Multidimensional Learning, Job Mobility, and Earnings Dynamics*

Joseph Han†

Korea Development Institute

November 7, 2017

Preliminary and Incomplete

Abstract

Job change is a primary way of wage growth for young workers—by improving the quality of their job match. At the same time, low-skilled young workers tend to change jobs more frequently than other workers without any clear sign of career progression. This paper introduces the possibility of multidimensional learning about worker ability and job match quality into a model of work decisions to explain two different types of job mobility—job shopping and job floundering—within an integrated framework. In this setup, worker ability can affect job change and individual employment through an information channel. This mechanism produces a unique prediction, which is testable if there exists a measure of ability which carries over some information unused by workers and employers. I estimate the life-cycle structural model, which also allows flexible general and job-specific skill accumulation, by indirect inference, using a sample from the NLSY79 data. From simulation results on life-cycle earnings dynamics, I find that job floundering is mostly explained by the process of resolving the uncertainty about ability; also, the importance of job shopping in earnings growth is even more highlighted in this extended framework.

Keywords: Uncertainty, ability, job match quality, job change, unemployment, earnings dynamics, indirect inference.

JEL Classification Numbers: J24, J31, J62.

*I am deeply grateful to Chris Taber, John Kennan, and Rasmus Lentz for their guidance and support. All remaining errors are mine.

†Associate Fellow, Korea Development Institute (KDI). e-mail: *han@kdi.re.kr*.

1 Introduction

Uncertainties characterize young workers' career. Most young workers begin their career without much knowledge on their job. It is also rare that the young workers who have just completed their education fully understand their own strength and weakness in the world of work. Nevertheless, they need to make their work decisions, which will affect their lifetime earnings and utility.

This paper tries to understand how uncertainties shape work decisions of young workers. In particular, the main focus here is the joint effect of uncertainties which can be different from their separate effects. This paper extends previous learnings models in the labor market by integrating two different types of uncertainty: worker ability[1], which is only partially known at labor market entry, both to workers and to employers; and job match quality[2], which is not known at job entry, either. In addition to this information channel, the model includes several important features related to work decisions: employment-related shocks such as exogenous job destruction and recall offer arrival[3], job search costs, general and job-specific skills which may accumulate differently by ability, and consumption/saving choice by risk-averse workers.

Although the main contribution of this paper is empirical, some theoretical predictions are new and noteworthy. First, the generalized learning mechanism predicts that job mobility and movement into and out of employment are heterogeneous by true ability—especially, less able workers are predicted to change jobs more frequently and work less than other

[1]Employer learning (e.g., Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lange, 2007) and learning about ability (e.g., James, 2011; Papageorgiou, 2014) literature suggest worker ability may not be fully known at labor market entry. This paper can be read in connection with employer learning when asymmetric information does not meaningfully affect workers' decisions. It is also interesting to see that this multi-dimensional learning under information symmetry has a very similar prediction on job mobility with asymmetric information models (e.g., Greenwald, 1986; Schönberg, 2007; Pinkston, 2009)

[2]In Jovanovic (1979, 1984), where workers gradually learn about the quality of a job match, job hazards are predicted initially increasing then decreasing; also, average wages are predicted to increase in job tenure. These predictions are widely supported by empirical evidence (e.g., Farber, 1999).

[3]Fujita and Moscarini (2013) show recall to a previous employer is important for understanding unemployment.

workers. This is basically a story of *misperceived* productivity[4] where workers cannot directly distinguish between ability and job match quality in their signals: less able workers get worse signals than others; they misinterpret the signals as evidence of a bad job match because their belief on their ability is closer to the median in the population than their true ability is; as a result, they are more likely to search and move. Moreover, less able workers work more than they would do under full information on ability although they still work less than other workers if leisure utility is positive. Second, these selective patterns by unknown ability are predicted to disappear over time. In particular, the differences in job mobility between less able and more able workers are expected to disappear over time. With more observations over the life cycle, workers' belief on their ability eventually reflects their true ability unless the speed of learning is too slow. The initially strong but disappearing negative selection into job mobility, conditional on endogenous job separation, provides a potentially testable implication of this learning story. Furthermore, the positive selection into employment by ability becomes stronger while the initially strong and negative selection by unknown ability disappears over time.

These predictions are indeed from the information side of the story.[5] Also, the predictions, especially the over-time variation in the difference in job hazard between less and more able workers, separately identify the information story from other possible mechanisms such as job-specific skills which are accumulated differently by ability.[6] These predictions are,

[4]Jovanovic (1979, 1984) shows how *perceived* productivity affects career decisions over the life cycle, using a model of learning about job match quality.

[5]Moscarini (2005) criticizes that Jovanovic (1979, 1984)'s predictions are from the selection side rather than the information side of the learning model. Moscarini (2005) shows, in his general equilibrium setup, that information friction uniquely predicts a Pareto upper tail of the empirical wage distribution.

[6]This is comparable to Nagypál (2007)'s. Learning-by-doing on the job increases the opportunity cost of moving as job tenure increases; learning about job match quality reduces the option value of a match as job tenure increases. The two mechanisms have similar predictions on average job mobility and wage growth. Nonetheless, the size of the match-specific surplus changes in different directions over time (on the job), so this can provide a source of identification. In her general equilibrium setting, she uses a negative productivity shock to change the value of a match and observes which matches are more fragile—old vs. new matches. Instead of the negative productivity shock in a general equilibrium setup, I use preference shocks in a partial equilibrium setup in order to trigger different responses in match continuation. In addition, I add one more dimension in the match surplus, other than job tenure, which is worker ability. Then I compare high and low ability groups over time.

however, only potentially useful for identification because it requires a measure of ability which carries over some information unknown to the decision-makers.

This paper takes an additional assumption on the nature of the Armed Forces Qualification Test (AFQT) score in the National Longitudinal Survey of Youth 1979 (NLSY79) data. I assume that the AFQT score carries over some information unused by the workers and employers. This assumption is similar in spirit to Altonji and Pierret (2001) in that the econometrician has more information on worker ability than the agents.[7] I modify the Altonji and Pierret (2001)'s assumption in two ways. First, I assume that the agents (both the workers and employers) have another source of information that is not available to the econometrician such as the SAT score or high school GPA. Second, I assume that the AFQT score was not properly understood (or rationally ignored) by the NLSY respondents except for those with some (future) military career. This is essentially about the nature of the data not related to the model, but this is important for the identification of the model.

