

Wage Dynamics with Private Learning-by-doing and On-the-job Search

Seung-Gyu Sim*
University of Tokyo

September, 2013

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. In order to encourage production and discourage job turnover, firms reward their workers having better performance and longer job tenure by bonus payments and long service allowances, respectively. The resulting wages grow with human capital accumulation (productive promotion), and job tenure (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 experiment 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%.

Key Words: Wage Dynamics, Human Capital Accumulation, Wage-Tenure Contract
JEL Classification: D82, E24, J31, J41

*This is the revised version of the second chapter in my Ph.D thesis. I deeply thank Shin-ichi Fukuda, Hugo Hopenhayn, Hidehiko Ichimura, John Kennan, Francis Kramarz, Rasmus Lentz, Shouyong Shi, Yongseok Shin, and Chris Taber for their helpful guidance and comments. I am also grateful to all participants at various seminars and conferences including Dallas Fed, Penn State University, SED Annual Meeting 2013, Seoul National University, St. Louis Fed, University of Iowa, University of Maryland, University of Tokyo, University of Washington, University of Wisconsin, and University of Virginia. All remaining errors are mine. Please address correspondence to: Seung-Gyu Sim, Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan, 113-0033. Tel: +81-3-5841-5620. E-mail: sgsim@e.u-tokyo.ac.jp.

1 Introduction

In the U.S. labor market, a typical male¹ 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 entering the labor market. However, given that human capital accumulation is neither sufficient nor necessary condition for wage growth in the frictional labor market, it is necessary to develop a structural model to identify and quantify the relevant sources of wage growth. This paper builds up and estimates an equilibrium job search model by focusing on firms' strategic responses to the worker's private learning-by-doing and on-the-job search behavior.

[Burdett and Mortensen \(1998\)](#) develop a wage-posting model with on-the-job search in which some firms post high wages to attract workers from other firms. In their model, workers climb up the wage ladder only through job-to-job transition. Departing from it, [Burdett and Coles \(2003\)](#) build up an equilibrium in which firms optimally back-load some portion of wage payments to discourage job turnover and to extract more surplus from early leavers. All firms post a common wage-tenure schedule and different firms choose different starting points on it. In their equilibrium, wages grow along with the back-loading schedule as well as job turnover. 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 the deterministic learning-by-doing process 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, it is unclear why each firm sticks to the piece rate sharing rule.

In reality, each worker privately accumulates human capital through learning-by-doing on the job and continually searches for a better paying job. The firm which is not able to monitor learning-by-doing process nor job search outcome, tries to provide right incentives by rewarding the worker having better performance and longer tenure through bonus payments and long-service allowances, respectively. Motivated by this, I incorporate private learning-by-doing into the equilibrium wage-tenure contract framework by [Burdett and Coles \(2003\)](#) and [Stevens \(2004\)](#). The firm offers a back-loading wage schedule which increases with the worker's performance and job tenure. In a market equilibrium, there are multiple wage-ladders and each of them is given to the workers with a particular level of human capital. Each worker gradually climbs his

¹Throughout the paper, I use the masculine pronoun for a worker because I deal with a male sample in the empirical analysis.

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.

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 through human capital accumulation or depreciation (productive promotion).

The private learning-by-doing process gives birth to two types of upward pressure in worker's compensation. First, the firm should commit a higher value for better output 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 latter dominates the former, the latter pushes up the wage payment beyond the incentive compatible wage level. The incentive compatibility constraint should be slack and the information asymmetry does not matter any more. Then, given the convexity of the offer distribution and earning distribution, the optimal productive promotion schedule after human capital accumulation may not be well defined. Hence, another wage determination mechanism such as the piece rate sharing rule is required in solving for it.³ In contrast, if the internal pressure dominates the external pressure, the (least cost) incentive compatibility constraints should be binding, which determine the optimal productive promotion schedule after human capital accumulation. 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 investigate returns to human capital and returns to tenure. 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 employment and wage history. 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 goes through unemployment. Hence, I take advantage of the difference between usual wage growth and re-employment wage⁴ 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 implies that returns to human capital is at best the other 46% wage growth in the first 20 years. Considering the interactions, one can think that it's just an upper bound of returns to

³The anti-discrimination legislation as in [Burdett, Carrillo-Tudela, and Coles \(2009\)](#) is also required. It assumes that a firm should pay the exactly same piece rate to all its workers regardless of their job tenure and performance. Basically, it prohibits firms from back-loading the whole value after human capital accumulation occurs and extracting all surplus before it occurs. This paper relaxes it and assumes that a firm cannot pay different values to two different workers having the same output and the same job tenure.

⁴Re-employment wage means the first wage after unemployment.

human capital. The limited effect of human capital accumulation is 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 usual wage-experience regression.

[Altonji and Shakotko \(1987\)](#) argue that returns to experience (general human capital), identified by ‘within job wage growth transferred to the next job’, take lion’s share in wage growth, while returns to tenure (job specific human capital), identified by ‘within job wage growth non-transferable to the subsequent job’, have limited effects on individual wage growth. This paper is not opposed to the empirical findings in [Altonji and Shakotko \(1987\)](#) and [Altonji and Williams \(2005\)](#) in the sense that ‘within job wage growth transferred to the next job’ is significant in wage growth. However, I argue that we need to be careful in interpreting ‘within job wage growth transferred to the next job’ as returns to human capital, because the non-productive promotion in my model is also transferred to subsequent jobs. In light of this, I redefine the wage growth through productive promotion as returns to human capital and the wage growth through non-productive promotion as returns to tenure.

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 \(2002\)](#). In their paper, it is assumed that the firm and the worker sign on a particular piece rate. If the 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. Wages grow through human capital accumulation, ‘search and stay’ and ‘search and switch’. They estimate the model and report that the wage growth through ‘search and stay’ has a substantial effect, which is consistent with my result. The key difference between two papers lies on the way of modeling the returns to tenure. In their work, they assume an ex post offer matching process,⁵ while I rely on ex ante preemptive back-loading scheme.

The paper proceeds as follows. In section 2, I build up the theoretical model and characterize the equilibrium I am 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.

⁵It is vulnerable to the criticism that the ex post offer matching gives workers the wrong incentives to search for an outside offer. See [Shimer \(2006\)](#).

2 The Model

2.1 Environment

Consider a labor market populated by a unit measure of risk-neutral firms and risk-averse workers. While firms are infinitely-lived and homogeneous, workers are mortal and heterogeneous in terms of the level of human capital, $y_i \in \mathcal{Y} := \{y_1, y_2, \dots, y_n\}$. For expositional convenience, I call the worker having y_i units of human capital ‘ y_i -type worker’. In particular, a newly-born worker enters the labor market with y_1 units of human capital, but he accumulates human capital throughout his career. He stochastically retires (or dies) and the retiree is replaced with another newly-born worker immediately. The model is set in continuous time, and all firms and workers discount the future at rate r .

