Abstract

This paper develops an equilibrium job search model in which the employed worker privately accumulates human capital and continually searches for a better paying job. Firms make wage payment based on both workers’ output and job tenure in order to encourage production and to discourage job turnover. In this environment, wages grow due to human capital accumulation (productive promotion), and the strategic back-loading scheme (non-productive promotion) as well as job-to-job transition. I estimate the model using indirect inference to investigate the effect of human capital accumulation on individual wage growth. In the NLSY79 data, the average wage of white male high school graduates after 20 years of market experience is 1.88 times larger than the average of the first full-time wages. A 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 41.8%.

JEL Classification: D82, E24, J31, J41
1 Introduction

In the U.S. labor market, a typical male\(^1\) worker works for 40 years, and his wage doubles over his career.\(^2\) In general, this wage growth is understood as the outcome of human capital accumulation after his entry into the labor market.\(^3\) However, given that human capital accumulation is neither sufficient nor necessary condition for wage growth in the frictional labor market\(^4\), it is necessary to develop a structural model to quantify returns to human capital and identify other relevant factors. This paper establishes and estimates an equilibrium job search model by focusing on firms’ strategic responses to the worker’s learning-by-doing and on-the-job search behavior.

In reality, each worker privately accumulates human capital through learning-by-doing and continually searches for a better paying job. Firms which cannot monitor learning-by-doing process nor job search outcome, try to provide right incentives by rewarding the worker having better performance and longer tenure through bonus payments and long-service allowances. Motivated by this, I build up a market equilibrium in which firms post different lifetime values for different output levels to encourage production, and deliver the committed values through back-loading wage schedules to discourage job turnover.\(^5\) Hence, there are multiple wage-ladders and each worker chooses one of them depending on his level of human capital. He gradually climbs his own wage ladder along the back-loading wage schedule (non-productive promotion), jumps up to a higher rung through job-to-job transition, and switches to a new ladder due to human capital accumulation or depreciation (productive promotion).

In their seminal work, Burdett and Mortensen (1998) develop a wage-posting model with on-the-job search in which some firms post high wages to attract a worker from other firm and workers jump up the wage ladder only through job-to-job transition. Departing from it, Burdett and Coles (2003) build up an equilibrium in which all firms post a common wage-tenure schedule and different firms choose different starting points on it. In their equilibrium, firms optimally back-load some portion of wage

---

\(^1\)Throughout the paper, I use the masculine pronoun for a worker because I deal with a male sample in the empirical application.

\(^2\)Topel and Ward (1992) report this fact in cross section data. Based on my sample of white male high school graduates in the 1979 National Longitudinal Survey of Youth, I also find that the average wage after 5, 10, and 20 years of market experience is 1.43, 1.61 and 1.88 times larger than the average of first full-time wages, respectively.

\(^3\)See Altonji and Shakotko (1987). They argue that returns to experience (general human capital), identified by wage growth transferred to the next job, take lion’s share in wage growth, while either job turnover or returns to tenure (job specific human capital), identified by wage growth non-transferable to the subsequent job, has limited effects on individual wage growth. However, their classification relies on the implicit assumption that the wage schedule on a job rises with and only with human capital, which is not consistent with the existence of (endogenous) job-to-job transition under labor market friction.

\(^4\)supported by the existence of unemployment and sequential job turnover

payments to discourage job turnover and to extract more surplus from early leavers. Wages grow due to job turnover and the back-loading schedule. As another attempt, Burdett, Carrillo-Tudela, and Coles (2009) extend the Burdett and Mortensen (1998) framework by adding human capital accumulation. They assume that each worker accumulates human capital through learning-by-doing on the job, and the firm pays for him following a single wage-output ratio (the piece rate sharing rule). In their model, wages grow due to job turnover and learning-by-doing. However, in their study, it is unclear why each firm sticks to a single piece rate.

Recently, Bagger, Fontaine, Postel-Vinay, and Robin (2006) study wage dynamics by combining learning-by-doing on the job with the ex post offer matching framework proposed by Postel-Vinay and Robin (2002). They assume that if an employed worker finds another recruiting firm, the existing firm and the recruiting firm bid new piece rates to attract him. He accepts the offer with the higher lifetime value. Their framework yields a tractable model with consistent implications for the wage earnings distribution and individual wage dynamics. However, it is also vulnerable to the criticism that the ex post offer matching gives workers the wrong incentives to search for an outside offer.\(^6\)

This paper incorporates private learning-by-doing into the equilibrium model by Burdett and Coles (2003). The private learning-by-doing process gives birth to two types of upward pressure in wage payments. First, it may be optimal for the firm to commit a higher value for better performance to induce truthful revelation (internal pressure). Second, since a more productive worker is more attractive to other poaching firms, the firm should pay more to retain the worker (external pressure). If the internal pressure dominates the external pressure, the incentive compatibility constraint should be binding, which together with anti-discrimination legislation\(^7\) determines the optimal productive promotion schedule after human capital accumulation. If the former is dominated by the latter, the incentive compatibility constraint (hence non-observability) should be slack. Then, besides the anti-discrimination legislation, additional restriction such as the piece rate sharing rule is required in determining the optimal productive promotion schedule. Based on the fact that learning-by-doing process is unobserved by firms, this paper focuses on the case in which the internal pressure dominates the external pressure.