With the additional data assumption mentioned above, I estimate the structural model of life-cycle labor supply and earnings, using a sample of only white men high school graduates without any further education or military experience from the NLSY79 data. For the estimation method, I use indirect inference, which is characterized by the use of auxiliary model. The auxiliary model in this paper comprises seven regression equations which explain conditional wages for job movers and stayers and job/employment transitions over the life cycle. This is in many ways similar to Altonji et al. (2013)'s auxiliary model although their structural model is totally different from mine. Nonetheless, I follow Sauer and Taber (2013) instead of Keane and Smith (2003) for the method of smoothing in indirect inference, as Keane and Smith (2003) was no longer applicable in this dynamic setting because of heavy dependence on past variables.

The estimation results are in general consistent with previous results; however, I have two new findings on earnings dynamics. First, the simulation results show that rapid average

[7] Lange (2007) has more discussions on this assumption.

earnings growth over the first 10 years (66 percent) can be attributed to general and job-specific skill accumulation and improved job match quality, at 33 percent, 10 percent, and 24 percent, respectively. While these findings are comparable to the results from other recent models of earnings dynamics (e.g., Altonji et al., 2013), the contribution of job shopping to average earnings growth through improved job match quality is much higher in this paper where the job shopping process is allowed to be heterogeneous across workers. Recall to the last previous employer also plays a role for preserving good matches and skills specific to the matches in late career.

Second, individual heterogeneity in earnings growth is mostly explained by the process of resolving uncertainties. Learning about ability and subsequent wage changes are obvious reasons. Individual earnings converge to the level associated with true ability with more information on ability. Heterogeneous job mobility by unknown ability is another important reason. More able workers are predicted to stay longer during their early career but move relatively more during their late career. Their average match quality increases slowly in their early career, and fast afterwards, compared to median-ability workers. Other channels of earnings growth are not very different by ability. Both general and job-specific skill accumulation processes are fairly homogeneous across workers.

The rest of this paper is structured as follows: Section 2 presents the model. Section 3 explains the data with several observed empirical patterns. Section 4 illustrates the identification and the estimation of the model. Section 5 performs several counterfactual experiments, and Section 6 concludes.

2 The Model

2.1 Overview

In this section, I present a life-cycle model of individual earnings and choices. This model is rich enough to capture many other important aspects of earnings dynamics and individual decisions, including general and job-specific skill accumulation (or "learning by doing"), complementarity between ability and skill accumulation (or "learning ability"), information frictions on ability and job match quality (or "multidimensional learning"), and frictional unemployment and recall .

The main novelty of this model is the multidimensional learning about uncertain ability and job match quality. While the implications of uncertainty about either ability or job match quality are extensively studied in the literature, the implications of the joint uncertainty is not studied in the literature. One notable difference from previous models of earnings dynamics is the multidimensional learning about ability and job match quality, which are permanent (or non-renewable) and transitory (or renewable) components in worker productivity, respectively. While preserving all predictions from the job matching theory based on learning (e.g., Jovanovic, 1984) that is well established by empirical evidence (e.g., Farber, 1999), this model extends the previous literature by introducing a source of individual heterogeneity—unknown ability—into the job matching process based on learning, partly motivated by the literature on learning about ability (e.g., Altonji and Pierret, 2001). By combining the two different types of learning in the labor market, it becomes possible to explain the negative selection into job mobility by ability that disappears over time in the labor market, which is observed in the data.

2.2 Time Schedule

Each individual enters into the labor market after their high school graduation at age 18.[8] They make various decisions about work and job change over T years. After the T years in the labor market, the workers retire from their work and consume what they have

[8]This model abstracts from educational choices—the sample used for estimation contains only high school graduates without any post-secondary education. The primary reason for this simplification is to focus on low-skill workers and their job choice in the labor market, as is often the case in other structural works (e.g., Kennan and Walker, 2011). Readers need to be cautious about the interpretation of the estimated parameters as they may not be generalizable to the entire population including all educational groups.

saved until they die.[9]

Individuals are heterogeneous in ability at age 18, but as long as it is unknown, this ability does not affect any decision—that is, there is no *perceived* difference across the individuals. At high school graduation (or at labor market entry), each individual receives a public signal on their ability which is called a "test score." This creates initial perceived heterogeneity among the high school graduates who are otherwise the same. Uncertainty about ability, however, still remains because test score is not a perfect signal on ability at workplace.

People make their participation/job change decisions annually, reflecting the data collection frequency. Each year in the labor market begins with job search. Then three different shocks related to job choice are realized: that is, each individual gets a new job offer; the individual's job is exogenously destroyed with a probability, δ, if he is employed; lastly, he gets a recall offer from the most recent employer with a probability, λ_R, if he is unemployed and has ever worked. After observing his preference shocks concerning jobs, he chooses the best career option available. After then, he chooses the level of consumption and savings given his current assets and earnings. Finally, the person observes a productivity signal at the end of the period, and his (and all employers') beliefs on his ability and job match quality are updated in the standard Bayesian way. This process is repeated for T years before he permanently retires at age 18+T.

2.3 Ability, Job Match Quality, and Prior Information

Individuals have different levels of ability at age 18: θ_i.[10] Ability is not directly observed and follows a normal distribution in the population. Without loss of generality, I standardize this distribution, while allowing the effect of standardized ability on productivity signal (ω_θ)

[9] I assume the length of life after retirement is 20 years for all people. This post-retirement period prevents more able workers from retiring earlier than less able workers.

[10] This model does not distinguish between innate ability and initial skill at age 18. As long as both innate ability and initial skill affect further skill accumulation, we can treat them equally in a model of post-schooling wage growth and career choices.

to be different from one:

$$\theta_i \sim N(0,1)$$

Before labor market entry (or at high school graduation), each individual receives a test score, θ_i^{TS}, which provides partial information on ability. That is, this test score is a noisy measure of true ability:

$$\theta_i^{TS} = \theta_i + \zeta_i^{TS},\ \zeta_i^{TS} \sim N\left(0, \sigma_\zeta^2\right),$$

where ζ_i^{TS} is an idiosyncratic error. This test score is publicly accessible in the labor market, and all distributions are common knowledge.