Workers A worker is either unemployed or employed. A y_i -type unemployed worker collects unemployment benefits b per instant, finds a job offer at rate λ_u , retires at rate ρ , and becomes a y_{i-1} -type by losing human capital at rate η_i . Denote by U_i the equilibrium asset value for the y_i -type unemployed worker. Also, let $F_i(\cdot)$ be the cumulative distribution function of lifetime values offered to y_i -type workers by recruiting firms. The equilibrium support as well as the cumulative distribution function itself will be endogenously determined later. The HJB equation for the y_i -type unemployed worker is given by

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

Since there is no worker having human capital less than y_1 units, it is natural to assume $\eta_1 = 0$ and ignore U_0 . It implies that a y_1 -type worker gets no depreciation shock. In asset value equation (1), the left-hand side is interpreted as the opportunity cost of holding the asset, y_i -type unemployment. The terms on the right-hand side are interpreted as the benefit flow from holding the asset U_i which consists of the dividend flow from the asset, the potential gains from job finding, the potential loss from retirement, the potential loss from human capital depreciation, respectively.

A y_i -type employed worker can produce y_i units of output at every instant. Since the type of the individual worker is time-varying and private information, firms cannot pay different wages by the worker’s type. Instead, it is assumed that the anti-discrimination legislation dictates that a firm should offer the same wage to its workers having the same job tenure and the same performance.⁶ Let $\phi : [0, \infty) \rightarrow \{y_1, y_2, \dots, y_n\}$ be a

⁶I borrow the anti-discrimination legislation concept from [Burdett, Carrillo-Tudela, and Coles \(2009\)](#) with modification. In their work, they assume that a firm should pay the exactly same piece rate to all its workers, which implies that workers producing the same output should be given the same wage. I relax it

mapping from the interval of job tenure to the set of types that the worker potentially pretends to be. Then, operating firms determine wages as a function of job tenure and performance, that is, $w(t, \phi(t))$. In addition, it is assumed that if a y_i -type employed worker mimics a $\phi(t)$ -type and produces $\phi(t)$ -units of output at time t , his flow disutility is given by

$$c_i(\phi(t)) = \begin{cases} \alpha_0 - \alpha_1(y_i - \phi(t)) & \text{if } \phi(t) \leq y_i \\ \infty & \text{otherwise} \end{cases}.$$

It implies that the disutility from working is proportional to hours worked. A y_i -type worker producing y_i units of output incurs the disutility of α_0 . But if a y_i -type worker decides to produce $\phi(t) (< y_i)$ units, he can finish his job earlier and save his disutility through leisure. The private benefit from misreporting is captured by $\alpha_1(y_i - \phi(t))$. Note that in the market equilibrium, there is no efficiency loss due to the information asymmetry in each match.

A y_i -type employed worker finds another job offer at rate $\lambda \in (0, \lambda_u)$ ⁷, privately accumulates human capital at rate μ_i , loses it at rate η_i , separates from his job at rate δ , and retires at rate ρ . The expected lifetime value of the y_i -type employed worker at tenure t choosing production schedule $\phi(\cdot)$ under contract m , $E_i(t; \phi, m)$, is given by

$$E_i(t; \phi, m) = \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_i+\eta_i)(s-t)} \left[u(w(s, \phi_i(s); m)) - c_i(\phi(s)) + z(s; m) \right] ds,$$

where

$$\begin{aligned} z_i(s; m) = & \delta U_i + \lambda \int_{E_i}^{\bar{E}_i} \max\{x, \max_{\phi(\cdot)}\{E_i(s; \phi, m)\}\} dF_i(x) + \mu_i \max_{\phi(\cdot)}\{E_{i+1}(s; \phi, m)\} \\ & + \eta_i \max_{\phi(\cdot)}\{E_{i-1}(s; \phi, m)\}. \end{aligned}$$

As before, assume $\eta_1 = \mu_n = 0$ and feel free to ignore $E_0(s; \phi, m)$ and $E_{n+1}(s; \phi, m)$ throughout the paper. A y_i -type employed worker under contract m can choose and update his own production schedule to maximize his lifetime value. His expected value at time t is characterized by

$$\max_{\phi(\cdot)} \{E_i(t; \phi, m)\}.$$

If a y_i -type employed worker truthfully chooses $\phi(\cdot) = y_i$, I use $E_i(t; m)$ instead of $E_i(t; \phi, m)$. The HJB equation for the y_i -type employed worker with the truthful

and assume that firms can pay different wages to their employees depending on the job tenure as well as performance.

⁷I assume that an unemployed worker contacts a recruiting firm more frequently than an employed worker.

production schedule is given by

$$rE_i(t; m) = u(w(t, y_i; m)) - c_i(y_i) + \dot{E}_i(t; m) + \lambda \int \max\{x - E_i(t; m), 0\} dF_i(x) - \rho E_i(t; m) \\ + \delta_i(U_i - E_i(t; m)) + \mu_i(E_{i+1}(t; m) - E_i(t; m)) + \eta_i(E_{i-1}(t; m) - E_i(t; m)). \quad (2)$$

Firms Each firm maintains one vacancy at every instant. To recruit a worker, the firm posts labor contract m . It specifies the action profile (or ‘terms of trade’) which consists of the output schedule by the worker and the lifetime value delivered by the firm under the truthful revelation assumption. That is, m is characterized by $\{(y_i(\cdot), E_i(\cdot; m))\}_{i=1}^n$.

Definition *Contract m is incentive compatible for y_i -type if*

$$E_i(t; m) \geq \max_{\phi} \{E_i(t; \phi, m)\} \quad \text{at each } t \in [0, \infty). \quad (3)$$

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 . In addition, I define the least cost incentive compatibility separately as follows.

Definition *Contract, m , is least cost incentive compatible for y_i -type if the following statements hold;*

- (i) *Contract m is incentive compatible for y_i -type workers.*
- (ii) *There exists at least one $\phi : [0, \infty] \rightarrow \mathcal{Y} \cap \{y_i\}^c$ such that $E_i(t; m) = E_i(t; \phi, m)$ at any $t \in [0, \infty)$.*

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

Once the menu of contracts is accepted by a worker, the firm together with the worker starts producing immediately. If y_i units are produced by an employee with job tenure t , the operating firm earns revenue y_i and makes wage payment $w(t, y_i; m)$. The match is destroyed when the worker leaves the job either voluntarily or involuntarily.⁸ Denote by $J_i(t; m)$ the expected value of an operating job with a y_i -type worker under

⁸In equilibrium, since all jobs yield positive expected profit to firms, there is no endogenous firing.

contract m . For expositional convenience, let

$$\hat{z}_i(s; m) = \mu_i J_{i+1}(s; m) + \eta_i J_{i-1}(s; m).$$

Given the promised value $\{E_i(0; m)\}_{i=1}^n$, the operating firm with a y_i -type worker chooses the schedule of $\{w(\cdot, \phi_i(\cdot); m)\}_{i=1}^n$ to maximize the expected value

$$\int_t^\infty e^{-\int_t^s [r + \rho + \delta_i + \lambda(1 - F_i(E_i(\tau; m))) + \mu_i + \eta_i] d\tau} [y_i(s) - w(s, y_i(s); m) + \hat{z}_i(s; m)] ds \quad (4)$$

subject to the set of least cost incentive compatibility constraints and the set of promise-keeping constraints. The least cost incentive compatibility constraints mean the condition (i) and (ii) above. The promise-keeping constraints are described in (2), which implies that once the firm commits $\{E_i(\cdot; m)\}_{i=1}^n$ in terms of the truthful revelation value, it should deliver it through wage schedules.