In the empirical analysis, I estimate the model and perform some counterfactual experiments to ask what would happen if a typical worker were not able to accumulate

---


\(^7\)The anti-discrimination legislation in this paper means that a firm cannot pay different values to two workers having the same performance and the same job tenure, which is a more relaxed assumption than in Burdett, Carrillo-Tudela, and Coles (2009). They assume that a firm should pay the exactly same piece rate to all its workers regardless of their job tenure. Basically, this anti-discrimination legislation prohibits firms from back-loading the whole value and extracting all surplus until human capital accumulation occurs.
human capital. To do this, I construct a sample of white male high school graduates from the 1979 National Longitudinal Survey of Youth, and keep track of them in terms of non-employment, hours worked, job tenure, employment tenure, work experience, market experience, wage and reemployment wage. Then, I estimate the structural parameters of the model using indirect inference. The model implies that human capital accumulation has a permanent effect on wage growth, while the effect of job tenure is reset once the worker becomes unemployed. Hence, I take advantage of the difference between usual wage growth and reemployment wage\(^8\) growth to capture the effect of human capital accumulation (productive promotion) separately from that of job tenure.

In the sample, the average wage after 20 years of market experience is 88% higher than the average of first full-time wages. The counterfactual analysis reports that without human capital accumulation, wages would grow by 41.8%. This result suggests that returns to human capital is not so large as in Altonji and Shakotko (1987) and Altonji and Williams (2005). The limited effect of human capital accumulation is also consistent with the fact that in the sample, the estimated slope coefficient in the reemployment wage-experience regression is almost a half of the coefficient in the normal wage-experience regression.

The paper proceeds as follows. In section 2, I build up the theoretical model and characterize the equilibrium we are interested in. In section 3, I construct the sample and define relevant variables, and in section 4, I provide the estimation protocol and results. Section 5 concludes. All proofs and data construction are given in the Appendix.

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. While firms are homogeneous and infinitely-lived, each worker privately accumulates human capital through learning-by-doing, and retires at rate \(\rho\). 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:

\[
\mathbb{E}_0 \left[ \int_0^\infty e^{-(\rho+r)t} [u(w(t; m)) - c(t)] dt \right],
\]

\(^8\)the first wage after unemployment
where \( u' > 0 \) and \( u'' < 0 \). He can neither borrow nor save. At every instant, an unemployed worker collects unemployment benefits \( w = b \) at no cost \( c = 0 \), while an employed worker produces output at cost \( c \) and receives a wage payment \( w(t;m) \) specified in the labor contract \( m \).

The unit of human capital is normalized so that it represents the units of output that the worker can produce. Each worker starts his career with \( y_1 \) units of human capital, and privately accumulates it by \( \Delta \) units at each learning shock with Poisson arrival rate \( \mu \).\(^9\) Also, he experiences human capital depreciation by \( \Delta \) units at rate \( \eta \). For expositional convenience, I call the worker with \( y_i \) units of human capital ‘\( y_i \)-type worker’, where

\[
y_i = y_1 + (i - 1)\Delta, \quad \text{for } i = 1, 2, \ldots, n.
\]

The type is private information of each worker. If a \( y_i \)-type worker produces \( y' \) units of output, his flow cost is given by

\[
c(y';y_i) = \begin{cases} 
\alpha_0 - \alpha_1(y_i - y') & \text{if } y' \leq y_i \\
\infty & \text{otherwise}
\end{cases},
\]

where \( 0 < \alpha_1 < 1 \). The cost function reflects that the disutility from working is proportional to hours worked. A \( y_i \)-type worker producing \( y_i \) units of output incurs disutility of \( \alpha_0 \). But if a \( y_i \)-type worker decides to produce \( y' (< y_i) \) units, he can finish his job earlier and save his disutility through leisure. The private benefit from misreporting is captured by \( \alpha_1(y_i - y') \). Note that in the market equilibrium, there is no efficiency loss due to the information asymmetry in production. The parameter \( \alpha_1 \) only affects how the surplus is split.

An unemployed worker finds a job offer at rate \( \lambda_u \), loses \( \Delta \) units of human capital at rate \( \eta \), and retires at rate \( \rho \). Let \( U_i \) denote the equilibrium asset value for a \( y_i \)-type unemployed worker. Also, let \( F_i(\cdot) \) denote the equilibrium distribution of lifetime values offered to \( y_i \)-type workers by recruiting firms. The HJB equation for the \( y_i \)-type unemployed worker is given by

\[
rU_i = u(b) + \lambda_u \int \max\{x - U_i, 0\} dF_i(x) - \rho U_i + \eta(U_{i-1} - U_i), \quad (1)
\]

where it is assumed that \( U_0 = U_1 \).

---

\(^9\)One may argue that human capital accumulation occurs more frequently among less productive workers. For example, I could assign \( (\mu_1, \mu_2, \ldots, \mu_{n-1}) \) instead of the unique \( \mu \). But, to keep the model simple, I assume that it is constant. Interestingly, the model predicts concave average wage growth in spite of the linear learning curve with a constant rate.
An employed worker finds another job offer at rate $\lambda \in (0, \lambda_u)$, accumulates $\Delta$ units of human capital at rate $\mu$, loses it at rate $\eta$, separates from his job at rate $\delta$, and retires at rate $\rho$. Denote the output schedule chosen by a $y_i$-type worker, by $\tilde{y}_i : [0, \infty) \to \{y_1, y_2, \ldots, y_n\}$, which represents a mapping from the space of job tenure to the set of output levels. The expected lifetime value of the $y_i$-type employed worker at time $t$ choosing output schedule $\tilde{y}_i$ under contract $m$, $E_i(t; \tilde{y}_i, m)$ is given by

$$E_i(t; \tilde{y}_i, m) = \int_t^\infty \phi_i(s; \tilde{y}_i, m) \left[ u(w(s; \tilde{y}_i, m)) - c(\tilde{y}_i; y_i) + E[ E'_i | E_i(s; \tilde{y}_i, m) ] \right] ds,$$

where

$$\phi_i(s; \tilde{y}_i, m) = e^{-\int_s^t \left[ r + \delta + \lambda(1 - F_i(E_i(t; \tilde{y}_i, m))) + \mu + \eta \right] dt}, \quad \text{and}$$

$$E[ E'_i | E_i(s; \tilde{y}_i, m) ] = \delta U_i + \lambda \int_{E_i(s; \tilde{y}_i, m)}^\infty x dF_i(x) + \mu E_{i+1}(s, m) + \eta E_{i-1}(s, m).$$