After individual i receives his test score, θ_i^{TS}, the individual's belief on his own ability is updated in a standard Bayesian way. Each individual's belief on ability at labor market entry changes according to his observed test score. The uncertainty shrinks down because of the new information, but the magnitude of the uncertainty remains the same across individuals. That is, after the initial signal on unknown ability, the individual i's believes that

$$\theta_i | \theta_i^{TS} \sim N\left(\frac{\theta_i^{TS}}{1+\sigma_\zeta^2}, \frac{\sigma_\zeta^2}{1+\sigma_\zeta^2}\right)$$

The quality of an employer-employee match, ϵ_{ij}, is randomly drawn from a normal distribution. This job match quality is constant over the job spell but (partially) uncertain at new job entry. I standardize the distribution of job match quality without loss of generality, while allowing the effect of standardized job match quality on productivity signal (ω_ϵ) to be different from one:

$$\epsilon_{ij} \sim N(0,1)$$

For expository convenience, I here briefly discuss on the nature of an ability measure in the data, the AFQT score, although this discussion is not really about the model. While the econometrician does not observe true ability, θ_i nor the test score described above, θ_i^{TS}, the econometrician does observe another ability measure, the AFQT score. I assume that the AFQT score is an ability measure which is similar to but not the same with the test score: that is, the AFQT score is also a noisy measure of true ability, whose measurement errors are independent of the measurement errors of θ^{TS}, but the AFQT score's variance of measurement errors is the same with θ^{TS}'s.

$$\theta_i^{AFQT} = \theta_i + \zeta_i^{AFQT}, \zeta_i^{AFQT} \sim N\left(0, \sigma_\zeta^2\right)$$

This assumption on the nature of the AFQT score is in line with Farber and Gibbons (1996) and Altonji and Pierret (2001), but different in two aspects: (1) agents have information not observed by the econometrician; (2) the econometrician has information not used by the agents, including the workers.[11]

2.4 Production Function

Individual i's log productivity (x_{ijt}) at job j at period t is explained by general work experience (E_{it}), job tenure (T_{ijt}), ability (θ_i) and job-specific match quality (ϵ_{ij}) at job j:

$$x_{ijt} = \omega_0 + g(E_{it}, \theta_i) + s(T_{ijt}, \theta_i) + \omega_\theta \theta_i + \omega_\epsilon \epsilon_{ij}$$

where E_{it} is individual i's experience at the beginning of period t (=0,..,t-1), T_{ijt} is i's job tenure at job j at the beginning of period t (=1,..,E_{it}), ω_0 is average productivity at labor market entry, $g(E_{it}, \theta_i)$ is general human capital which is a function of experience and ability, $s(T_{ijt}, \theta_i)$ is job-specific human capital which is a function of job tenure and ability, θ_i is individual i's ability, ϵ_{ij} is i's job match quality at job j.

Since investment is not directly observed, I use a simplified skill accumulation process,

[11] This second aspect sounds strong, but it is not really different from the Altonji and Pierret (2001)'s assumption on the nature of the AFQT score. If the AFQT scores were not available (or not verifiable) to the employers, the information carried out by the AFQT scores would not be considered in wage bargaining even if the workers had the information. Roughly speaking, this is a symmetric information version of the Altonji and Pierret (2001) assumption. In fact, the current assumption on the AFQT score is a weaker one than Altonji and Pierret (2001)'s because of the first aspect.

which is “learning by doing”—that is, productivity automatically grows over time. The two functions in the productivity equation capture learning by doing over experience and job tenure, respectively.

Individual i’s ability, θ_i , affects not only his productivity level but also his productivity growth over experience and job tenure (“learning ability”).[12] For example, high ability workers have higher productivity than other workers, and they can also learn both general and job-specific skills faster than others. This potential complementarity between ability and skill accumulation imposes nonnegativity restrictions on the first order derivatives of $g(E_{it}, \theta_i)$ and $s(T_{ijt}, \theta_i)$ with respect to θ_i.

For each learning by doing process, I use the standard quadratic function. Additionally, I assume that the peak time to be equal regardless of ability, which seems the most natural assumption.

$$g(E_{it}, \theta_i) = (1 + \omega_{g,\theta}\theta_i)q_g(E_{it})$$

$$s(T_{ijt}, \theta_i) = (1 + \omega_{s,\theta}\theta_i)q_s(T_{ijt})$$

where $q_l(x) = \omega_{l,1}x + \omega_{l,2}x^2$ if $x \geq 0$ for $l = g, s$.

The distributions of (normalized) θ and ϵ follow the standard normal distribution. The effects of these variables on productivity $(\omega_\theta, \omega_\epsilon, \omega_{g,\theta}, \omega_{s,\theta})$, however, can be different from one, so there is no loss of generality in the normalization of θ and ϵ. These coefficients are interpreted as the effects of one standard deviation (SD) increase in ability or job match quality on productivity.

[12] Previous studies distinguish between “ability to earn” and “ability to learn.” Also, many studies suggest that the two may be positively correlated (Rubinstein and Weiss, 2006).

The log productivity can be rewritten as a sum of certain and uncertain components:

$$\begin{aligned} x_{ijt} &= \omega_0 + q_g(E_{it}) + q_s(T_{ijt}) \\ &\quad + (\omega_\theta + \omega_{g,\theta} q_g(E_{it}) + \omega_{s,\theta} q_s(T_{ijt}))\theta_i + \omega_\epsilon \epsilon_{ij} \\ &= \underbrace{h(E_{it}, T_{ijt})}_{certain\ component} + \underbrace{\omega_{it}(E_{it}, T_{ijt})'\Theta_{ij}}_{uncertain\ component} \end{aligned}$$

Using a vector notation, I denote two unknown objects—ability and job match quality—as Θ_{ij}, which follows a bivariate standard normal distribution in the population in the first period. This is the prior beliefs of all people before they receive the test score information at high school graduation.

$$\Theta_{ij} = \begin{pmatrix} \theta_i \\ \epsilon_{ij} \end{pmatrix} \sim N(\mathbf{0}, I).$$

After the second period, the distribution of Θ_{ij} in the population is no more normal as job matches are selectively destroyed. Nonetheless, one's prior belief on Θ_{ij} at new job entry always follows a bivariate normal distribution because the renewed job match quality is independently drawn from the standard normal distribution.