Let u_i and $G_i(x)$ be the proportion of y_i -type unemployed workers and employed workers receiving the value less than x , respectively. Also, denote by \bar{m} and \underline{m} the contract offered by the most generous recruiting firm and the least generous firm, respectively.⁹ Then, for each type,

$$G_i(E_i(0; \underline{m})) = 0, \quad \frac{\partial G_i(\cdot)}{\partial x} \geq 0, \quad \text{and} \quad \sum_{i=1}^n [u_i + G_i(E_i(0; \bar{m}))] = 1.$$

Denote by \mathcal{M} the set of the equilibrium contracts. The equal profit condition implies that

$$\sum_{i=1}^n (\lambda G_i(E_i(0; m)) + \lambda_u u_i) J_i(0; m) \begin{cases} = \pi, & \text{if } m \in \mathcal{M}, \\ < \pi & \text{otherwise.} \end{cases} \quad (5)$$

By aggregating all recruiting firms' strategies, I get the distribution of lifetime values offered to each type, $\{F_i\}_{i=1}^n$. Given $\{F_i\}_{i=1}^n$, both employed and unemployed workers optimally behave. Given $\{F_i\}_{i=1}^n$ and workers' optimal behaviors, operating firms choose both productive and non-productive promotion schedules, which determine the steady state distribution $\{u_i, G_i(\cdot)\}_{i=1}^n$. Then, $\{u_i, G_i(\cdot), J_i, E_i\}_{i=1}^n$ should be consistent with the equal profit condition (5). 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 $\{(y_i, E_i(0; m))\}_{i=1}^n$*

⁹One 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.

if and only if

$$\max_{\phi} \{E_i(0; \phi, m)\} \geq U_i.$$

(ii) Given $\{F_i\}_{i=1}^n$, a y_i -type employed worker optimally chooses the production schedule, and accepts a new contract m' if and only if

$$\max_{\phi} \{E_i(0; \phi, m')\} \geq \max_{\phi} \{E_i(t; \phi, m)\}.$$

(iii) Given $\{F_i\}_{i=1}^n$, an operating firm with contract m optimally chooses the wage schedules to deliver $\{E_i(0; m)\}_{i=1}^n$. The contracts described by $\{y_i, E_i(t; m)\}_{i=1}^n$ are least cost incentive compatible at any $t \in [0, \infty)$.

(iv) Given $\{F_i\}_{i=1}^n$, the optimal behaviors of each economic agents determine $\{u_i, G_i(\cdot)\}_{i=1}^n$.

(v) Given $\{u_i, G_i(\cdot)\}_{i=1}^n$ and a recruiting firm optimally posts contract m given the equal profit condition described in (5).

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

2.2 Equilibrium Characterization

In this subsection, I characterize the equilibrium in which all firms offer least cost incentive compatible contracts. Technically, I assume that all least cost incentive compatible constraints are binding, and then numerically check whether or not firms have incentives to deviate from the least cost incentive compatible contracts. To make the model tractable, I impose the following equilibrium restriction.

Eq'm Restriction Let \mathcal{M} be the set of contracts offered on equilibrium. For all $E_i(t; m) \in (\underline{E}_i, \bar{E}_i)$,

- (i) F_i is continuously differentiable and satisfies $F'_i(E_i)$ is bounded away from zero.
- (ii) $F_1(E_1(t; m)) = F_2(E_2(t; m)) = \dots = F_n(E_n(t; m))$ for any $m \in \mathcal{M}$ and $t \in [0, \infty)$.

It implies that the acceptance probability (and also retention probability) under contract m is same across all types. These are based on pre-imposed equilibrium condition that firms have no incentive to screen out any type. Also, note that condition (ii) makes the least generous contract \underline{m} and the most generous contact \bar{m} well-defined.

Lemma 1 *Contract m is least cost incentive compatible if and only if*

$$u(w(t, y_i; m)) = u(w(t, y_1; m)) + \alpha_1(y_i - y_1), \quad \text{at any } t \in [0, \infty). \quad (6)$$

Lemma 1 characterizes the least cost incentive compatible contracts. In addition, I need to characterize the set of least cost incentive compatible contracts. Since no firm has incentives to screen out any particular type, I obtain that

$$\min_i \{E_i(0; \underline{m}) - U_i\} = 0. \quad (7)$$

It implies that the lifetime value of employment delivered by the least generous firm should be greater than the values of unemployment and one of them should be equal. Also, let \bar{m} be the most generous contract posted by recruiting firms on equilibrium. The equal profit condition dictates that given \underline{m} ,

$$\sum_{i=1}^n \lambda_u u_i J_i(0; \underline{m}) = \sum_{i=1}^n (\lambda G_i(E_i(0; \bar{m})) + \lambda_u u_i) J_i(0; \bar{m}). \quad (8)$$

Since there is no contract dominating contract \bar{m} , the operating firm with contract \bar{m} has no incentive to increase the worker's value over $\{E_i(0; \bar{m})\}_{i=1}^n$. That is, for any $i = 1, 2, \dots, n$,

$$E_i(t; \bar{m}) = E_i(0; \bar{m}), \quad \text{at every } t \in [0, \infty). \quad (9)$$

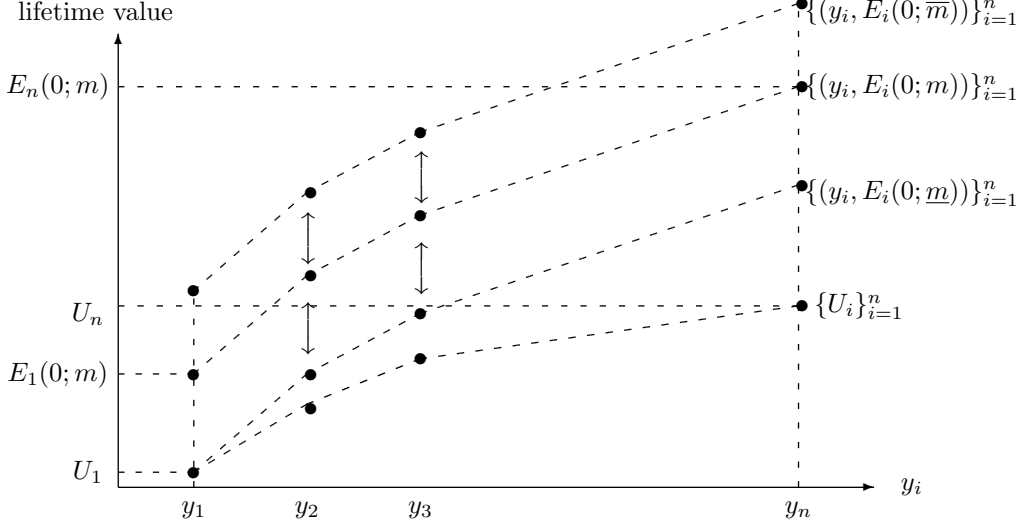
It implies that given i ,

$$E_i(t; \bar{m}) = \frac{u(w(t, y_i; \bar{m})) - c_i(y_i) + \delta U_i + \mu_i E_{i+1}(t; \bar{m}) + \eta_i E_{i-1}(t; \bar{m})}{r + \rho + \delta + \mu_i + \eta_i} = \bar{E}_i, \quad (10)$$

$$J_i(t; \bar{m}) = \frac{y_i - w(t, y_i; \bar{m})}{r + \rho + \delta + \mu_i + \eta_i} = \bar{J}_i, \quad \text{and} \quad (11)$$

$$w(t, y_i; \bar{m}) = w(0, y_i; \bar{m}) = \bar{w}_i \quad \text{for any } t \in [0, \infty). \quad (12)$$