As before, let $E_0(s, m) = E_1(s, m)$ and $E_{n+1}(s, m) = E_n(s, m)$. A $y_i$-type employed worker under contract $m$ chooses and updates his own production schedule at every instant to maximize his lifetime value. His expected value at time $t$ is characterized by

$$\max \{ E_i(t; \tilde{y}_i, m) \mid \tilde{y}_i \in \mathcal{Y} \}.$$

If he truthfully produces $y_i$ at each instant, the HJB equation for the $y_i$-type employed worker is given by

$$rE_i(t, m) = u(w_i(t, m)) - c(y_i; y_i) + \bar{E}_i(t, m) + \lambda \int \max \{ x - E_i(t, m), 0 \} dF_i(x) - \rho E_i(t, m)$$

$$+ \delta (U_i - E_i(t, m)) + \mu (E_{i+1}(t, m) - E_i(t, m)) + \eta (E_{i-1}(t, m) - E_i(t, m)). \quad (2)$$

A firm can create and maintain one vacant job. The firm with a vacancy posts an incentive compatible menu of contracts to attract workers. Once the menu of contracts is accepted by a worker, the firm together with the worker starts producing immediately. It is assumed that a firm should offer the same wage to its workers having the same performance and the same job tenure. The anti-discrimination legislation is borrowed from Burdett, Carrillo-Tudela, and Coles (2009) to avoid corner solutions. It prohibits firms from committing the largest value after human capital accumulation and extracting more surplus before it, which results in a huge mass at the upper bound of the support of the wage-earning distribution and violates the equilibrium condition.

The private learning-by-doing process, under the anti-discrimination legislation,
gives birth to two types of upward pressure in wage payments. First, it may be optimal for the firm to pay the worker a high value conditioning on output to induce truthful revelation (internal pressure). Second, since a more productive worker is more attractive to other poaching firms, the firm should pay a high value to retain the worker (external pressure). Suppose that the former is dominated by the latter. Then, the incentive compatibility constraints are not binding and information asymmetry in production plays no role, which requires another optimization in choosing the continuation value. However, the existence of the interior solution is not guaranteed. In most cases\textsuperscript{11}, the offer distribution is convex so that the optimal choice on the continuation value may end up with corner solution. To avoid this, Burdett, Carrillo-Tudela, and Coles (2009) assume that firms should offer a constant piece rate regardless of the worker’s productivity. If the former dominates the latter as assumed in this paper, the incentive compatibility constraint should be binding, which presents an interior solution.

Contract \( m \) specifies the action profile (or ‘terms of trade’) which consists of the output schedule by the worker and the lifetime values delivered by the firm under the truthful revelation assumption. Then, \( m \) is characterized by \( \{(y_i, E_i(\cdot, m))\}_{i=1}^n \).

**Definition** A contract, \( m \), is incentive compatible for \( y_i \)-type workers if

\[
E_i(t; m) \geq \max \{ E_i(t; \tilde{y}_i, m) \mid \tilde{y}_i \in \mathcal{Y} \cap \{y_i\}^c \} \quad \text{for any} \quad t \in [0, \infty). \tag{3}
\]

In particular, when the contract is incentive compatible for all types, I say that it is incentive compatible.

As a tie-breaking rule, it is assumed that when a \( y_i \)-type worker is indifferent, he truthfully produces \( y_i \) at every instant. Let \( \mathcal{M}_i \) and \( \mathcal{M} \) be the set of contracts incentive compatible for \( y_i \)-type workers and contracts incentive compatible for all types, respectively.

**Definition** A contract, \( m \in \mathcal{M}_i \), is least cost incentive compatible for \( y_i \)-type workers if for any \( m \in \mathcal{M}_i \) with \( E_1(t; m') = E_1(t; m) \),

\[
E_i(t; m) \leq E_i(t; m') \quad \text{for any} \quad t \in [0, \infty). \tag{4}
\]

In particular, when the contract is least cost incentive compatible for all types, I say that it is least cost incentive compatible.

\textsuperscript{11}As in Burdett and Mortensen (1998), Burdett and Coles (2003) and Burdett, Carrillo-Tudela, and Coles (2009)
If \( y_i \) units are produced, the operating firm earns revenue \( y_i \) and makes wage payments from it. The match is destroyed when the worker leaves the job either voluntarily or involuntarily.\(^\text{12}\) Let
\[
\psi_i(t, s) := e^{-\int_t^s [\rho + \delta + \lambda (1 - F_i(E_i(\tau, m))) + \mu + \eta] d\tau}.
\]
Given the promised value \( \{ E_i \}_{i=1}^n \), the operating firm chooses the schedule of \((w(\cdot, m), E_{i-1}(\cdot, m), E_{i+1}(\cdot, m))\) to maximize the expected value
\[
J_i(t, E_i(t, m)) = \int_t^\infty \psi_i(t, s)[y_i(s) - w_i(s, m) + \mu J_{i+1}(s, E_{i+1}(s, m)) + \eta J_{i-1}(s, E_{i-1}(s, m))] ds
\]
subject to the set of incentive compatibility constraints and the set of promise-keeping constraints. The incentive compatibility constraints are presented in \((3)\). The promise-keeping constraints are described in \((2)\), which implies that once the firm commits \( \{ E_i(\cdot, m) \}_{i=1}^n \), it should deliver it through wage and promotion schedule.