True (log) productivity (x_{jit}) cannot be directly observed; instead, a noisy signal (y_{ijt}) of the true productivity is observed by the worker and all employers, at the end of each period.

$$y_{ijt} = x_{ijt} + \eta_{ijt}, \ \ \eta_{ijt} \sim N\left(0, \sigma_\eta^2\right)$$

2.5 Multidimensional Learning

After observing a new signal on productivity, each individual (and all employers) updates his (and their) beliefs on two unknown objects in a standard Bayesian way. This multidimensional learning mechanism has two major differences from the previous learning mechanisms. First, this generalized mechanism has two unknown objects, and only one of them is renewable. That is, this mechanism extends learning about job match quality (e.g., Jovanovic, 1979, 1984; Moscarini, 2005) by introducing permanent individual heterogeneity (ability) and learning about ability (e.g., Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lange, 2007) by incorporating endogenous job changes and a grouped signaling structure. Second, and more importantly, the signals on the unknown objects are correlated with each other. Unlike multiple unidimensional learning processes, this "multidimensional" learning process explicitly takes into account the possible correlation between the signals. Although job match quality is drawn independently of ability, workers' *belief* on job match quality can be affected by workers' unknown ability because the signal on job match quality is (perfectly) positively correlated with the signal on ability. As a result, (unknown) ability affects job change negatively—which occurs when perceived job match quality is below a certain cutoff.

I first consider a case of a worker who has stayed at his first job from his labor market entry until the beginning of period t without any spell of nonemployment ($1 \leq T_{ijt} = E_{it} = t-1$). The certain component is irrelevant so omitted, and subscripts i and j are suppressed where they are obvious.

$$\text{Signals up to the beginning of period t: } y_\tau = \omega'_\tau \Theta + \eta_\tau, \tau = 1, ..., T_{ijt} (= E_{it} = t-1)$$

where ω_τ (2×1 vector) is the weights on the unknown objects in the productivity signal observed in period τ.

$$\begin{aligned}
\text{Prior distribution: } & \Theta \sim N(\mathbf{0}, I) \\
\text{Posterior distribution: } & \Theta | \theta^{TS}, y_1, ..., y_t \sim N(\Theta_t, \Sigma_t) \\
\text{(first job entry) } & \Theta_0 = \begin{pmatrix} \theta^{TS}/(1+\sigma^2_\zeta) \\ 0 \end{pmatrix} \\
& \Sigma_0 = \begin{pmatrix} \sigma^2_\zeta/(1+\sigma^2_\zeta) & 0 \\ 0 & 1 \end{pmatrix} \\
\text{(after } t-1 \text{ signals) } & \Theta_{t-1} = \Sigma_{t-1}\left(\Sigma_0^{-1}\Theta_0 + (\omega' y)/\sigma^2_\eta\right) \\
& \Sigma_{t-1} = \left(\Sigma_0^{-1} + (\omega'\omega)/\sigma^2_\eta\right)^{-1}
\end{aligned}$$

where $\omega = (\omega_1\ \omega_2 ... \omega_{t-1})'$ ($(t-1) \times 2$ matrix), $y = (y_1\ y_2 ... y_{t-1})'$, ($(t-1) \times 1$ vector)

This expression is similar to Bayesian linear regression. Intuitively, we can interpret the signals (y) and the deterministic weights (ω) as "data" given to a worker and the unknown objects as unknown "parameters" which the Bayesian-rational agent wants to estimate. Note that the agent cannot separately identify the two unknown components from this on-the-job information if there is no over-time variation in the weights, although he can still learn about the weighted sum.

When a new signal y_t arrives at the end of period t, the beliefs are updated as follows:

$$\begin{aligned} \text{Posterior distribution after } t \text{ signals: } \Theta | y_t, \Theta_{t-1}, \Sigma_{t-1} &\sim N(\Theta_t, \Sigma_t) \\ \text{where } \Theta_t &= \Sigma_t \left(\Sigma_{t-1}^{-1} \Theta_{t-1} + \omega_t y_t / \sigma_\eta^2 \right) \\ \Sigma_t &= \left(\Sigma_{t-1}^{-1} + (\omega_t \omega_t') / \sigma_\eta^2 \right)^{-1} \end{aligned} \tag{1}$$

The equation (1) shows that the belief process is a first-order Markov process. Θ_{t-1} and Σ_{t-1} (and Σ_0) are the sufficient statistics for all $t-1$ signals and (weights).

The following lemmas are useful for describing the dynamic programming problem of each individual. In particular, the predictive distribution of the future posterior means is important for forming expectations and making decisions. Lemma 2 shows that the predicted future posterior mean is equal to current posterior mean, which is typical in Bayesian updating. This implies on-the-job wage growth due to stochastic components is a martingale, a key property shared across almost all learning models.

Lemma 1a: (sufficient statistics, $T_{ijt} = E_{it} = t-1$)

$$Pr(\Theta | \theta^{TS}, y_1, ..., y_{t-1}) = Pr(\Theta | \Theta_{t-1}, \Sigma_{t-1})$$

Lemma 2a: (the predictive distribution of the next period's posterior mean, $T_{ijt} = E_{it} = t-1$)

$$\Theta_t | \Theta_{t-1}, \Sigma_{t-1} \sim N\left(\Theta_{t-1}, (\Sigma_t \omega_t / \sigma_\eta^2)(\omega_t' \Sigma_{t-1} \omega_t + \sigma_\eta^2)(\omega_t' \Sigma_t / \sigma_\eta^2) \right)$$

It is not difficult to extend the above lemmas for a general case, $T_{ijt} \leq E_{it} \leq t-1$. First, if non-working periods exist ($E_{it} < t-1$), experience (E_{it}), instead of time in the labor market (t), matters for the learning process. If an individual does not work, there is no new information on his productivity provided to the labor market. In that case, the predictive distribution of the future posterior means degenerates to the current posterior means.[13] Second, when job changes are allowed ($T_{ijt} \leq E_{it}$), Lemma 1 and 2 still holds true with only slight modifications. If a worker chooses to stay ($D_{it} = 0$), Lemma 2b is equivalent to Lemma 2a. If the worker chooses to move ($D_{it} = 1$ or $j(i, E_{it}+1) \neq j(i, E_{it})$), his job match quality is replaced with the new job match quality that is randomly drawn from the standard normal distribution. This replacement is immediately reflected in the posterior beliefs. Once this immediate updating is completed, the Bayesian updating formula of Equation (1) is applicable within the new job spell. Subscript i is suppressed where it is not necessary.