The second statement implies that once $E_i(0; \bar{m})$ by the most generous recruiting firm is given, no firms have incentives to pay more than the value. Therefore, the optimal wage schedule for each type by the most generous firm is constant. In equilibrium, the wage schedules by all firms are bounded.



[Figure 1]

[Figure 1] summarizes the above. If both m and m' are least cost incentive compatible, then $\{(y_i, E_i(0; m))\}_{i=1}^n$ and $\{(y_i, E_i(0; m'))\}_{i=1}^n$ cannot cross each other. Lemma 1 determines the size of information rent, which is represented by the slope of the contract curve in [Figure 1]. The gap between $E_i(0; \underline{m})$ and U_i should be zero at at least one y_i .¹⁰ Given $\{F_i(\cdot), G(i, \cdot)\}_{i=1}^n$ and \underline{m} , the equal profit condition determines \bar{m} .

Now, I look into the strategy by the least generous firm. Given \underline{m} , the firm chooses an optimal wage schedule to deliver the committed value $\{E_i(0; \underline{m})\}_{i=1}^n$. For expositional convenience, denote by w_i the wage schedule to the y_i -type worker producing truthfully under the least cost incentive compatible contract \underline{m} . Let

$$\hat{\psi}(t) = \exp\left[\int_0^t (r + \rho + \delta + \lambda(1 - F_1(E_1(s; m))))ds\right].$$

Lemma 2 *Given $\{F_i\}_{i=1}^n$ and $E_1(0; \underline{m})$, the optimal wage-tenure schedules solve for*

$$\begin{aligned} \dot{J}_i &= -(y_i - w_i + \eta_i J_{i-1} + \mu_i J_{i+1}) + [r + \rho + \delta + \lambda(1 - F_i(E_i)) + \mu_i + \eta_i] J_i, \\ \dot{E}_i &= -u(w_i) + \alpha_0 + (r + \rho + \delta + \lambda(1 - F_i(E_i)) + \mu_i + \eta_i) E_i \\ &\quad - \delta U_i - \lambda \int_{E_i}^{\bar{E}} x dF_i(x) - \mu_i E_{i+1} - \eta_i E_{i-1}, \\ \dot{w}_1 &= \left[\frac{\psi_1 \dot{\psi}_1 u''(w_1)}{[u'(w_1)]^2} - \sum_{i=2}^n \frac{x_{ji} \dot{\psi}_1 u''(w_i)}{[u'(w_i)]^2} \cdot \frac{u'(w_1)}{u'(w_i)} \right]^{-1} \\ &\quad \left[\hat{\psi}(t) \sum_{i=1}^n x_{ji} \lambda F'_i(E_i) J_i + \frac{\psi_1 \dot{\psi}}{u'(w_1)} - \sum_{i=2}^n \frac{x_{ji} \dot{\psi}}{u'(w_i)} + \frac{\dot{\psi}_1 \hat{\psi}_1}{u'(w_1)} - \sum_{i=2}^n \frac{\dot{x}_{ji} \hat{\psi}_1}{u'(w_i)} \right] \end{aligned}$$

¹⁰In most numerical experiments, $E_1(0; \underline{m}) = U_1$ and $E_{-1}(0; \underline{m}) > U_{-1}$.

$$\begin{aligned}
w_i &= u^{-1}(u(w_{i-1}) + \alpha_1 \Delta), \quad \text{and} \\
\dot{x}_i &= x_{i-1}\mu_i + x_{i+1}\eta_i - x_i[r + \rho + \delta + \lambda(1 - F_i(E_i)) + \mu_i + \eta_i]
\end{aligned}$$

subject to the boundary conditions:

$$\lim_{t \rightarrow \infty} \{J_i, E_i, w_i, x_i\}_{i=1}^n = \{\bar{J}_i, \bar{E}_i, \bar{w}_i, 0\}_{i=1}^n.$$

Now consider the general optimal wage-tenure contract $m (\neq \underline{m})$ by other recruiting firms. I borrow the concept of the baseline salary scale from [Burdett and Coles \(2003\)](#). The baseline salary scale is important as it can be extended to prescribe the wage schedules offered by all firms in a steady state. The optimal paths represented by the system of differential equations and their boundary conditions in Lemma 2 are uniquely determined. The initial value $E_1(0; \underline{m})$ determines only a starting point on them. It means that other firms posting $E_1(0; m) > E_1(0; \underline{m})$ also move along the same path but from a different starting point. Different recruiting firms just choose different starting points on the baseline salary scale. That is,

$$\begin{aligned}
E_i(0; m) = E_i(t; \underline{m}) = E_i(t), \quad J_i(0; m) = J_i(t; \underline{m}) = J_i(t), \quad \text{and} \\
w_i(0; m) = w_i(t; \underline{m}) = w_i(t)
\end{aligned}$$

Given the baseline property, define $F : [0, \infty] \rightarrow [0, 1]$ as the distribution of starting points on the baseline offered by firms. Let $G_i(t)$ be the proportion of y_i -type employed workers receiving less than $E_i(t; \underline{m})$. Then,

$$F_i(E_i(0; m)) = F_i(E_i(t; \underline{m})) = F(t) \quad \text{and} \quad G(i, E_i(0; m)) = G(i, E_i(t; \underline{m})) = G_i(t).$$