Let \( u_i \) and \( G_i(x) \) be the proportion of \( y_i \)-type unemployed worker and \( y_i \)-type employed workers receiving the value less than \( x \), respectively. Then,
\[
\frac{\partial G_i(\cdot)}{\partial x} \geq 0, \quad \text{for each } i = 1, 2, \ldots, n, \quad \text{and} \quad \sum_{i=1}^n [u_i + G_i(E_i(0, m))] = 1.
\]
Denote by \( M^* \) the set of the equilibrium contracts. The equal profit condition implies that
\[
\sum_{i=1}^n (\lambda G_i(E_i(0, m)) + \lambda_u u_i) J_i(t, E_i(0, m)) \begin{cases} = \pi, & \text{if } m \in M^*, \\ < \pi & \text{otherwise.} \end{cases}
\]

By aggregating all recruiting firms’ strategies, I get the distribution of lifetime values offered to each type, \( \{ F_i \}_{i=1}^n \). Suppose that the most generous recruiting firm and the least generous firm post \( \overline{m} \) and \( m \), respectively.\(^\text{13}\) To make the model tractable, I restrict my attention to the equilibrium which satisfies the following restriction.

**Eq’m Restriction** \textit{For all } \( E_i \in (E_i(0, \overline{m}), E_i(0, m)) \), \( F_i \) is continuously differentiable and satisfies \( F_i'(E_i) > 0 \).

\(^\text{12}\)In equilibrium, since all jobs yield positive expected profit to firms, there is no endogenous firing.

\(^\text{13}\)One may argue that some firms make the most generous offer to some types and others may do to other types. But the least cost incentive compatibility removes such cases. See Lemma 1.
Operating firms choose wage and promotion schedules, which determines the steady state distribution of expected lifetime values received by each type, \(\{G_i\}_{i=1}^n\). In turn, given \(\{G_i\}_{i=1}^n\), recruiting firms choose optimal contracts, which result in \(\{F_i\}_{i=1}^n\).

The equal profit condition (6) must hold. Consequently, the equilibrium is defined as follows.

**Definition** A market equilibrium requires:

(i) Given \(\{F_i\}_{i=1}^n\), a \(y_i\)-type unemployed worker accepts the contract \(\{(y_i, E_i(t, m))\}_{i=1}^n\) if and only if

\[
\max_{\hat{y}_i \in Y} \{E_i(t, \hat{y}_i, m)\} \geq U_i.
\]

(ii) Given \(\{F_i\}_{i=1}^n\), a \(y_i\)-type employed worker optimally chooses the level of output, determines whether he will search or not, and accepts a new contract \(m'\) if and only if

\[
\max_{\hat{y}_i} \{E_i(t; \hat{y}_i, m')\} \geq \max_{\hat{y}_i} \{E_i(t; \hat{y}_i, m)\}.
\]

(iii) Given \(\{F_i\}_{i=1}^n\), an operating firm with contract \(m\) optimally chooses \(\{(w_i^m, E_{i+1}^m(t; m))\}_{i=1}^n\) to deliver \(\{E_i^m(t; m)\}_{i=1}^n\).

(iv) Given \(\{G_i\}_{i=1}^n\) and a recruiting firm optimally posts contract \(m\) given the equal profit condition described in (6).

(v) The equilibrium distributions \(\{F_i, G_i\}_{i=1}^n\) are stationary.

### 2.2 Equilibrium Characterization

In this subsection, I characterize the equilibrium in which all least cost incentive compatible constraints are binding. To do so, I assume that firms should offer least cost incentive compatible contracts and numerically check whether they have incentives to deviate from their equilibrium strategy.

**Lemma 1** Suppose that \(m \in M\) is the contract delivering the lowest expected value to each type of workers. For any \(m \in M\),

(i) \(E_1(0, m) - E_1(0, \overline{m}) = E_2(0, m) - E_2(0, \overline{m}) = \cdots = E_n(0, m) - E_n(0, \overline{m})\) and,

(ii) \(F_1(E_1(0, m)) = F_2(E_2(0, m)) = \cdots = F_n(E_n(0, m))\).
The first statement says that least cost incentive compatible contracts are parallel to each other. That is, contract \( m \) pays an additional value to \( y_i \)-types than \( m' \), if and only if it pays the additional value to the other types. Also, it implies that the information rents are same across different least cost incentive compatible contracts. In particular,

\[
E_i(0, m) - E_{i-1}(0, m) = E_i(0, m') - E_{i-1}(0, m'), \quad \text{for each } i = 2, 3, \ldots, n
\]

The second statement implies that the probability of finding a contract delivering a smaller value than contract \( m \) are same across all types. These are based on pre-imposed equilibrium condition that firms have no incentive to screen out any type. Lemma 1 allows me to drop the subscript of the value offer distribution,

\[
F_1(E_1(0, m)) = F_2(E_2(0, m)) = \cdots = F_n(E_n(0, m)) =: F(E_1(0, m)),
\]

where \( F : [E_1(0, \underline{m}), E_1(0, \overline{m})] \to [0, 1] \).

**Lemma 2** In equilibrium,

(i) \( \min \{E_i(0, m) - U_i\} = 0 \), and

(ii) given \( \overline{m} \), no firm pays more than \( E_i(0, \overline{m}) \) for any type \( y_i \).