Lemma 1b: (sufficient statistics, general case)

$$\begin{aligned} &Pr\left(\Theta_{j(E_t)} | \theta^{TS}, y_{j(1),1}, ..., y_{j(E_t),E_t}\right) \\ &= Pr\left(\Theta_{j(E_t)} | \Theta_{E_t}, \Sigma_{E_t}\right) \end{aligned}$$

where $j(e)$ is the job at experience e, $\Theta_{j(E_t)}$ is the vector of ability and job match quality at job $j(E_t)$, Θ_0 and Σ_0 are the posterior mean and covariance at first job entry (after the initial signals).

Lemma 2b: (the predictive distribution of the next period's posterior mean, general case)

$$\begin{aligned} \Theta_{j(E_t+1),E_t+1} | \Theta_{j(E_t),E_t}, \Sigma_{j(E_t),E_t}, D_t, P_t &\sim N\left(\Theta_{j(E_t),E_t}, P_t ABA'\right) \\ A &= \left(\Sigma_{j(E_t+1),E_t+1}\right)\left(\omega_{j(E_t+1),E_t+1}\right)/\sigma_\eta^2 \\ B &= \left(\omega_{j(E_t+1),E_t+1}\right)'\left(\Sigma_{j(E_t),E_t}\right)\left(\omega_{j(E_t+1),E_t+1}\right) + \sigma_\eta^2 \end{aligned}$$

where D_t is an indicator of job change at period t, that is $j(E_t+1) \neq j(E_t)$, and P_t is an

[13] The belief on job match quality is renewed at entry into a new job, not at the timing of the separation from the previous job because of recall possibility.

indicator of participation (or employment) at period t.

Although the worker's decision making process is not fully explained, it is possible to infer about the relationship between ability and job change.

Proposition 1 *(Ability and Job Change) Given the structure of the model and the above lemmas, unknown individual heterogeneity, ability, do not affect job change conditional on current beliefs. Ability, however, positively influences (new) productivity signals and mean belief on both ability and job match quality, so it negatively affects job change, unconditionally. As the uncertainty about ability dissolves over time, this initial negative selection into job change fades away. That is, if job change is characterized by a cutoff strategy* $(D_{it} = 1 \iff \epsilon_{ijt} < \epsilon_t^*(\Omega_{it}))$,

$$1. Pr[D_{it} = 1|\Omega_{it}, \theta_i] = Pr[D_{it} = 1|\Omega_{it}, \theta_{i'}], \forall i'$$

$$2. Pr[D_{it} = 1|\Omega_{it}^o, \theta_i] > Pr[D_{it} = 1|\Omega_{it}^o, \theta_{i'}] \text{ as } \theta_i < \theta_{i'} \text{ and } \sigma_{\theta,t}^2 > 0$$

$$3. Pr[D_{it} = 1|\Omega_{it}^o, \theta_i] \to Pr[D_{it} = 1|\Omega_{it}^o, \bar{\theta}] \text{ as } \sigma_{\theta,t}^2 \to 0$$

where $\Omega_{it} = \{\Theta_{i,j(i,E_{it}),E_{it}}, \Sigma_{i,j(i,E_{it}),E_{it}}, E_{it}, T_{ij(i,E_{it})t}, P_{it-1}, P_{it}\})$,
$\Omega_{it}^o = \{\Sigma_{i,j(i,E_{it}),E_{it}}, E_{it}, T_{ij(i,E_{it})t}, P_{it-1}, P_{it}\}$.

The multidimensional learning about ability and job match quality has interesting implications for job mobility and wage growth. First, unlike general or specific skill accumulation which explains wage growth *either* across all jobs *or* within a job, this generalized learning process explains both types of wage growth. Information acquired while working on the job, which was entirely job-specific in Jovanovic (1979), is relevant for productivity prospect at other jobs. This corresponds to public evaluation on worker ability in the real world—for example, recommendation letters. Unlike Altonji and Pierret (2001), this information is only partially transferable to other jobs.

Second, there are two different kinds of job experimentation, which are related to experience and job tenure, respectively. The first kind of job experimentation is learning about ability. Workers need precise information about their ability to plan a right career path for themselves, and the fastest way to attain this information is try many jobs as early as possible, in most parametrization.[14] This generates job hazard decreasing in work experience. Workers move more frequently when their experience is lower, and they become stabilized with more work experience.[15] The second kind of job experimentation is learning about job match quality. As in Jovanovic (1979), workers move more frequently when their job tenure is lower. Job hazard is thus (first increasing then) decreasing in job tenure.

Third, a negative selection into job mobility is predicted under this symmetric but multidimensional learning structure. Initially similar individuals receive different productivity signals over their career. Their beliefs on ability and job match quality eventually converge to the true values, but there exist non-negligible (finite sample) biases in their beliefs toward their prior beliefs until the uncertainty about ability is completely resolved. Especially, one's belief on job match quality is positively correlated with *unknown ability*—residual ability conditional on attained information. As a result, a job move is negatively correlated with unknown ability and "suboptimal" compared to the case where ability is fully known. A bad match can survive for a long time because of positively biased belief on job match quality (or negatively biased belief on ability), whereas a good match can be terminated very early because of underestimated match quality (or overestimated ability). This implies the phase of "job shopping" can be heterogeneous across individuals due to unknown ability. While

[14]This point is clearly presented in an extreme case where $\sigma_\eta^2 = 0$ with time-invariant weights on Θ. In that case, staying at a job would not add any more knowledge on ability. One can learn only by moving to another job. If $\sigma_\eta^2 > 0$, people learn considerably in the initial phase of new employment, but they get less and less information as job tenure increases.

[15]Miller (1984) and Antonovics and Golan (2012) describe occupational experimentation, which is essentially a Multi-Armed Bandit (MAB) problem. In their models, occupations are different in their degree of match uncertainty. People try more risky occupations when they are young because of the option value arising from occupation/job separation. While there is only one occupation in this paper, the amount of information on ability is a choice variable because the weights on unknown objects in productivity vary depending on one's chosen career path. The multidimensional learning mechanism in this paper, therefore, has a very similar implication to the previous studies: young people love risks—even if they are risk-averse.

this idea of negative selection into mobility based on learning has once been introduced in another context such as college-major choice (e.g., Arcidiacono, 2004), this paper is new in the sense that such a prediction is extended to a much more general situation. Moreover, the negative selection into job mobility gradually disappears as people eventually learn about their ability, which is directly testable if information on unknown ability is available to the econometrician.