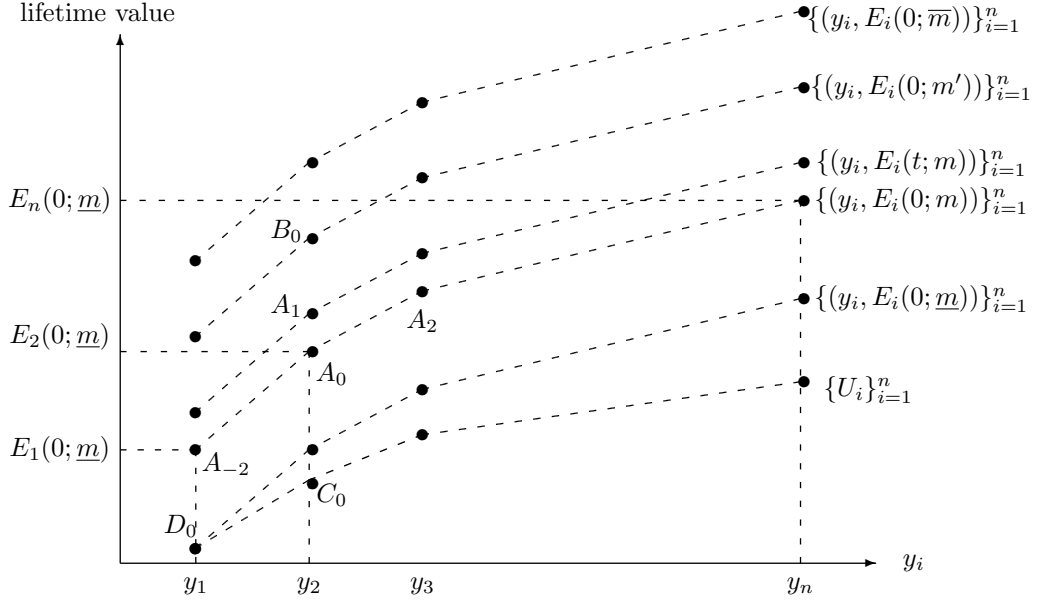
The y_i -type employed workers retire, are laid off, accumulate human capital or lose it at rate $\rho + \delta + \mu_i + \eta_i$ so that the outflow from \bar{G}_i is $(\rho + \delta + \mu_i + \eta_i)\bar{G}_i$, where $\bar{G}_i = G_i(\infty)$. At the same time, y_i -type unemployed workers get jobs at rate λ_u . Also, y_{i-1} -type and y_{i+1} -type employed workers enter into \bar{G}_i at rate μ_i and η_i , respectively. Then, equating the outflow and inflow at steady state, I can solve for \bar{G}_i and also u_i . By the same reasoning, I also obtain for $\dot{G}_i(t)$. Finally, the baseline property presents steady state $\{(u_i, G_i)\}_{i=1}^n$ in what follows.

Lemma 3 *In the steady state equilibrium,*

$$\dot{G}_i(t) = \lambda_u F(t) u_i + \mu_{i-1} G_{i-1}(t) + \eta_{i+1} G_{i+1}(t) - (\rho + \delta + \lambda(1 - F(t)) + \mu_i + \eta_i) G_i(t),$$

where

$$\begin{aligned}
u_1 &= \frac{\delta \bar{G}_1 + \rho + \eta_2 u_2}{\lambda_u + \rho}, \\
u_i &= \frac{\delta \bar{G}_i + \eta_{i+1} u_{i+1}}{\lambda_u + \rho + \eta_i}, \quad i = 2, 3, \dots, n \quad \text{and} \\
\bar{G}_i &= \frac{u_i \lambda_u + \mu_{i-1} \bar{G}_{i-1} + \eta_{i+1} \bar{G}_{i+1}}{\delta + \rho + \mu_i + \eta_i}, \quad i = 1, 2, \dots, n.
\end{aligned}$$



[Figure 2]

[Figure 2] summarizes Lemma 1 through Lemma 3. 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 after some time. If he accumulates (or loses) Δ 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 who starts his career at point D_0 . Lemma 1 and 2 determines the gains from human capital accumulation (depreciation) and the speed of the non-productive promotion. Finally, lemma 3 presents the earning distribution over the type-value space.

Proposition 1 *The sufficient conditions for a market equilibrium are:*

- (i) *Given F , Lemma 1-3 are satisfied.*

(ii) F is stationary and 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 μ is small so that a relatively large mass is of the y_1 -type, the constraints are binding.¹¹

3 Data

I use data from the 1979 National Longitudinal Survey of Youth (NLSY79), which contains weekly work records from 1978 through 2010. The model implies that workers receive different wages based on their (unobserved) productivity and (observed) job tenure. NLSY79 is well suited to analyze careers (and human capital accumulation) because it reports weekly labor force status from the high school period, which enables me 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 way of constructing the sample is presented in Appendix B in details. In this section, I mainly focus on the definition of variables that I have adopted.

Workers The survey starts with individuals who were 14-22 years old in 1979. Among those individuals, I 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 17-20 after 1978, and have never reported more than 12 years of education until the most recent survey. I put the age restriction and year restriction to rule out individuals who have unusual and hidden experience. The individuals who got the high school diploma earlier than their age 17, later than their age 20, or before 1978 are expected to have different and/or hidden experience, which is not captured in the survey. By the similar reasoning, I also discard the individuals who enrolled for military service. Following this selection rule, I collect 776 individuals in the sample.

Full-time Employment and Non-employment Full-time employment is employment in which the worker works more than 30 hours per week in average. The average

¹¹In my estimates, I get $\mu = 0.023$.

hours worked is calculated by the weighted (by the number of weeks) mean of hours worked at the match (with the same employer) throughout his career. The criteria precludes a transition between part-time and full-time within the same employer. But it includes the case in which the worker works less than 30 hours per week for a short period as a trainee or intern and works more than 30 hours for a long period as a regular worker. According to the criteria, the periods reported as out of labor force, no information, unemployment, and employment with the average hours worked less than 30 hours per week are recoded as non-employment, which is the counterpart of unemployment in the model. As for hours worked, I mainly use hours worked per week. When it is not available, I calculate it through ‘hours worked per day’ times five working days per week.

Labor Market Entry Following [Farber and Gibbons \(1996\)](#) and [Yamaguchi \(2009\)](#), if the worker is employed full-time for more than half of three consecutive years for the first time, I assume that the worker makes transition from school to work. I keep track of the work history of each individual from their first transition. Note that by construction, all workers start their career as an employed worker in the sample. To mitigate any risk of potential bias, this paper ignores the first unemployment period before the first job.

Tenures Job tenure is defined by the length of a continuous working period within one employer. Employment tenure means the duration of consecutive job spells without non-employment. 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 starting 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 on-the-job search rather than an employment-unemployment-employment transition. This selection integrates 1,702 short term non-employment into the subsequent job tenure.

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. 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. Finally, my sample contains 4,325 employer-employee matches¹² and 4,880 jobs.

Experiences 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, I 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.¹³ 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.

¹²NLSY79 does not distinguish ‘job’ from ‘employer-employee match’. Therefore all returning cases are considered as one job.

¹³For 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

I use indirect inference as in [Bagger, Fontaine, Postel-Vinay, and Robin \(2006\)](#), because maximum likelihood inference is not numerically feasible. Indirect inference requires¹⁴ 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 by the parameter vector θ . To estimate θ , I minimize the distance between the set of the sample moments from NLSY79 and 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 θ_0 is defined as

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

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 four¹⁵ different starting points. If I get the same estimates for the structural parameters, I take it to be a global minimizer.

¹⁴For details on ‘indirect inference’, see [Gourieroux, Monfort, and Renault \(1993\)](#) and [Gourieroux and Monfort \(1997\)](#).

¹⁵Actually, we would need more than four.

4.2 Estimation Specification

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

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

I normalize $y_{n_j} = 1.0$. 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 ($n_j = 7$).¹⁶ 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 levels¹⁷ and I set $s = 0.01$. 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 ρ 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 ρ 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 γ , the unemployment benefit b (or w_{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

¹⁶The level of human capital is discretized into 7 levels.

