{F_j\}_{j=1}^{n_j}$. In equilibrium, the job turnover rate of a y_j-type employed worker with lifetime value x is given by $\lambda(1 - F_j(x))$.[8] When F_j has the discrete support evenly spaced with increment s, the operating firm can discretely lower the turnover rate into $\lambda(1 - F_j(x+s))$ by increasing the lifetime value by sufficiently small $\varepsilon > 0$. In this sense, the operating firm with contract m^f has incentive to immediately promote the worker to the contract with $E_j(m^f) + \varepsilon$. To prevent the immediate ε-promotion, we allow the empty threat strategy by the recruiting firm. Under this strategy by the recruiting firm, the operating firm should increase the lifetime value of the worker at least by s to lower the turnover rate. Note that all these arguments are irrelevant on equilibrium path. In addition, as s goes to zero, the potential distortion from this empty threat is mitigated.

Let $g : \mathcal{Y} \to [0, 1]$ be the steady state probability mass function of $\mathcal{Y}$. It has n_j mass points and $\sum g(y_j) = 1$. Also, let $G_j(x)$ be the cumulative distribution of the equilibrium lifetime values received by y_j-type workers. Denote by $\mathcal{M}$ the set of the equilibrium contracts. The equal profit condition implies that

$$\sum_{j=1}^{n_j} (\lambda G_j(E_j(m^f) - s)g(y_j) + \lambda_u G_j(U_j)g(y_j))J_j(m^f) \begin{cases} = \pi, & \text{if } \ m^f \in \mathcal{M}, \\ < \pi & \text{otherwise.} \end{cases}$$

[8] We will see this in detail later.

By aggregating all recruiting firms' strategies, we get the distribution of lifetime values offered to each type $\{F_j\}_{j=1}^{n_j}$. Operating firms choose a wage and promotion schedule, which determine the steady state distribution of expected lifetime values received by each type, $\{G_j\}_{j=1}^{n_j}$. Given $\{G_j\}_{j=1}^{n_j}$, recruiting firms choose optimal contracts, which form $\{F_j\}_{j=1}^{n_j}$ in turns. The equilibrium is defined as follows.

Definition A market equilibrium requires:

(i) A y_j-type unemployed worker accepts contract $\{(y_j, E_j(m^f))\}_{j=1}^{n_j}$ if and only if

$$\max_{\hat{y}}\{E(\hat{y}; y_j, m^f)\} - s \geq U_j.$$

(ii) A y_j-type employed worker optimally chooses the level of output, and accepts a new contract $m^{f''}$ if and only if

$$\max_{\hat{y}}\{E(\hat{y}; y_j, m^{f''})\} - s \geq \max_{\hat{y}}\{E(\hat{y}; y_j, m^f)\}.$$

(iii) An operating firm with contract m^f optimally chooses $\{(w^f(y_j), \eta^f(y_j))\}_{j=1}^{n_j}$ and $m^{f'}$ to deliver $\{E_j(m^f)\}_{j=1}^{n_j}$.

(iv) A recruiting firm optimally posts contract m^f given the equal profit condition described in (6). If it is rejected, the firm make a one-time revised offer by adding s to the lifetime value of each type.

(v) The equilibrium distributions $\{F_j, G_j\}_{j=1}^{n_j}$ are stationary.

2.2 Equilibrium Characterization

In this subsection, we characterize a market equilibrium in which all firms offer least cost incentive compatible contracts. To this end, we define the least cost incentive compatible contract first and construct an equilibrium under the restriction that all firms offer a contract from the set of least cost incentive compatible contracts. Then, we examine under what condition the strategy set of the firm can be generalized without affecting the equilibrium outcome.

Definition *A contract, m^f, is* least cost incentive compatible *if*

(i) $E_j(m^f) \geq E(\hat{y}; y_j, m^f)$ *for any* $y_j, \hat{y}_j \in \mathcal{Y}$*, and*

(ii) for each $y_j(\neq y_1)$*, there exists at least one* $\hat{y}(\neq y_j)$ *such that the equality is binding.*

To distinguish it from m^f, denote by m^i an arbitrary least cost incentive compatible contract. Lemma 1 characterizes the set of least cost incentive compatible contracts.

Lemma 1

(i) If the contract characterized by $\{(y_j, E_j(m^f))\}_{j=1}^{n_j}$, is least cost incentive compatible, the contract characterized by $\{(y_j, E_j(m^f) + x)\}_{j=1}^{n_j}$, is also least cost incentive compatible.