Interestingly, this learning mechanism under symmetric information has very similar predictions with asymmetric employer learning (e.g., Greenwald, 1986; Schönberg, 2007; Pinkston, 2009). It is possible to construct two observationally-equivalent mechanisms, with regard to the patterns of selection into mobility and wage growth, based on symmetric multidimensional learning and based on asymmetric learning, respectively. Intuitively, public information on nontransferable productivity (ϵ_{ij}) is functionally similar to private information on transferable productivity (θ_i). One obvious advantage of this symmetric, multidimensional setting is the feasibility of extensions into a full structural model.

The multidimensional learning, especially the third point above, offers an interesting behavioral interpretation of the coexistence of match-enhancing job moves of young workers (e.g., Topel and Ward, 1992) and excessive job moves of low skilled workers (e.g., Gladden and Taber, 2009). It is well known that job mobility is an important way to increase job match quality and wages (job shopping). At the same time, some workers seem to move too much (job churning/floundering)—Gladden and Taber (2009) find low skilled workers change their job too frequently under lifetime income maximization hypothesis. Both kinds of job moves are well explained under the multidimensional learning hypothesis.

2.6 Search Technology and Partial Insurance

Each individual has to go through costly job search to get a new job offer in this frictional economy. Job search costs are the same on and off the job, c in utility unit. After job search, a new job offer arrives with a probability λ, again, on and off the job. The new job match

(ϵ') is randomly drawn from a distribution of job match quality ($\epsilon' \sim N(0, \sigma_\epsilon^2)$), and there exists no additional information on the quality of the new match at job offer, other than the population distribution which is common knowledge.

Whereas being unemployed does not provide any advantage for searching for a *new* job in this setting, an unemployed status does have an advantage for searching for an *old* job. A recall offer from the most recent employer arrives with a probability of λ_R only for the unemployed who ever worked. This possibility creates an option value of being unemployed as in Fujita and Moscarini (2013).[16] Recall possibility is also potentially more important in this multidimensional learning setting because it could be the only way get back to work for some workers. This also matters more with job-specific human capital accumulation.[17]

There also exist involuntary job separations. δ proportion out of currently working population undergoes an exogenous job destruction shock such as plant closing. Unemployment insurance partially insures against this shock. People who become involuntarily unemployed are eligible for unemployment benefits, which are 50 percent of previous earnings capped at \$400 (weekly) and \$10,000 (annually).[18] In addition, nonemployed workers including unemployed workers get welfare. I have a very simplified version of welfare: the benefits are \$5,000[19] (not means-tested); every individual who is not working is eligible for the benefits; and the take-up rate is assumed 100 percent. This number is in fact not important—I have another parameter, b, the leisure utility, which essentially does the same role—but the benefits guarantee individual income is above zero.

[16]Fujita and Moscarini (2013) report 20 percent of the workers who were permanently separated from their previous employer eventually return back to the employer. Following Fujita and Moscarini (2013), I assume that the recall offer is only from the most recent employer. That is, if the worker takes a new job offer, it is not possible to go back to the previous employer.

[17]The possibility of recall in a dynamic setting creates an incentive to stay unemployed, especially for workers with a good previous match or a high level of job-specific skill. If job-specific skill accumulation is faster for more able workers, they in this case have a stronger incentive to stay unemployed when their job has been destroyed by an exogenous shock, compared to other workers.

[18]The cap of unemployment benefits is different from state to state, but the average is about \$400 per week. Also, the benefits are provided up to 26 weeks in the U. S. The income replacement rate is approximately 50 percent.

[19]This number is borrowed from French and Jones (2011)—they find the average welfare benefits for households without a member over 65 is about 3,500 in 1998 dollars. Their number is converted into 2010 dollars using the CPI index.

2.7 Preference

Individual preference is time-separable, and the time discount factor, $\beta(< 1)$, is constant over time. Also, consumption and leisure are additively separable. The periodic utility from consumption follows a Constant Relative Risk Aversion (CRRA) function.

The periodic utility function is

$$U(C_{it}, S_{it}, D_{it}, P_{it}) = \frac{C_{it}^{1-\gamma} - 1}{1-\gamma} + c_s S_{it} + bP_{it} + \xi(S_{it}) + \kappa\xi(D_{it}, P_{it})$$

where S_{it} is an indicator of job search at period t, D_{it} is an indicator of job change at period t, P_{it} is an indicator of participation/employment, $\xi_t(S_{it})$ is a preference shock associated with job search, $\xi_t(D_{it})$ is a preference shock to work decisions.

The relative risk aversion, γ, is greater than or equal to 0. If $\gamma = 0$, the CRRA function becomes a linear utility function. If $\gamma = 1$, the utility function becomes a logarithm function. b denotes utility from leisure.

Each individual experiences sequential and transient preference shocks across job search and work alternatives ($\xi(S_{it})$ and then $\xi(D_{it}, P_{it})$), which all follow the standard Type I Extreme Value (Gumbel) distribution: $\xi(S_{it}), \xi(D_{it}, P_{it}) \sim i.i.d.$ Gumbel(0,1).[20] A sequential preference shock structure generates the same choice probabilities with the Nested Logit choice structure (Kennan, 2015). Preference shocks are often for a technical purpose, to make choice probabilities non-trivial ones, but they play economically meaningful roles here: the relative value of old and new matches over leisure utility change exogenously.

2.8 Wage Determination

The equilibrium wages are set at the *ex ante* expected productivity level given current beliefs, experience and job tenure. That is, $w^*_{ijt} = E[x_{ijt}|\Theta_{ijt}, \Sigma_{ijt}, E_{it}, T_{ijt}]$. As in Jovanovic (1979) and Jovanovic (1984), I assume firms have zero profits on average. In this case,

[20] CDF: $F(x) = e^{-e^{-x}}$.

workers fully internalize all benefits and costs, and their decisions on the match continuation/termination are efficient.

This "competitive" wage scheme is a (problematic) special case of generalized Nash-bargaining setup if vacancy creation is costless (Moscarini, 2005). The existence of match-specific surplus under search frictions leaves a difficult issue of how to divide the surplus between a worker and an incumbent employer in the presence of outside employers. As Jovanovic (1979) notes, there can be more than one possible wage function (mapping of productivity to a wage level). Following Jovanovic (1984), I assume that the worker gets all surplus all the time. By doing so, worker's decision problem becomes a much simpler one—there is no strategic motive arising from bargaining with incumbent and outside employers. In this case, (*ex ante*) efficient separation is achieved since workers fully internalize all benefits and costs into their decisions.