Lemma 2 presents the condition for \((m, \overline{m})\). The first statement implies that the least generous firm should offer the value of unemployment for at least one type in equilibrium. The second statement implies that once \( E_i(0, \overline{m}) \) by the most generous recruiting firm is given, no firms including recruiting and operating firms have incentives to pay more than the value. In turn, it implies that the optimal wage schedule for each type is bounded on equilibrium. Then, Lemma 3 and 4 jointly determine the size of information rent.

**Lemma 3** The following statements are equivalent.

(i) Contract \( m \) is least cost incentive compatible.

(ii) For each \( i = 2, 3, \ldots, n \),

\[
E_i(t, m) = E_i(t, y_{i-1}, m) \quad \text{at any } t \in [0, \infty).
\]  

(iii) For each \( i = 2, 3, \ldots, n \),

\[
u(w(t, y_i; m)) = u(w(t, y_{i-1}; m)) + \alpha_1 \Delta \quad \text{at any } t \in [0, \infty).\]
Lemma 4  Contract \( \overline{m} \) is least cost incentive compatible if and only if for \( i = 2, 3, \ldots, n \),

\[
(r + \rho + \lambda_u + \eta)(U_i - U_{i-1}) = \eta(U_{i-1} - U_{i-2}) + \lambda_u(\overline{E}_i(\overline{m}) - \overline{E}_{i-1}(\overline{m})),
\]

(9)

and

\[
E_i(0, \overline{m}) - E_{i-1}(0, \overline{m}) = \frac{\alpha_1 \Delta + \delta(U_i - U_{i-1}) + \mu(E_{i+1}(0, \overline{m}) - E_i(0, \overline{m})) + \eta(E_{i-1}(0, \overline{m}) - E_i(0, \overline{m}))}{r + \delta + \rho + \mu + \eta},
\]

(10)

where \( U_0 = U_1, E_0(0, \overline{m}) = E_1(0, \overline{m}) \) and \( E_{n+1}(0, \overline{m}) = E_n(0, \overline{m}) \).

[Figure 1] summarizes Lemma 1 - 4. If both \( m \) and \( m' \) are least cost incentive compatible, then \( \{(y_i, E_i(0, m))\}_{i=1}^n \) and \( \{(y_i, E_i(0, m'))\}_{i=1}^n \) are parallel to each other. The gap between \( E_i(0, m) \) and \( U_i \) should be zero at least one \( y_i \). But in many numerical experiments, \( E_1(0, m) = U_1 \) and \( E_{-1}(0, m) > U_{-1} \) as long as the rate of depreciation \( \eta \) is sufficiently smaller than the rate of accumulation \( \mu \). Give \( (F(\cdot), \{G_i(\cdot)\}_{i=1}^n) \) and \( \overline{m} \), the equal profit condition determines \( \overline{m} \). Lemma 3 and Lemma 4 jointly determines the marginal information rent. In [Figure 1], it is represented by the slope of the contract curve.

Now, I look into the strategy by the least generous firm. Given \( \overline{m} \), the firm chooses
an optimal wage schedule to deliver the committed value \( \{E_i(0, m)\}_{i=1}^n \) at tenure 0.

**Lemma 5** Given \( F \) and \( E_1(0, m) = E_0 \), the optimal schedules solve for

\[
\begin{align*}
\dot{J}_i &= -(p_i - w_i + \eta J_{i-1} + \mu J_{i+1}) + [r + \rho + \delta + \lambda(1 - F) + \mu + \eta]J_i \\
\dot{E}_i &= -u(w_i) + \alpha_0 + (r + \rho + \delta + \lambda(1 - F) + \mu + \eta)E_i - \delta U_i - \lambda \int_{E_i}^{E} xdF(x) - \mu E_{i+1} - \eta E_{i-1} \\
\dot{x}_i &= x_{i-1} + x_{i+1} \eta - x_i[r + \rho + \delta + \lambda(1 - F) + \mu + \eta], \\
\dot{w}_1 &= \left[ \sum_{i=1}^{n} \frac{u''(w_i)}{u'(w_i)^2} \cdot \frac{u'(w_1)}{u'(w_i)} \right]^{-1} \left[ \sum_{i=1}^{n} \frac{\dot{x}_i}{u'(w_i)} - x_i \lambda J_i \dot{F} + \frac{(r + \rho + \delta + \lambda(1 - F))x_i}{u'(u^{-1}(u(w_1) + (i-1)\Delta))} \right], \text{ and} \\
\dot{w}_i &= u^{-1}(u(w_{i-1}) + \alpha_1 \Delta)
\end{align*}
\]

subject to the boundary conditions:

\[ E_1(0, m) = E_0, \quad \text{and} \quad \lim_{t \to \infty} \{J_i, E_i, w_i, x_i\}_{i=1}^n = \{\overline{J}_i, \overline{E}_i, \overline{w}_i, 0\}_{i=1}^n. \]

I borrow the concept of the baseline salary scale from Burdett and Coles (2003). The optimal schedule presented by the system of differential equations in lemma 5 does not depend on the initial value \( E_0 \). The initial value \( E_0 \) determines only a starting point. It means that other firms posting \( E_1(0, m) > E_0 \) also move along the path. The baseline salary scale is important as it can be extended to prescribe the wage schedules offered by all recruiting firms in a steady state. Different recruiting firms just choose different starting points on the baseline salary scale.