(ii) If the contract characterized by $\{(y_j, E_j(m^f))\}_{j=1}^{n_j}$, is least cost incentive compatible, all other least cost incentive compatible contracts are represented by $\{(y_j, E_j(m^f) + x)\}_{j=1}^{n_j}$.

The first statement tells us that if a contract is least cost incentive compatible, all other contracts which pay the same amount of information rent are also least cost incentive compatible. The second statement implies that all least cost incentive compatible contracts should pay the same amount of information rent. Intuitively, these are true by the definition of the least cost incentive compatibility. Then, Lemma 1 also implies that in equilibrium,

$$F_j(E_j(m^i)) = F_{j'}(E_{j'}(m^i)), \quad \forall y_j, y_{j'} \in \mathcal{Y}. \tag{6}$$

Denote by m^0 the equilibrium contract which delivers the lowest value to each type of worker. Then, by definition, the contract m^0 is least cost incentive compatible and attracts only unemployed workers. Lemma 2 shows that m^0 is uniquely defined.

Lemma 2

(i) $s = E_1(m^0) - U_1 < E_2(m^0) - U_2 < \cdots < E_{n_j}(m^0) - U_{n_j}$.

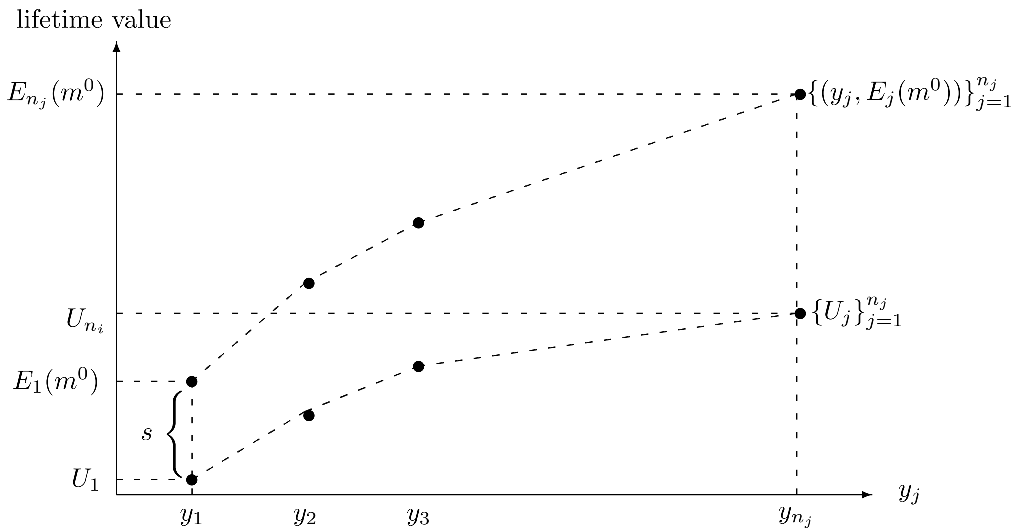
(ii) The equilibrium contract, $\{(y_j, E_j(m^0))\}_{j=1}^{n_j}$, uniquely exists.

To attract all types of unemployed workers, the contract m^0 must have $E_j(m^0) \geq U_j + s$ for each $y_j \in \mathcal{Y}$, At the same time, it should be least cost incentive compatible. Lemma 2 shows that there exists a unique least cost incentive compatible contract such that

$$E_1(m^0) = U_1 + s, \quad \text{and} \quad E_j(m^0) > U_j + s, \quad \forall j = 2, 3, \cdots, n_j.$$

By Lemma 1, we also know that m^0 delivers the lowest contingent values among all equilibrium contracts. [Figure 1] explains this. The gap between $E_1(m^0)$ and U_1 is

s, and the gap between $E_j(m^0)$ and U_j increases in y_j. Any contract that pays less than m^0 cannot attract y_1-type unemployed workers because of the switching cost. Also, any contract that pays more than m^0 cannot be optimal or cannot be least cost incentive compatible.



[Figure 1]

Lemma 3

(i) The set of equilibrium contracts $\mathcal{M}$ is finite. In particular,

$$\mathcal{M} = \{m^0, m^1, \cdots, m^{n_i}\}, \quad \text{where} \quad m^i = \{(y_j, E_j(m^0) + is)\}_{j=1}^{n_j}$$

(ii) The operating firm with contract $m^i \in \mathcal{M}$ promotes the worker to the contract characterized by $\{(y_j, E_j(m^i) + s)\}_{j=1}^{n_j}$.