¹⁷We will also report the case with 19 and 21 equilibrium contracts later.

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 ε_k is a statistical residual. I 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 the model, human capital accumulation occurs at rate μ 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 distribution, and the 2nd quartile to the 1st quartile. 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, 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 α_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.¹⁸

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

¹⁸The asymptotic standard errors will be reported soon.

Table 1: Auxiliary Moments

	sample moment	simulated moment
average unemployment duration (yr)	0.471 (0.013)	0.467 (0.052)
average job duration (yr)	2.175 (0.044)	2.185 (0.458)
average unemployment periods in the first 5years	0.117 (0.004)	0.123 (0.025)
$\Delta \log(\tilde{w}) / \Delta work\ experience$	0.023 (0.002)	0.025 (0.020)
3rd/1st quartile ratio of reemployment wage dist.	1.775 (0.031)	1.774 (0.467)
2nd/1st quartile ratio of reemployment wage dist.	1.281 (0.019)	1.282 (0.091)
$\Delta \log(w) / \Delta market\ experience$	0.052 (0.004)	0.054 (0.034)
w_{20}/w_1	1.881 (0.042)	1.913 (0.361)
w_{10}/w_1	1.616 (0.033)	1.645 (1.001)
w_5/w_1	1.430 (0.026)	1.478 (0.070)

*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.

Table 2: Parameter Estimation

Parameter	Estimates
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
b (unemployment benefit)	0.413
c_1 (cost parameter)	0.302
γ (risk aversion parameter)	0.450

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] 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

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	1.354	1.416	1.418

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%.

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

- ALTONJI, J., AND R. SHAKOTKO (1987): “Do Wages Rise with Job Seniority?,” *Review of Economic Studies*, 54(3), 437–459. 4, 21
- ALTONJI, J., AND N. WILLIAMS (2005): “Do Wages Rise with Job Seniority? A Reassessment,” *Industrial and Labor Relations Review*, 58(3), 370–397. 4, 21
- 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. 4, 17
- BURDETT, K., C. CARRILLO-TUDELA, AND M. G. COLES (2009): “Human Capital Accumulation and Labor Market Equilibrium,” Manuscript, June 2009. 2, 3, 5
- BURDETT, K., AND M. COLES (2003): “Equilibrium Wage-Tenure Contracts,” *Econometrica*, 71(5), 1377–1404. 2, 12
- BURDETT, K., AND D. T. MORTENSEN (1998): “Wage Differentials, Equilibrium Size, and Unemployment,” *International Economic Review*, 39(2), 257–273. 2
- FARBER, H., AND R. GIBBONS (1996): “Learning and Wage Dynamics,” *Quarterly Journal of Economics*, 111(4), 1007–1047. 15, 30
- GOURIEROUX, C., AND A. MONFORT (1997): *Simulation-Based Econometric Methods*. Oxford University Press, New York, NY. 17
- GOURIEROUX, C., A. MONFORT, AND E. RENAULT (1993): “Indirect Inference,” *Journal of Applied Econometrics*, 8, S85–S118. 17
- PAVAN, R. (2008): “A flexible Model of Individual Wage Dynamics and Job Mobility Outcomes,” August 2008, University of Rochester. 16
- POSTEL-VINAY, F., AND J.-M. ROBIN (2002): “The Distribution of Earnings in an Equilibrium Search Model with State-Dependent Offers and Counter-Offer,” *International Economic Review*. 4
- SHIMER, R. (2006): “On-the-job Search and Strategic Bargaining,” *European Economics Review*, 50(4), 811–830. 4
- 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. WARD (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).
15, 30

Appendices

A Mathematical Appendix

Proof of Lemma 1 i) First, suppose that contract m is least cost incentive compatible. We want to show that condition (6) should be true for all types. Since contract m is least cost incentive compatible, it should be least cost incentive compatible for each type y_i . Consider the case of $i = 2$ first. There should be $\hat{y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)$ with $\hat{y}_2 : [0, \infty] \rightarrow \mathcal{Y} \cap \{y_2\}^c$ and $\hat{y}_i : [0, \infty] \rightarrow \mathcal{Y}$ for each $i \neq 2$, such that $E_2(\cdot; \hat{y}, m) = E_2(\cdot; m)$. At every $t \in [0, \infty)$,

$$\begin{aligned}
& E_2(t; m) - E_2(t; \hat{y}_1, m) \\
&= \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_2+\eta_2)(s-t)} [u(w(s, y_2; m)) - c_2(y_2) + z_2(s; m)] ds \\
&\quad - \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_2+\eta_2)(s-t)} [u(w(s, y_1; m)) - c_2(y_1) + z_2(s; m)] ds \quad (\text{A1}) \\
&= \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_2+\eta_2)(s-t)} [u(w(s, y_2; m)) - c_2(y_2) - u(w(s, y_1; m)) + c_2(y_1)] ds \\
&= \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_2+\eta_2)(s-t)} [u(w(s, y_2; m)) - u(w(s, y_1; m)) - \alpha_1(y_2 - y_1)] ds \\
&= 0,
\end{aligned}$$

which implies that

$$u(w(\cdot, y_2; m)) = u(w(\cdot, y_1; m)) + \alpha_1(y_2 - y_1). \quad (\text{A2})$$

Now, we want to show that once condition (6) is true for $i \leq k$, it is also true for $i = k + 1$. Assume that condition (6) is true for all $i = 1, 2, \dots, k$. That is,

$$u(w(\cdot, y_i; m)) = u(w(\cdot, y_1; m)) + \alpha_1(y_i - y_1) \quad \text{for any } i = 1, 2, \dots, k. \quad (\text{A3})$$

Let $i = k + 1$. Since contract m is least cost incentive compatible for the y_i -type, there exists at least one $\hat{y}' : [0, \infty] \rightarrow \mathcal{Y}^n$ with $\hat{y}'_{k+1} : [0, \infty] \rightarrow \mathcal{Y} \cap \{y_{k+1}\}^c$ and $\hat{y}'_i : [0, \infty] \rightarrow \mathcal{Y}$ for each $i \neq k + 1$, such that $E_i(\cdot; m) = E_i(\cdot; \hat{y}', m)$. Define set $T_j \subset [0, \infty)$ such that if $\hat{y}'_{k+1}(t) = y_j$, then $t \in T_j$, where $j = 1, 2, \dots, k$. Then, for every $t \in [0, \infty)$,

$$\begin{aligned}
& E_i(t; m) - E_i(t; \hat{y}', m) \\
&= \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - c_i(y_i) + z_i(s; m)] ds \\
&\quad - \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, \hat{y}'_i(s); m)) - c_i(\hat{y}'_i(s)) + z_i(s; m)] ds
\end{aligned}$$

$$\begin{aligned}
&= \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - c_i(y_i) - u(w(s, \hat{y}'_i(s); m)) + c_i(\hat{y}'_i(s))] ds \\
&= \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - u(w(s, \hat{y}'_i(s); m)) - \alpha_1(y_i - \hat{y}'_i(s))] ds \\
&= \sum_{j=1}^k \int_{T_j \cap [t, \infty]} e^{-(r+\rho+\delta+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - u(w(s, y_j; m)) - \alpha_1(y_i - y_j)] ds \\
&= \sum_{j=1}^k \int_{T_j \cap [t, \infty]} e^{-(r+\rho+\delta+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - u(w(s, y_1; m)) - \alpha_1(y_i - y_1)] ds \\
&= \int_t^\infty e^{-(r+\rho+\delta+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - u(w(s, y_1; m)) - \alpha_1(y_i - y_1)] ds \\
&= 0.
\end{aligned}$$