[Figure 2] summarizes Lemma 1 through Lemma 5. Consider an \( y_2 \)-type employed worker under contract \( m \), denoted by \( A_0 \) in the figure. If he switches to a better paying job with contract \( m'' \), he jumps up to point \( B_0 \) within his value ladder. If he stays without any shocks, he gets a gradual promotion and will be found at \( A_1 \) under contract \( m' \) after some time. If he accumulates (or loses) \( \Delta \) units of human capital, he moves to point \( A_2 \) (\( A_{-2} \)) in a neighboring ladder. When he becomes unemployed, he comes down to \( C_0 \). If he retires, he is replaced by a newly born worker at point \( D_0 \). Lemma 1 and 2 determines the size and location of each ladder. Lemma 3-4 jointly determine gains from human capital accumulation and depreciation and Lemma 5 specifies the speed of climbing the ladder through non-productive promotion.
Given $F$ and the optimal wage schedule by each firm, the following lemma presents steady state $\{(u_i, G_i)\}_{i=1}^n$.

**Lemma 6** In the steady state equilibrium,

\[
\begin{align*}
\dot{G}_1(E_1) &= \lambda_0 F(E_1) u_1 + \eta G_2(E_2) - (\rho + \delta + \lambda(1 - F(E_1)) + \mu) G_1(E_1), \\
\dot{G}_i(E_i) &= \lambda_0 F(E_i) u_i + \mu G_{i-1}(E_{i-1}) + \eta G_{i+1}(E_{i+1}) - (\rho + \delta + \lambda(1 - F(E_i)) + \mu + \eta) G_i(E_i), \\
\dot{G}_n(E_n) &= \lambda_0 F(E_n) u_n + \mu G_{n-1}(E_{n-1}) - (\rho + \delta + \lambda(1 - F(E_n)) + \eta) G_n(E_n),
\end{align*}
\]

where

\[
u_i = \begin{cases} 
\frac{\delta G_1(E_1)+\rho+\eta u_2}{\lambda_0+\rho} & \text{if } i = 1 \\
\frac{\delta G_i(E_i)+\eta u_{i+1}}{\lambda_0+\rho+\eta} & \text{if } i = 2, 3, \ldots, n-1 \\
\frac{\delta G_n(E_n)}{\lambda_0+\rho+\eta} & \text{if } i = n
\end{cases}
\]

**Proposition 1** Necessary and sufficient conditions for a market equilibrium are:

(a) a vector of functions $\{w_{si}, E_{si}, J_{si}, F\}_{i=1}^n$ satisfying the system of differential equations and the equal profit condition subject to

\[
\lim_{t \to \infty} (w_{si}(t), E_{si}(t), J_{si}, F) = (\bar{w}, \bar{E}_{si}, \bar{J}_{si}, 1)
\]

(b) $\{(u_i, G_{si})\}_{i=1}^n$ solving the system of differential equations in lemma 5;
(c) \( \min \{ E_i(0, m) - U_i \} = 0 \), and \( F \) satisfies the equilibrium restriction.

Although I provide a fixed point algorithm to find a market equilibrium, it does not guarantee the existence of an equilibrium. Moreover, it is not clear when the least cost incentive compatibility constraints are binding. Instead of a theoretical proof, I solve the model numerically and check whether or not the implied equilibrium outcome satisfies the sufficient condition. Here, I report that in a broad range of model parameters, I obtained a unique fixed point with all constraints binding. In particular, when \( \mu \) is small so that a relatively large mass is of the \( y_1 \)-type, the constraints are binding.\(^{14}\)

3 Data

I 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. To estimate the model, I need to keep track of entire work histories from their first full time job. NLSY79 is well suited to analyze careers because it 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, it keeps track of five jobs in each survey round.

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 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 I cannot check what they did immediately after graduation. I also discard individuals who enrolled for military service because their experience and work decisions are different from others. Following this selection rule, I 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, I 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

\(^{14}\)In our estimates, we get \( \mu = 0.023 \).
average. We keep track of the work history of each individual from the first transition.\footnote{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.} Note that by construction, all individuals start their career as an employed worker in my sample. To mitigate any risk of potential bias, I also ignore the first unemployment period before the first job in our simulation.

The model implies that there is neither recalled jobs nor returned workers. However, in NLSY79, workers frequently returned to their former jobs after leaving for some periods. If it was planned by both parties in advance such as unpaid vacation or hospitalization, the former job and the recalled job should be considered as one job, as long as 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 as two different jobs, because the fact that he returned affects neither the previous labor contract nor the new labor contract. To distinguish these, as in Pavan (2008), it is nature to think 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, I drop the intermediate work history and connect the two jobs as one continued job. Otherwise, I 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\footnote{NLSY79 does not distinguish ‘job’ from ‘employer-employee match’. Therefore all returning cases are considered as one job.} and 4,880 jobs.

I 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, I 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, I use hours worked per week. If it is not available, I 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 spells. The difference comes from job-to-job transition. If a worker switches to a new job from an old job, job tenure is reset, but employment tenure continues. However, it is not clear how to determine job-to-job transition in the sample. Workers may have short term vacations before switching to new jobs, even though they made the switching decision on the old job. Thus, I 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 rather than an employment-unemployment-employment transition. This selection integrates 1,702 short term non-employment into the subsequent job tenure.

The model assumes that a worker accumulates human capital only on the job. Hence, I should distinguish work experience from market experience. I refer to worker experience as the sum of all employment spells. Market experience is calculated by subtracting the age at entry from the current age of the worker. In the NLSY79 data set, wages are reported at the interview date and the end date if the job was ended. In addition, they began asking the first wage on the job from the 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, first wage data might be biased. To mitigate the potential bias, in my simulation I use the first wage only when the job started after or continued until the 1985 survey. Then, among the first wages reported, we define the reemployment wage as the first wage after non-employment. The reemployment wage is a key variable to estimate the effect of human capital accumulation. In my sample, I have 13,735 wage observations. These include some observations with potential coding errors. Moreover, it is hard to fit all data points (especially data points at both ends) using a simple model. Hence I discard both the top and bottom 2.5% and focus on the remaining 95% of wage observations.