A wage contract is signed right after job acceptance—a new productivity signal, y_{ijt}, is observed at the end of the period. The econometrician observes a noisy measure of the latent earnings. The logarithm of the observed earnings (w_{ijt}) is a sum of the logarithm of latent earnings (w^*_{ijt}) and measurement error: $w_{ijt} = w^*_{ijt} + \nu_{ijt}$, where $\nu_{ijt} \sim i.i.d.N(0, \sigma^2_\nu)$.

2.9 Dynamic Programming Problem

Given the wage equation and learning process, the worker maximizes his lifetime utility, looking forward up to 60 years. The work decisions are only for the first T(=40) years.

$$V_t(\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1}) = \max_{\{(S_\tau, D_\tau, P_\tau)_{\tau=t}^T, (C_t, A_{t+1})_{\tau=t}^{T+R}\}} \sum_{\tau=t}^{T+R} \beta^{\tau-t} E[u(C_\tau)|\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1}]$$

The initial conditions other than the beliefs at period 1 are set all zeros: $A_1 = 0, E_1 = 0, T_{1,1} = 0, d^W_0 = 0$.

In each period, the first choice is job search (right after a new productivity signal on the last period's productivity). The value function at the beginning of period t is a Emax function of choice specific values:

$$V_t(\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1}) = \text{Emax } \{V_t^S(\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1}) + \xi_t^S - c_s,$$

$$V_t^{NS}(\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1}) + \xi_t^{NS}\}$$

where V_t^S is the choice-specific value of job search, V_t^{NS} is the choice-specific value of no job search, $\xi_t^{S,NS}$ are transient preference shocks (known at search choice), c is the sum of job search costs, Θ_t is posterior beliefs (mean), Σ_t is posterior beliefs (covariance matrix), A_t is assets, E_t is work experience at the beginning of period t, T_{jt} is job tenure at job j at the beginning of period t, and P_{t-1} is employment status at the beginning of period t (0: not working, 1: working).

The sets of possible shocks and available options change depending on current employment status and job search. First, the option of a new job is available only after a (successful) job search. A success of job search can be stochastic with a probability of λ.[21] Second, only the currently employed ($P_{t-1} = 1$) experience an exogenous job destruction with a probability of δ. Third, only the currently non-employed ($P_{t-1} = 0$) can have a recall offer from their most recent employer with a probability of λ_R.

$$V_t^S(\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1} = 1) = \lambda[(1-\delta)\text{Emax } \{V_t^{Old}, V_t^{New}, V_t^U\} + \delta\text{Emax}\{V_t^{New}, V_t^{UB}\}]$$

$$+ (1-\lambda)[(1-\delta)\text{Emax } \{V_t^{Old}, V_t^U\} + \delta V_t^{UB}]$$

$$V_t^{NS}(\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1} = 1) = (1-\delta)\text{Emax } \{V_t^{Old}, V_t^U\} + \delta V_t^{UB}$$

[21] In a discrete-time setting with annual frequency, it must be rare for anyone not to receive any offer. This is one reason why I use job search costs.

$$V_t^S(\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1} = 0) = \lambda[\lambda_R \text{Emax}\ \{V_t^{Old}, V_t^{New}, V_t^U\} + (1-\lambda_R)\text{Emax}\{V_t^{New}, V_t^U\}]$$

$$+ (1-\lambda)[\lambda_R \text{Emax}\ \{V_t^{Old}, V_t^U\} + (1-\lambda_R) V_t^U]$$

$$V_t^{NS}(\Theta_t, \Sigma_t, A_t, E_t, T_{jt}, P_{t-1} = 0) = \lambda_R \text{Emax}\ \{V_t^{Old}, V_t^U\} + (1-\lambda_R) V_t^U$$

where V_t^{New} is the choice-specific value of job change, V_t^{Old} is the choice-specific value of working with no job change, V_t^U s the choice-specific value of not working, V_t^{UB} is the choice-specific value of not working with unemployment benefits, λ is the probability of job offer arrival, δ is the probability of job destruction, and λ_R is the probability of recall offer arrival.

Individuals adjust consumption level after their employment status is fixed. Each work-choice specific value is given as follows:

$$\begin{aligned}
V_t^{New} =& \underset{A_{t+1}}{\text{Max}}\ \{u((1+r)A_t + W^*_{j't} - A_{t+1}) \\
&+ \beta E_{\Theta_{t+1}|\Theta_t,\Sigma_t,move}[V_{t+1}(\Theta_{t+1}, \Sigma_{t+1}, A_{t+1}, E_t + 1, 0, P_t = 1)]\} + \kappa\xi^{New} \\
V_t^{Old} =& \underset{A_{t+1}}{\text{Max}}\ \{u((1+r)A_t + W^*_{jt} - A_{t+1}) \\
&+ \beta E_{\Theta_{t+1}|\Theta_t,\Sigma_t,stay}[V_{t+1}(\Theta_{t+1}, \Sigma_{t+1}, A_{t+1}, E_t + 1, T_{jt} + 1, P_t = 1)\} + \kappa\xi^{Old} \\
V_t^{U} =& \underset{A_{t+1}}{\text{Max}}\ \{u((1+r)A_t + 5,000 - A_{t+1}) + b \\
&+ \beta V_{t+1}(\Theta_t, \Sigma_t, A_{t+1}, E_t, T_{jt}, P_t = 0)\} + \kappa\xi^{U} \\
V_t^{UB} =& \underset{A_{t+1}}{\text{Max}}\ \{u((1+r)A_t + 5,000 + UI - A_{t+1}) + b \\
&+ \beta V_{t+1}(\Theta_t, \Sigma_t, A_{t+1}, E_t, T_{jt}, P_t = 0)\} + \kappa\xi^{UB}
\end{aligned}$$

where $W^*_{j't}$ is (latent) labor earnings at a new job j , W^*_{jt} is (latent) labor earnings at current job j, b is leisure utility, UI is unemployment benefits (only for exogenous job separations), r is real interest rate, ξ's are transient preference shocks (realized right before work choice), and κ is a scale parameter.