Thus, we obtain that when $i = k + 1$,

$$u(w(t, y_i; m)) = u(w(t, y_1; m)) + \alpha_1(y_i - y_1) \text{ at any } t \in [0, \infty). \quad (\text{A4})$$

The mathematical induction yields condition (6).

ii) Conversely, I want to show that if condition (6) holds under contract m , it should be least cost incentive compatible. It is sufficient to show that the y_i -type worker has no incentive to deviate. Consider an arbitrary downward deviation with

$$\phi''(t) = \begin{cases} y_1 & \text{if } t \in T_{i1} \\ \vdots & \quad \quad \quad \vdots \\ y_{i-1} & \text{if } t \in T_{i,i-1} \end{cases}. \quad (\text{A5})$$

and $\bigcup_{j=1}^{i-1} T_{ij} = [0, \infty)$. For any $t \in [0, \infty)$ and $i \in \{1, 2, \dots, n\}$,

$$\begin{aligned}
&E_i(t; m) - E_i(t; \phi'', m) \\
&= \int_t^\infty e^{-(r+\rho+\delta_i+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - c_i(y_i) + z_i(s; m)] ds \\
&\quad - \int_t^\infty e^{-(r+\rho+\delta_i+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, \phi''_i(s); m)) - c_i(\phi''_i(s)) + z_i(s; m)] ds \\
&= \int_t^\infty e^{-(r+\rho+\delta_i+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - c_i(y_i) - u(w(s, \phi''_i(s); m)) + c_i(\phi''_i(s))] ds \\
&= \sum_{j=1}^i \int_{T_{ij}} e^{-(r+\rho+\delta_i+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - c_i(y_i) - u(w(s, y_j; m)) + c_i(y_j)] ds \\
&= \sum_{j=1}^{i-1} \int_{T_{ij}} e^{-(r+\rho+\delta_i+\lambda+\mu_i+\eta_i)(s-t)} [u(w(s, y_i; m)) - u(w(s, y_1; m)) - \alpha_1(y_i - y_1)] ds \\
&= 0.
\end{aligned} \quad (\text{A6})$$

There is no profitable deviation and there exists $\phi'' : [0, \infty] \rightarrow \mathcal{Y} \cap \{y_i\}^c$ such that $E_i(\cdot; m) = E_i(\cdot; \phi'', m)$. Since this is true for all $i = 1, 2, \dots, n$, contract m is least cost incentive compatible. Q.E.D.

Proof of Lemma 2 Rewrite the operating firm's problem:

$$\begin{aligned} \max_{w_1(\cdot)} \int_0^\infty \psi_1 [y_1 - w_1 + \mu_1 J_2] ds \\ \text{s.t. } \dot{\psi}_1 &= -[r + \rho + \delta + \lambda(1 - F_1(E_1)) + \mu_1] \psi_1 \\ \dot{E}_i &= -u(w_i) + \alpha_0 + (r + \rho + \delta + \lambda(1 - F_i(E_i)) + \mu_i + \eta_i) E_i - z_i^*(E_i, m) \\ \dot{J}_i &= -y_i + w_i + (r + \rho + \delta + \lambda(1 - F_i(E_i)) + \mu_i + \eta_i) J_i - \tilde{\varphi}_i(m) \\ u(w_i) &= u(w_{i-1}) + \Delta = u(w_1) + (i-1)\Delta \end{aligned}$$

The Hamiltonian of the problem is

$$\begin{aligned} \mathcal{H} &= \psi_1 [y_1 - w_1 + \mu_1 J_2] - x_\psi [r + \rho + \delta + \lambda(1 - F_1(E_1)) + \mu_1] \psi_1 \\ &+ \sum_{i=1}^n x_{ei} [-u(w_1) + \alpha_0 + \alpha_1(y_i - y_1) + (r + \rho + \delta + \lambda(1 - F_i(E_i)) + \mu_i + \eta_i) E_i - \varphi_i(E_i; m)] \\ &+ \sum_{i=2}^n x_{ji} [-y_i + u^{-1}(u(w_1) + \alpha_1(y_i - y_1)) + (r + \rho + \delta + \lambda(1 - F_i(E_i)) + \mu_i + \eta_i) J_i - \tilde{\varphi}_i(m)] \end{aligned}$$

Applying the maximum principle yields the following first order condition and differential equations.

$$0 = -\psi_1 - \sum_{i=1}^n x_{ei} u'(w_1) + \sum_{i=2}^n \frac{x_{ji} u'(w_1)}{u'(u^{-1}(u(w_1) + (i-1)\Delta))} \quad (\text{A7})$$

$$\dot{x}_\psi = -[y_1 - w_1 + \mu_1 J_2] + x_\psi [r + \rho + \delta + \lambda(1 - F_1(E_1)) + \mu_1] \quad (\text{A8})$$

$$\dot{x}_{j i'} = x_{j i' - 1} \mu_{i' - 1} + x_{j i' + 1} \eta_{i' + 1} - x_{j i'} [r + \rho + \delta + \lambda(1 - F_{i'}(E_{i'})) + \mu_{i'} + \eta_{i'}] \quad (\text{A9})$$

$$\dot{x}_{ei} = x_{ei-1} \mu_{i-1} + x_{ei+1} \eta_{i+1} - x_{ei} [r + \rho + \delta + \lambda(1 - F_i(E_i)) + \mu_i + \eta_i] + x_{ji} \lambda F_i'(E_i) J_i \quad (\text{A10})$$

where $x_{j1}(t) = -\psi(t)$, $i' = 2, 3, \dots, n$ and $i = 1, 2, \dots, n$. From (A8), I obtain

$$\begin{aligned} \dot{x}_\psi \psi_1(t) + x_\psi \dot{\psi}_1(t) &= -[y_1 - w_1 + \mu_1 J_2] \psi_1(t) \\ \iff x_\psi &= \int_t^\infty [y_1 - w_1 + \mu_1 J_2] \frac{\psi_1(0, \tau)}{\psi_1(0, t)} d\tau + A_\psi \psi_1^{-1}(t) \\ &= \int_t^\infty \psi_1(t, \tau) [y_1 - w_1 + \mu_1 J_2] d\tau = J_1. \end{aligned}$$

Plugging (A11) into (A8) yields

$$\dot{J}_1 = -[y_1 - w_1 + \mu_1 J_2] + J_1[r + \rho + \delta + \lambda(1 - F_1(E_1)) + \mu_1]. \quad (\text{A11})$$

Summing up all equations in (A10) and reordering yields

$$\sum_{i=1}^n \dot{x}_{ei} = \sum_{i=1}^n x_{ji} \lambda F'_i(E_i) J_i - (r + \rho + \delta + \lambda(1 - F_1(E_1))) \sum_{i=1}^n x_{ei}.$$