Lemma 3 implies that the set of lifetime values offered on equilibrium path has a value-ladder structure with increment s in which the operating firms promote their workers one rung at a time. The number of equilibrium contracts, n_i, can be any positive integer less than the upper limit which is determined by the condition that

$$J_j(m^{n_i}) \geq 0, \quad \forall y_j \in \mathcal{Y} \tag{7}$$

Hence there are multiple equilibria with different n_i. No matter what n_i is, the worker with contract m^{n_i} believes that there is no better offer so that he has no incentive to search. Given this, firms also have no reason to offer any contract which delivers higher contingent values than contract m^{n_i} does. In the numerical simulation, we set

$n_i = 20$ and $s = 0.01$.

Lemma 4

(i) For any least cost incentive compatible contract m^i,

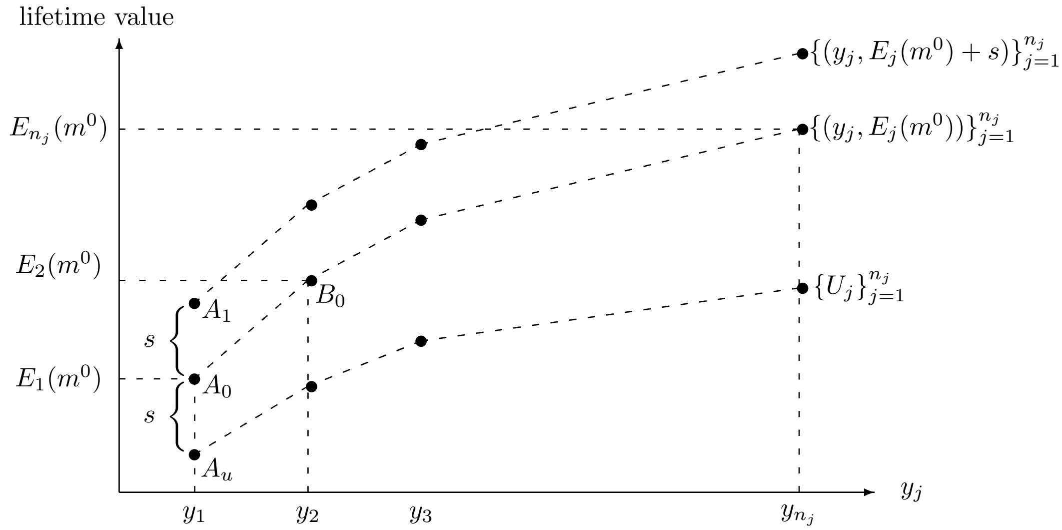
$$E_j(m^i) = E(y_{j-1}; y_j, m^i), \quad \forall j = 2, 3, \cdots, n_j.$$

(ii) The contract characterized by $\{(y_j, E_j(m^f))\}_{j=1}^{n_j}$, is least cost incentive compatible if and only if for $j = 2, 3, \cdots, n_{j-1}$,

$$\begin{aligned} E_j(m^i) &= E_{j-1}(m^i) + \frac{c_1\Delta + \delta(U_j - U_{j-1}) + \mu(E_{j+1}(m^i) - E_j(m^i))}{r + \delta + \rho + \mu}, \quad \text{and} \\ E_{n_j}(m^i) &= E_{n_j-1}(m^i) + \frac{c_1\Delta + \delta(U_{n_j} - U_{n_j-1})}{r + \delta + \rho + \mu}. \end{aligned}$$

The first statement in Lemma 4 says that in equilibrium, downward adjacent incentive compatibility constraints are binding. The second statement shows the amount of additional lifetime values given to a more productive type in order to induce truthful revelation.

[Figure 2] summarizes Lemma 1 through Lemma 4. A newly born worker starts his career with y_1 units of human capital at point A_u. He finds a job among $\{A_1, A_2, \cdots, A_{n_i}\}$ (ladder A), and receives lifetime value A_i. The employed worker climbs up one rung at a time through non-productive promotion or multiple rungs at a time through job turnover. Also, he switches to the next higher-valued ladder through productive promotion. For example, the worker at A_0 moves to B_0 through productive promotion when he accumulates human capital. The number of ladders is determined by n_j, while the number of rungs in each ladder is given by n_i. Lemma 3 says that the gap between two neighboring rungs is s, and Lemma 4 specifies gains from productive promotion.