The finalized sample contains 665 individuals, 4,325 employer-employee matches, 4,880 jobs, and 14,298 observations. Details of the construction of the data set are contained in the Appendix.

4 Estimation

4.1 Estimation Procedure

I 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 $\theta_0$. Denote by $g(\theta)$ the vector of the target moments simulated by the parameter vector $\theta$. To estimate $\theta$, I minimize the distance between the set of the sample moments from NLSY79 and

\footnote{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).}

\footnote{For details on ‘indirect inference’, see Gourieroux, Monfort, and Renault (1993) and Gourieroux and Monfort (1997).}
the set of the moments from our simulations. I simulate and calculate the moment vector \( k \) times and take their average. Then, the simulated moments estimator of \( \theta_0 \) is defined as

\[
\hat{\theta} = \arg \min_{\theta} \quad (\bar{g}_k(\theta) - g(\theta_0))^T \hat{w}_n (\bar{g}_k(\theta) - g(\theta_0)),
\]

where \( \hat{w}_n \) is a positive definite matrix that converges in probability to a deterministic positive definite matrix \( W \). For this, I use the inverse of the covariance matrix of the auxiliary statistics. I 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 I do not have any theoretical evidence on the uniqueness of the minimum value, I minimize the objective function using both the Nelder-Mead and simulated annealing algorithms. First, I use the Nelder-Mead method repeatedly. When the distance reaches a local minimum, I reset the size of the simplex and restart from the local minimum. If the program stops at a point sufficiently close to the local minimum, I start the simulated annealing method. Although this requires heavy computation, I can increase the probability that I reach a global minimum by applying the simulated annealing method repeatedly. I repeat this process with four\(^{19} \) different starting points. If I get the same estimates for the structural parameters, I take it to be a global minimizer.

4.2 Estimation Specification

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

\[
u(w) = -\exp(-\gamma w)
\]

I normalize \( y_{nj} = 1.0 \). The most productive worker produces one unit of output. Then, I set \( y_1 = 0.4 \), and \( \Delta = 0.1 \) so that I have 7 types of workers \( (n_j = 7) \).\(^{20} \) This choice is arbitrary, but without output data, it is hard to get inference on these parameters. The number of equilibrium contracts are fixed to 20 levels\(^{21} \) and I set

\(^{19}\)Actually, we would need more than four.

\(^{20}\)The level of human capital is discretized into 7 levels.

\(^{21}\)We will also report the case with 19 and 21 equilibrium contracts later.
s = 0.01. In my sample, the highest wage is almost eight times larger than the lowest wage. My choice makes the highest wage eight or nine times larger than the lowest wage depending on parameter values. I fix the interest rate \( r \) at 0.012.

It is hard to estimate the arrival rate of the retirement shock \( \rho \) based on the NLSY79 data set because of its short history. I assume that the average worker stays in the labor market for 40 years, which fixes \( \rho \) at 1/160. Instead, to match the actual survival probability, I introduce an ‘attrition probability’ in each survey round. I 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%.

### 4.3 Estimation

I have seven structural parameters to be estimated: four Poisson arrival rates, \((\delta, \lambda_u, \lambda, \mu)\), the risk aversion parameter \( \gamma \), the unemployment benefit \( b \) (or \( w_{\text{max}} \)), and cost function parameter \( c_1 \).

First, to capture the dynamic flow of workers, I 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 an average unemployment duration of 0.471 years, job spell of 2.175 years. The average worker keeps a full time job during 88.3% of the first five years.

Second, one of the main tasks in this empirical study is to estimate the effect of human capital accumulation separately from the effect of strategic promotion and job turnover. I take advantage of the reemployment wage which is defined by the first wage after unemployment. I regress log reemployment wage \( \hat{w} \) on work experience, \( \hat{w}_k = \beta_0 + \beta_1 \times \text{work experience}_k + \varepsilon_k \), where \( \varepsilon_k \) is a statistical residual. I adopt \( \beta_1 \) 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 the model, human capital accumulation occurs at rate \( \mu \) and it increases workers’ wages by a certain amount, which is affected by \( c_1 \). To capture the frequency and magnitude of each shock separately, I also take advantage of information on the re-employment wage distribution. From the sample, I calculate the ratio of the 3rd quartile to the 1st quartile of the dis-
tribution, and the 2nd quartile to the 1st quartile. The auxiliary regression indicates a coefficient $\beta_1$ of 0.109 and the two quartile ratios are 1.775 and 1.281, respectively.

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

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

I adopt $\alpha_1$ as one auxiliary moment. The reason that we focus on the wages reported in the first five years is that promotion rates are different depending on the level of human capital. I want to focus on a narrow and identical group to capture the slope more accurately. In the sample, $\alpha = 0.052$.

Finally, to capture overall wage growth (or wage-age profile), I add some additional auxiliary moments. Denote by $w_1$ the first wage reported within the first 6 months after the transition to work. Also denote by $w_5$, $w_{10}$, and $w_{20}$ the average of wages reported first after 5 years, 10 years, and 20 years of market experience, respectively. I take the ratios $w_5/w_1$, $w_{10}/w_1$, and $w_{20}/w_1$, which 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.$^{22}$

<table>
<thead>
<tr>
<th>Table 1: Auxiliary Moments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>average unemployment duration (yr)</td>
</tr>
<tr>
<td>average job duration (yr)</td>
</tr>
<tr>
<td>average unemployment periods in the first 5 years</td>
</tr>
<tr>
<td>$\Delta \log(\tilde{w}) / \Delta \text{work experience}$</td>
</tr>
<tr>
<td>3rd/1st quartile ratio of reemployment wage dist.</td>
</tr>
<tr>
<td>2nd/1st quartile ratio of reemployment wage dist.</td>
</tr>
<tr>
<td>$\Delta \log(w) / \Delta \text{market experience}$</td>
</tr>
<tr>
<td>$w_{20}/w_1$</td>
</tr>
<tr>
<td>$w_{10}/w_1$</td>
</tr>
<tr>
<td>$w_5/w_1$</td>
</tr>
</tbody>
</table>

*Standard errors in the second column are estimated using bootstrap. The asymptotic standard error of the estimated moments are reported in the parenthesis in the third column.