Let

$$\hat{\psi}(t) = \exp\left[\int_0^t (r + \rho + \delta + \lambda(1 - F_1(E_1(s; m)))) ds\right]. \quad (\text{A12})$$

Multiplying by integrating factor on both sides of equation (A12) yields

$$\hat{\psi}(t) \sum_{i=1}^n \dot{x}_{ei} + (r + \rho + \delta + \lambda(1 - F_1(E_1))) \hat{\psi}(t) \sum_{i=1}^n x_{ei} = \hat{\psi}(t) \sum_{i=1}^n x_{ji} \lambda F'_i(E_i) J_i. \quad (\text{A13})$$

Then, multiplying by on both sides of equation (A7), dividing by $u'(w_1)$, taking derivative with respect to t , and combining with equation (A13) yields

$$\begin{aligned} \hat{\psi}(t) \sum_{i=1}^n x_{ji} \lambda F'_i(E_i) J_i &= \left[-\frac{\psi_1}{u'(w_1)} + \sum_{i=2}^n \frac{x_{ji}}{u'(u^{-1}(u(w_1) + (i-1)\Delta))} \right] \dot{\psi} - \frac{\dot{\psi}_1}{u'(w_1)} \\ &+ \sum_{i=2}^n \frac{\dot{x}_{ji}}{u'(w_i)} + \left[\frac{\psi_1 \hat{\psi}_1 u''(w_1)}{[u'(w_1)]^2} - \sum_{i=2}^n \frac{x_{ji} \hat{\psi}_1 u''(w_i)}{[u'(w_i)]^2} \cdot \frac{u'(w_1)}{u'(w_i)} \right] \dot{w}_1 \end{aligned}$$

Rewriting this, I obtain

$$\begin{aligned} \dot{w}_1 &= \left[\frac{\psi_1 \hat{\psi}_1 u''(w_1)}{[u'(w_1)]^2} - \sum_{i=2}^n \frac{x_{ji} \hat{\psi}_1 u''(w_i)}{[u'(w_i)]^2} \cdot \frac{u'(w_1)}{u'(w_i)} \right]^{-1} \\ &\left[\hat{\psi}(t) \sum_{i=1}^n x_{ji} \lambda F'_i(E_i) J_i + \frac{\psi_1 \dot{\psi}}{u'(w_1)} - \sum_{i=2}^n \frac{x_{ji} \dot{\psi}}{u'(u^{-1}(u(w_1) + (i-1)\Delta))} + \frac{\dot{\psi}_1 \hat{\psi}_1}{u'(w_1)} - \sum_{i=2}^n \frac{\dot{x}_{ji} \hat{\psi}_1}{u'(w_i)} \right] \end{aligned}$$

It derives Lemma 2.

Q.E.D

Proof of Lemma 3 Consider the outflow from and inflow into y_i -type unemployment for any arbitrarily small time interval $dt > 0$. By equating them, I obtain that when $i = 1$,

$$\delta dt \bar{G}_1 + \eta_2 dt u_2 + \rho dt (1 - u_1) = (\lambda + \mu_1) dt u_1 \quad (\text{A14})$$

and when $i > 1$,

$$\delta dt\bar{G}_i + \mu_{i-1}dtu_{i-1} + \eta_{i+1}dtu_{i+1} = (\rho + \lambda + \mu_i + \eta_i)dtu_i. \quad (\text{A15})$$

Also, consider the proportion of y_i -type employed workers. Equating the inflow and outflow yields

$$\lambda dtu_i + \mu_{i-1}dt\bar{G}_{i-1} + \eta_{i+1}dt\bar{G}_{i+1} = (\rho + \delta + \mu_i + \eta_i)dt\bar{G}_i, \quad i = 1, 2, \dots, n \quad (\text{A16})$$

Sending $dt \rightarrow 0$ and combining (A14), (A15) and (A16) all together yields the initial values in Lemma 3. By the similar reasoning, the differential equations in Lemma 3 are obtains. *Q.E.D*

B Construction of the Sample

Appendix B reports the steps I used to construct my own sample.

1. The analysis is restricted to white, male, high school graduates, since they are the largest demographic group in the NLSY79 dataset. In the first survey round, 8,736 respondents out of 12,686 are marked as ‘white’,¹⁹ and 6,403 as ‘male’. Combining those responses together, I get 4,393 white males in the initial sample. Among them, 1,990 individuals have completed their 12th grade or received the equivalent degree (GED) without reporting further education until the most recent survey in 2010.
2. To exclude the individuals with hidden or unusual experience, I drop 994 individuals who graduated before January 1st 1978, or graduated before their age 17 (the 204th month) or after their age 20 (the 240th month). Also, 220 individuals are dropped who enrolled in military service at least once until the last survey. This selection rule leaves 776 workers and 34,010 work-records in the sample.²⁰
3. I define as full-time employment the match with weekly hours worked more than 30 in average. The average hours worked is calculated by the weighted (by the number of weeks) mean of hours worked at all the matches with the same employer throughout his career. When weekly hours worked are less than 10, I take the maximum number between the weekly hours worked and the daily hours worked times 5 working days. When only daily hours worked are reported, I use daily hours worked times 5 working days. Before calculating the average, I top-code the hours worker per week so that it cannot exceed 96.²¹
4. Following [Farber and Gibbons \(1996\)](#) and [Yamaguchi \(2009\)](#), if a worker works full-time for more than 78 weeks in three consecutive years for the first time, it is assumed that the worker has made the transition from school to work. The work records before this transition are dropped. This leaves 752 workers, 5,955 full-time employers, 573 part-time employers, and 27,606 work records.

¹⁹To get the sample of ‘white’, I use question (R01727.00) rather than (R02147.00).

²⁰In the non-military sample, 20 individuals reported more than one graduation date. Assuming that these are coding errors, I select the graduation date closest to the age of 18 (the 222nd month).

²¹Indeed, 98 cases are top-coded.

5. I recode as non-employment all non-full-time employment such as ‘out of labor force’, ‘no information’, ‘unemployment’, and employment with the average weekly hours worked less than 30. Then, I merge all consecutive non-employment spells together. It gives me 752 workers, 5,955 full-time employers, 5,918 non-employment spells, and 24,808 work records in my sample.²²
6. After long term non-employment, I assume that the worker is no longer attached to the regular labor market. I drop subsequent observations after 3 years of non-employment spells. This leaves 752 workers, 5,801 (full-time) employers, 5,413 non-employment spells, and 24,019 work records.
7. If a worker quitted an old job and started a new job within 3 weeks, I consider it a job-to-job transition. In this case, I include the intermediate period in the next job spell. If a worker returned to an old job within 13 weeks, I assume that it was a planned return, and recode it as a single continuous job. I finally have 665 workers, 4,796 jobs, and 14,298 observations in the sample.

²²Although the NLSY79 does not distinguish employers from job, I separately define each. In particular, when a worker makes a return to an old employer, it can be considered as a planned return or a random re-match. If it is not a planned return, it should be considered as a different job by the same employer.