[Figure 2]

Given $\{F_j\}_{j=1}^{n_j}$, we can get U_1 and $E_1(m^0)$. Then, Lemma 3 and 4 determine $\{E_j(m^i)\}$ for all $y_j \in \mathcal{Y}$ and for all $m^i \in \mathcal{M}$. All workers with contract m^{n_i} do not search another job because there is no better offer. Then, the operating firm with contract m^{n_i} sets $\eta^{n_i}(\cdot) = 0$. Then,

$$w^{n_i}(y_{n_j}) = u^{-1}((r+\rho+\delta)E_{n_j}(m^{n_i}) - \delta U_{n_j}), \quad \text{and} \tag{8}$$

$$w^{n_i}(y_j) = u^{-1}((r+\rho+\delta)E_j(m^{n_i}) - \delta U_j - \mu E_{j+1}(m^{n_i})), \quad \forall j = 1, 2, \cdots, n_j - 1 \tag{9}$$

Plugging these into equation (5) yields $\{J_j(m^{n_i})\}_{j=1}^{n_j}$. When $i < n_i$, given $\{(E_j(m^i), E_j(m^{i+1}), J_j(m^{i+1}))\}_{j=1}^{n_j}$, we can solve for $w^i(\cdot)$, and $\eta^i(\cdot)$ backward. Then, plugging $w^i(\cdot)$, and $\eta^i(\cdot)$ into equation (5) leads $\{J_j(m^i)\}_{j=1}^{n_j}$. Repeating this argument, we can solve for all $\eta^i(\cdot)$, which gives us $\{G\}$. In turn, given G, we can update $\{F_j\}_{j=1}^{n_j}$ using condition (6). This completes a fixe point argument.

Now, we examine when we can remove the restriction on the firm's strategy set without affecting the equilibrium outcome. It is trivial to show that if $J_1(m^{n_i}) > 0$, and $J_j(m^i)$ is increasing in y_j, it is not profitable to deviate by pooling any types. Let $\pi_j^i := (\lambda G_j(E_j(m^i) - s) + \lambda_u G_j(U_j))J_j(m^i)$, which means the expected profit when a recruiting firm posting m^i meets a y_j-type worker. A problem may occur when the firm pays more information rent to a certain type and attracts more workers of the type. Indeed, a 'least cost' incentive compatible contract may not be an optimal. However, it never occurs if π_1^i is increasing in i, and π_j^i is decreasing in i for $j = 2, 3, \cdots, n_j$. Lemma 5 proves this.

Lemma 5 (sufficient condition) It is optimal for firms to offer a least cost incentive compatible contract if the following conditions hold.

(i) $J_1(m^{n_i}) > 0$, and $J_j(m^{n_i})$ is increasing in j.

(ii) π_1^i is increasing in i, and π_j^i is decreasing in i for $j = 2, 3, \cdots, n_j$.

Although we provide a fixed point algorithm to find a market equilibrium above, it does not guarantee the existence of an equilibrium. Instead of theoretical proof, we report that in many numerical simulations, we obtained a unique fixed point. Moreover, we can say that our final estimates are consistent with the sufficient conditions in Lemma 5.

3 Data

We use data from the 1979 National Longitudinal Survey of Youth (NLSY79), which contains weekly work records from 1978 through 2006. The model implies that workers receive different wages based on their job tenure and level of human capital. Moreover, human capital is accumulated only on the job. To estimate the model, we need to keep track of whole work histories from their first full time job and job tenure of each job from the starting date. NLSY79 is well suited to analyze careers because of its rich work history. In particular, NLSY79 reports weekly labor force status from the high school period, which enables us to investigate the whole work history of individual workers from their first jobs. Also, NLSY79 keeps track of five jobs in each survey round. It reveals the starting date of a new labor contract more accurately than other data sets.

The survey consists of individuals who were 14-22 years old in 1979. Among those workers, we construct the sample with white male high school graduates, which is the largest demographic group in NLSY79. The sample includes individuals who completed 12th grade or received the equivalent degree(GED) at their age 17-19 after 1978, and who have never reported more than 12 years of education until the most recent survey. The reason that we put the age restriction is to rule out individuals who have long (full time) work experience before graduation. Workers who graduated before the survey started are dropped, because we cannot check what they did immediately after graduation. We also discard individuals who enrolled for the military service because their experience and work decision is different from others. Following this selection rule, we start with 773 individuals in our sample.