The second column shows auxiliary moments calculated from NLSY79. The number in the parenthesis is the bootstrapping standard error to use in estimating the weight matrix. The third column provides the estimates of the moments from simulation.\footnote{The asymptotic standard errors will be reported soon.}
[Table 2] reports the estimates of the structural parameters.

Table 2: Parameter Estimation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta$ (separation shock)</td>
<td>0.091</td>
</tr>
<tr>
<td>$\lambda_u$ (offer finding rate by unemployed workers)</td>
<td>0.580</td>
</tr>
<tr>
<td>$\lambda$ (offer finding rate by employed workers)</td>
<td>0.446</td>
</tr>
<tr>
<td>$\mu$ (human capital accumulation shock)</td>
<td>0.023</td>
</tr>
<tr>
<td>$b$ (unemployment benefit)</td>
<td>0.413</td>
</tr>
<tr>
<td>$c_1$ (cost parameter)</td>
<td>0.302</td>
</tr>
<tr>
<td>$\gamma$ (risk aversion parameter)</td>
<td>0.450</td>
</tr>
</tbody>
</table>

4.4 Counterfactual Analysis

In this section, I conduct a counterfactual experiment to understand how human capital accumulation contributes to wage growth. The counterfactual experiment is designed to show how much a representative worker would earn if he were not able to accumulate any human capital. To this end, I need to keep all players’ strategies unchanged. As before, firms optimally choose their strategies assuming that workers stochastically accumulate human capital. But it is assumed that the worker just stays in the same state when he is hit by the human capital accumulation shock. I repeat this experiment with 665 workers and construct a artificial data set.

Table 3: Counter Factual Analysis

<table>
<thead>
<tr>
<th></th>
<th>$w_5/w_1$</th>
<th>$w_{10}/w_1$</th>
<th>$w_{20}/w_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>1.430</td>
<td>1.614</td>
<td>1.881</td>
</tr>
<tr>
<td>estimation with human capital accumulation</td>
<td>1.478</td>
<td>1.645</td>
<td>1.913</td>
</tr>
<tr>
<td>without human capital accumulation</td>
<td>1.354</td>
<td>1.416</td>
<td>1.418</td>
</tr>
</tbody>
</table>

[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, 10, and 20 years, respectively. In our estimation, it grows by 47.8%, 64.5% and 91.3%, respectively. Without human capital accumulation, the average wage 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%.  

20
The conclusion that wage still grows by 41.8% without human capital accumulation is somewhat surprising. One may argue that it contradicts to the conclusion of Altonji and Shakotko (1987) and Altonji and Williams (2005). They show that ‘returns to job tenure’ accounts for 11% of individual wage growth at most, while ‘returns to experience’ takes the lion’s share of wage growth. Based on this, they conclude that general human capital accumulation accounts for most of wage growth. The seemingly different result comes from different definitions. They define the job tenure effect as the wage loss the worker would suffer if he were to move to a new job with the same values for the error components. They interpret all partial effects of market experience as general human capital accumulation effects. In the model, wage growth within one employer, through both productive and non-productive promotion, are transferred to the next job through the reservation value. If I strictly apply their definition to my model, ‘returns to job tenure’ is zero because workers lose nothing in job-to-job transition. Instead, the effect of human capital accumulation is overestimated because ‘returns to experience’ also includes wage growth through non productive promotion.

Another interesting point is that wage growth through non-productive promotion and job-to-job transition reveals a concave pattern. Wage grows by 35% without human capital accumulation in the first five years. After that, the growth rate becomes moderate. When firms are allowed to back-load wage payments, the recruiting firm has the incentive to post a contract with low contingent values and to promote the worker to a higher valued contract later. As the worker stays longer, the wage increases, the possible job turnover rate decreases, and the promotion rate also declines. This shows that a faster wage growth in early periods can be explained by the strategic back-loading scheme of the firm as well as a concave learning curve.

5 Conclusion

This paper develops and estimates an equilibrium job search model with unobserved human capital. In the theoretical part, I build up an equilibrium with multiple wage-ladders that a worker can climb or switch from one to another. The worker jumps up to a higher rung through job-to-job transition, while he climbs up gradually through non-productive promotion. If he accumulates human capital, he switches to a higher-valued ladder through productive promotion. In the empirical study, I estimate the model using indirect inference. I capture the effect of human capital accumulation using the reemployment wage after unemployment. After estimating the model, I perform a counterfactual experiment which reports that if a typical worker were not able to accumulate human capital, his wage would grow by 41.8%.

The next aim is to add ex ante heterogeneity on the worker side to the framework.
developed in this paper. Although I restrict our attention to the sample of white male high school graduates, they are hardly expected to be homogeneous. The model proposes that the lifetime value received by the worker is not a function of his type, but a function of what he actually produces. Therefore, adding ex ante heterogeneity in the worker’s level of human capital does not require any additional state variables. But more careful attention is required on the sufficient condition.
References


TOPEL, R., AND M. WARD (1992): “Job Mobility and the Careers of Young Men,”
Quarterly Journal of Economics, 107(2), 439–479. 2

14, 38