Following Farber and Gibbons (1996) and Yamaguchi (2009), if the worker holds any full time jobs for more than half of three consecutive years for the first time, we assume that the individual worker makes transition from school to work. A full time job is defined by a job at which the worker worked for more than 30 hours per week in average. We keep track of the work history of each individual from the first transition.[9] Note that by construction, all individuals start their career as an employed worker in my sample. To mitigate any risk of potential bias, we also ignore the first unemployment period before the first job in our simulation.

The model assumes that there is neither recalled jobs no returned workers. However, in NLSY79, workers frequently returned to their former jobs after leaving for some period. If it was planned by both parties in advance such as unpaid vacation, hospitalization, the two jobs, the formed job and the recalled job, should be considered as one job, because the previous labor contract already considered his return. The new contract after returning is also affected by the previous contract. If it is not planned, it should be considered two different jobs, because the fact that he returned affects neither the previous labor contract nor the new labor contract. To distinguish these, following Pavan (2008), we assume that if the intermediate period is sufficiently short, it is more likely to be planned. If the worker returned to a previous job within one quarter, we drop the intermediate work history and connect the two jobs as one continued job. Otherwise, we consider them as two different jobs. This consideration drops 923 short-term (less than one quarter) non-employment and 84 temporary jobs under the name of 'planned return'. In 555 cases, workers return to an old job after one quarter. Thus, our sample contains 4,325 employer-employee matches[10] and 4,880 jobs.

We keep track of individual workers in terms of non-employment, hours worked, job tenure, employment tenure, work experience, market experience, wage and re-employment wage. First, we consider non-employment as the periods in which the worker still stays in the survey but does not report any full time job. The periods reported as part time job, out of labor force, no information, and unemployment are recoded as non-employment, which is the counterpart of unemployment in the model. As for hours worked, we use hours worked per week. If it is not available, we calculate it through 'hours worked per day' times five working days per week.

Job tenure is defined by the length of a continuous working period within one employer. Employment tenure means the duration of consecutive job tenure. The difference follows from job-to-job transition. If a worker switches to a new job from an old

[9]We discard all work history ended before high school graduation since it is hard to think that jobs before and after graduation are homogeneous in terms of work decision, wage payment, and experience accumulation.

[10]NLSY79 does not distinguish 'job' from 'employer-employee match'. Therefore all returning cases are considered as one job.

job, job tenure is reset, but employment tenure continues. However, to determine job-to-job transition is not clear in the actual sample. Workers may have short term vacations before switching to new jobs, even though they made the switching decision on the old job. Thus, we discard the short term non-employment spells between two different jobs if the non-employment spell is less than three weeks. These cases are more likely to be an outcome of job-to-job transition rather than employment-unemployment-employment flow. This selection integrates 1702 short term non-employment into the subsequent job tenure.

The model assumes that a worker accumulates human capital only on the job. Hence, we should distinguish work experience from market experience. We refer to worker experience the sum of all employment spells. Market experience is calculated by subtracting the age at the entry from the current age of the worker.

In NLSY79 data set, wages are reported at the interview date and at the end date if the job was ended. In addition, they asked the first wage on the job from 1985 survey. Also, if the worker started a job before 1985 and kept the job until the 1985 survey, the first wage of the job was reported. In that sense, the first wage data might be biased. To mitigate the potential bias, in simulation we use the first wage only when the job started after or continued until the 1985 survey. Then, among the first wages reported, we define reemployment wage as the first wage after non-employment. Since there is no depreciation of human capital in the model, the reemployment wage is a key variable to estimate the effect of human capital accumulation. In our sample, we have 13,735 wage observations. These include some observations with potential coding errors.[11] Moreover, it is hard to fit all data points (especially data points at both ends) using a simple model. Hence we discard both the top and bottom 2.5% and focus on the remaining 95% of wage observations. This choice sets the lowest wage at \$2.84 per hour and the highest wage at \$16.31 per hour.

The finalized sample contains 665 individuals, 4325 employer-employee matches, 4880 jobs, and 14298 observations. Details of the construction of the data set are contained in appendix.

[11]For example, the lowest wage reported is \$0.03 per hour and the highest wage reported is \$862.69 per hour (after adjustment by monthly CPI).

4 Estimation

4.1 Estimation Procedure

We use indirect inference as in Bagger, Fontaine, Postel-Vinay, and Robin (2006), because maximum likelihood inference is not numerically feasible. Indirect inference requires that the structural model replicates the true data generating process in terms of some target moments given a true value of the structural parameter vector θ_0. Denote by $g(\theta)$ the vector of the target moments simulated with parameter vector θ. To estimate θ, we minimize the distance between the set of the sample moments from NLSY79 and the set of the moments from our simulations (auxiliary model). We simulate and calculate the moment vector k times and take their average. Then, the simulated moments estimator of θ_0 is defined as

$$\hat{\theta} = \arg\min_{\theta} \quad (\overline{g}_k(\theta) - g(\theta_0))^T \hat{W}_n(\overline{g}_k(\theta) - g(\theta_0)),$$

where $\hat{W}_n$ is a positive definite matrix that converges in probability to a deterministic positive definite matrix W. For this, we use the inverse of the covariance matrix of the auxiliary statistics. We estimate the covariance matrix of the auxiliary statistics by re-sampling 665 number of individuals with replacement 1,000 times (n=1000), and take the inverse of it. If a particular individual i is selected, his entire wage and employment history are included in the sample. For each set of simulated moments, we repeat the simulation 200 times and take the average of the moments from each simulation (k=200).

Although we do not have any theoretical evidence on the uniqueness of the minimum value, we minimize the objective function by using both Nelder-Mead method and the simulated annealing algorithm. First, we use the Nelder-Mead method repeatedly. If the program reaches to a local minimum point, we reset the size of simplex and restart from the local minimum point. If the program stops at a point sufficiently closed to the local minimum, we start the simulated annealing method. Although it requires heavy computation burden, we can increase the probability that we reach to a global minimum point by applying the simulated annealing method repeatedly. We repeat this process with four[12] different starting points. If we get the same estimates of the structural parameters, it can be understood as a global minimizer.

[12]Actually, we need more than four.

4.2 Estimation Specification

For our empirical implementation, we assume CARA (exponential) utility with a risk aversion parameter γ.

$$u(w) = -\exp(-\gamma w)$$

The level of human capital is discretized into 7 levels ($n_j = 7$). We set $y_1 = 0.4$, $y_{n_j} = 1.0$ and $\Delta = 0.1$. The number of equilibrium contracts are fixed to 20 levels and we set $s = 0.01$. In our sample, the highest wage is almost eight times larger than the lowest wage. Our choice makes the highest wage eight or nine times larger than the lowest wage depending on parameter values. We fix interest rate r at 0.012.

Also, it is hard to estimate the arrival rate of retirement shock ρ based on NLSY79 data set, because of the short history of NLSY79 data set. We assume that average worker stays in the labor market for 40 years, which fixes ρ at 1/160 to set the average duration equal to 160 quarters. Instead, to match the actual survival probability, we introduce 'attrition probability' in each survey round. We assume that although workers stay in the labor market, the survey loses some of them with the attrition probability as time goes on. The implied attrition probability per each survey round is 2.5%.

To solve the model, we use the reverse shooting algorithm, which substantially saves the computation time. The reverse shooting algorithm requires the end points. In our simulation, we consider the highest equilibrium wage w_{max} as a structural parameter to be estimated, and consider b an equilibrium outcome rather than a parameter to be estimated. If the equilibrium wage offer distribution is unique, w_{max} and b have a exact one-to-one relationship. Therefore, once we estimate w_{max} and other structural parameters, we can recover b. There are some advantages in this method. First, when $i = n_i$, we simply get $\{w_j^{n_i}, \eta_j^{n_i}, J_j^{n_i}\}_{j=1}^{n_j}$ regardless of $\{F_j\}_{j=1}^{n_j}$. Second, the non-negativity constraints of $\{\{\eta_j^i, J_j^i\}_{j=1}^{n_j}\}_{i=0}^{n_i-1}$ and the monotonicity constraints of value $\{\{J_j^i\}_{j=1}^{n_j}\}_{i=0}^{n_i-1}$ are automatically satisfied, once they are satisfied at $i = n_i$. Therefore, by choosing w_{max} properly, we can ignore the constraints.

4.3 Estimation

We have seven structural parameters to be estimated: Four Poisson arrival rate parameters, $(\delta, \lambda_u, \lambda, \mu)$, risk aversion parameter γ, the flow of unemployment benefit b (or w_{max}), and cost function parameters c_1.

First, to capture the dynamic flow of workers, we use the average nonemployment spell, the average job spell, and the average length of unemployment in the first five years. The model implies that as workers accumulate human capital, they are promoted at a faster rate and job turnover is more likely to happen among young workers with less human capital. Thus, we examine the total nonemployment(or employment) period in the first five years. The sample reports that the average unemployment duration is 0.471 year, job spell is 2.175 years, and the average worker keeps a full time job during 88.3% of the first five years.

Second, one of main task in this empirical study is to estimate the effect of human capital accumulation separately from the effect of strategic promotion and job turnover. Since our model implies that human capital is neither depreciated nor accumulated when unemployed, we take advantage of the reemployment wage which is defined by the first wage after unemployment. We regress log reemployment wage($\hat{w}$) on the work experience,

$$\hat{w}_k = \beta_0 + \beta_1 \times work\ experience_k + \varepsilon_k,$$

where ε_k is a statistical residual. We adopt β_1 as our auxiliary moment, which captures the wage growth due to work experience accumulation. However, the regression coefficient by itself is not sufficient to distinguish how frequently the human capital accumulation shock arrives and how large each shock is. In our model, human capital accumulation occurs at rate μ and it increases workers' wages by a ceratin amount, which is affected by c_1. To capture the frequency and the magnitude of each shock separately, we also take advantage of information on the re-employment wage distribution. From the sample, we calculate the ratio of the 3rd quartile to the 1st quartile of the distribution, and the 2nd quartile to the 1st quartile. In the sample, the auxiliary regression indicates a coefficient β_1 of 0.109 and the two quartile ratios are 1.775 and 1.281, respectively.

To capture the slope of the wage-tenure profile, we regress wages reported in the first five years ($\tilde{w}$) on market experience.

$$\tilde{w}_k = \alpha_0 + \alpha_1 \times market\ experience_k + u_k.$$

We adopt α_1 as one auxiliary moment. The reason that we focus on the wages reported in the first five years is that the promotion rates are so different depending on the level of human capital. We want to focus on a narrow and identical group to capture the slope more accurately.

Finally, to capture overall wage growth (or wage-age profile), we put some additional auxiliary moments. Denote by $w01$ the first wage reported within the first 6 months after the transition to work. Also denote by $w05$, $w10$, and $w20$ the average of wages reported first after 5 years, 10 years, and 20 years of market experience, respectively. Then, we take the ratios of $w05/w01$, $w10/w01$, and $w20/w01$. In the sample, the regression coefficient α_1 is 0.052. The ratios, $w05/w01$, $w10/w01$, and $w20/w01$ are 1.430, 1.616, and 1.881, respectively. The auxiliary moments from the sample and the bootstrapping standard errors are summarized in the second column of [Table 1]. It also reports the estimates of corresponding moments from the simulation based on estimates of structural parameters.[13]

Table 1: Auxiliary Moments

	sample moment	simulated moment
average unemployment duration (yr)	0.471 (0.013)	0.467(-)
average job duration (yr)	2.175 (0.044)	2.185 (-)
average unemployment periods in the first 5years	0.117 (0.004)	0.123(-)
$\Delta log(\tilde{w})$ / $\Delta work\ experience$	0.023 (0.002)	0.025 (-)
3rd/1st quartile ratio of reemployment wage dist.	1.775 (0.031)	1.774 (-)
2nd/1st quartile ratio of reemployment wage dist.	1.281 (0.019)	1.282 (-)
$\Delta log(w)$ / $\Delta market\ experience$	0.052 (0.004)	0.054 (-)
$w20/w01$	1.881 (0.042)	1.913 (-)
$w10/w01$	1.616 (0.033)	1.645 (-)
$w05/w01$	1.430 (0.026)	1.478 (-)

*Standard errors in the second column are estimated using bootstrap.

[Table 2] reports the estimates of the structural parameters.

Table 2: Parameter Estimation

Parameter	Estimates	std. error
δ(separation shock)	0.091	-
λ_u(offer finding rate by unemployed workers)	0.580	-
λ(offer finding rate by employed workers)	0.446	-
μ(human capital accumulation shock)	0.023	-
w_{max}	0.940	-
c_1(cost parameter)	0.302	-
γ(risk aversion parameter)	0.450	-

4.4 Counterfactual Analysis

In this section, we conduct a counterfactual experiment to understand how human capital accumulation contribute to wage growth. The counterfactual experiment is

[13]The asymptotic standard errors will be reported soon.

designed to shows that how much a representative worker would earn if he were not able to accumulate any human capital. To this end, we need to keep all players' strategies same. As before, firms optimally choose their strategies assuming that workers stochastically accumulate human capital. But, it is assumed that the worker just stay in the same state when he is hit by the human capital accumulation shock. We repeat this experiment with 665 workers and construct a artificial data set.

Table 3: Counter Factual Analysis

		w_5/w_1	w_{10}/w_1	w_{20}/w_1
data		1.430	1.614	1.881
estimation	with human capital accumulation	1.478	1.645	1.913
	without human capital accumulation	1.350	1.416	1.418

[Table 3] compares the average wage growth of two groups. It shows that the average wage grows by 43%, 61.4% and 88.1% in the first 5 years, 10 years, and 20 years, respectively. In our estimation, it grows by 47.8%, 64.5% and 91.3%, respectively. Without human capital accumulation, the average wages grows partly due to non-productive promotion and partly due to job turnover. The growth rates without productive promotion are reported as 35%, 41.6%, and 41.8%. Based on the counterfactual analysis, we can infer that the human capital accumulation increases 12.8% in the first five years, 22.9% in the first 10 years, and 49.5% in the first 20 years.

5 Conclusion

The main contribution of this paper is to develop and estimate an equilibrium job search model with unobserved human capital to investigate the effect of human capital accumulation on individual wage dynamics using . In equilibrium, there are multiple wage-ladders that a worker can climb up or switch from one to another. After estimating the model, through a counterfactual experiment, we show that if a typical worker were not able to accumulate human capital, his wage would grow by 42%.

Our next aim is to add ex ante heterogeneity on the worker side to the framework developed in this paper. The model proposes that the employment contract specifies the lifetime value received by the worker as a function of what he actually produces. Therefore we can easily extend the model in this direction. However, estimation would not be an easy task. If the distribution of first wages on the first jobs is available, we can estimate the distribution of human capital endowments using it. Unfortunately, the NLSY79 data did not collect the first wages until the 8th survey rounds in 1986. For this reason, we have not included the analysis on the ex ante heterogeneity case.

We can also think of a match specific productivity component. This extension could yield many interesting results. First, this may introduce equilibrium wage cuts in job-to-job transition. Due to non-productive promotion, a higher value job does not always pay a higher wage. Second, this would also introduce efficient job turnovers. In our model, job turnover is inefficient because all jobs are homogeneous. This leads to a low turnover rate and a faster (non-productive) promotion rate for more productive workers. By introducing a match specific productivity, we can stimulate efficient job turnover and lower the average non-productive promotion rate. However, it requires another dimension in the state space. Therefore, we leave this for future research.

References

Altonji, J., and R. Shakotko (1987): "Does Wages Rise with Job Seniority?," *Review of Economic Studies*, 54(3), 437–459. 4

Altonji, J., and N. Williams (2005): "Does Wages Rise with Job Seniority? A Reassessment," *Industrial and Labor Relations Review*, 58(3), 370–397. 4

Bagger, J., F. Fontaine, F. Postel-Vinay, and J. Robin (2006): "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation," Manuscript, December 2006. 17

Burdett, K., C. Carrillo-Tudela, and M. G. Coles (2008): "Human Capital Accumulation and Labor Market Equilibrium," Manuscript, June 2008. 3

Burdett, K., and M. Coles (2003): "Equilibrium Wage-Tenure Contracts," *Econometrica*, 71(5), 1377–1404. 2, 7, 31

Burdett, K., and D. T. Mortensen (1998): "Wage Differentials, Equilibrium Size, and Unemployment," *International Economic Review*, 39(2), 257–273. 2

Delacroix, A., and S. Shi (2006): "Directed Search on the Job and the Wage Ladder," *International Economic Review*, 47(2), 651–699. 3

Farber, H., and R. Gibbons (1996): "Learning and Wage Dynamics," *Quarterly Journal of Economics*, 111(4), 1007–1047. 14

Pavan, R. (2008): "A flexible Model of Individual Wage Dynamics and Job Mobility Outcomes," August 2008, University of Rochester. 15

Shi, S. (2009): "Directed Search for Equilibrium Wage-Tenure Contracts," *Econometrica*, 77(2), 561–584. 31

Stevens, M. (2004): "Wage-Tenure Contracts in a Frictional Labor Market: Firms' Strategies for Recruitment and Retention," *Review of Economic Studies*, 71(2), 535–551. 2

Topel, R., and M. Wald (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107(2), 439–479. 2

Yamaguchi, S. (2009): "Job Search, Bargaining, and Wage Dynamics," (2007-03). 14