3 Data

3.1 The National Longitudinal Study of Youth 1979 (NLSY79)

The data used in this analysis is the National Longitudinal Study of Youth 1979 (NLSY79), Round 1-25. The NLSY79 surveys a nationally representative sample of young men and women, including minority, poor and military oversamples, who were 14-22 years old in 1979. The data was collected annually from 1979 to 1993 and then biannually from 1994.

The NLSY data offers an opportunity to study how ability interacts with career decisions of young people in the labor market. First, almost all respondents took the Armed Services Vocational Aptitude Battery (ASVAB) at the very beginning of the survey.[22] In addition, several scales of non-cognitive traits near labor market entry are available. For example, Rosenberg's Self-Esteem Scale and Rotter's Internal-External Locus of Control Scale in 1980 are available—these scores are recently used as measures of non-cognitive ability in several studies (e.g., Heckman et al., 2006).

Second, the NLSY data provides an event history of each respondent's education and work decisions. For example, the NLSY79 has kept weekly arrays of all jobs ever held for all respondents, from which job changes can be detected very accurately. Even though someone have missed several rounds of interviews, the respondent was asked about the missing work history at a later interview. More details on the construction of ability measures and education/work history are in Appendix 1.

To avoid any complication from oversampling, gender and racial issues, I focus on the Cross-sectional White Male sample from the NLSY 1979 data (2,236). I remove all records from the respondents whose highest grade completed was less than 12 years at their first

[22] The ASVAB was administered to 11,914 NLS respondents (94 percent) during July through October of 1980 for the purpose of establishing a new national norm of the test (NLSY Attachment 106). ASVAB score are used to determine eligibility and assignment qualifications for specific military jobs for new enlistees, and the AFQT score, the sum of four subsection scores (word knowledge, paragraph comprehension, arithmetic reasoning and numeric operations), is a general measure of trainability and a primary criteria of enlisted eligibility for the Armed Forces (NLSY Attachment 106).

entry into the labor market or from those with a G.E.D. (-425) because high school dropouts and G.E.D.'s are reported to show very different behavioral patterns from other educational groups. The respondents whose post-schooling work history cannot be correctly constructed are also dropped from the sample (-24)—this is either because their first graduation year is not specified from their records or because their first graduation year is before 1975. I further drop individuals if they have ever joined the Military Services (-263) throughout the surveys. This is very conventional but has a specific meaning in this paper, which is related to the assumption on the AFQT score. All records from individuals are dropped if any of five ability measures—AFQT, ASVAB Verbal, ASVAB Math, Rotter's Internal-External Scale and Rosenberg's Self-Esteem Scale(1980)—are missing (-145), or if any parent's education is missing (-61). Finally, I consider weekly earnings less than $1 or greater than $10,000 as missing to eliminate potentially influential data points.

The final sample (all education levels) has 1,294 individuals and 27,218 person-year observations after high school graduation (25,257 observations after the first labor market entry). Table 1 has the descriptive statistics of the final sample. Among them, 634 individuals entered the labor market as a high school graduate. Among the initially high school graduates, 131 finish at least some post-secondary education before the last interview.

I use the final sample of all education levels for the following discussion of observed patterns in the data—for the comparison purpose with the previous literature. In the estimation of the structural model, I use only the sample of high school graduates who never attained extra college education (503 individuals) because the model does not have educational decisions. Restricting the sample is also a (crude) way of controlling for occupational heterogeneity.

Table 1: Descriptive Statistics

Variable	Obs.	Mean	SD	Min	Max
Birth Year	27,218	60.53	2.19	57	65
Age at Interview	27,218	31.44	9.39	16	56
Father's Education	27,218	13.04	3.16	0	20
Mother's Education	27,218	12.54	2.23	0	20
Siblings	27,218	2.69	1.69	0	15
AFQT (age-adjusted)	27,218	0.03	0.98	-4.44	1.75
ASVAB Math (age-adjusted)	27,218	0.02	0.99	-3.22	2.77
ASVAB Verbal (age-adjusted)	27,218	0.02	0.98	-2.85	3.02
Rosenberg's Scale (age-adjusted)	27,218	-0.01	1.00	-2.89	2.65
Rotter's Scale (age-adjusted)	27,218	0.01	1.00	-2.28	3.48
College Enrollment (May 1)	27,218	0.14	0.35	0	1
Highest Grade Completed	27,218	13.98	2.24	12	20
H.G.C. at the First Graduation	27,218	13.79	2.13	12	20
H.G.C. at Last Interview	27,218	14.44	2.51	12	20
Potential Experience	25,264	12.95	9.19	0.04	38.5
Work Experience	25,264	11.60	8.69	0	38.1
Working	25,258	0.91	0.28	0	1
New Employer (within 52 weeks)	25,258	0.26	0.44	0	1
Job Turnover (within 52 weeks)	25,258	0.21	0.40	0	1
Employer Tenure	26,050	4.99	5.89	0.02	35.5
Weekly Earnings ($,'10)	24,793	996.1	881.2	1.17	9,992
Annual Labor Earnings ($,'10)	25,560	49,049	49,915	0	326,609
Net Assets ($, 2010)	15,627	184,052	415,343	-745,165	3,506,059
Married	27,218	0.51	0.50	0	1
Residential Area (Census)	27,218	2.36	1.01	0	4
SMSA	27,218	0.47	0.50	0	1
Urban	27,218	0.73	0.45	0	1
Unemployment Spell	27,218	5.66	18.73	0	1451

3.2 Observed Correlation Patterns from Linear Regressions

In this subsection, I report observed correlation patterns between an ability measure, the Armed Forces Qualification Test (AFQT) score, and various labor market decisions/outcomes. This explains the motivation why I introduce individual heterogeneity in job change and employment into a model of work decisions and earnings over the life cycle. Moreover, this intuitively shows how the structural model in the previous section can be identified from the data. Many correlation patterns presented here are already introduced in the literature, especially in the context of employer learning (e.g., Farber and Gibbons, 1996; Altonji and Pierret, 2001; Lange, 2007; Schönberg, 2007; Pinkston, 2009). I add more patterns unnoticed by previous literature and discuss the limitations of previous interpretations.

The Armed Forces Qualification Test (AFQT) score is a well-established measure of ability in the literature. It is a special[23] measure of ability which is highly correlated with both the entry level and the growth rate of individual earnings (Table 2, Column 1). The AFQT score and earnings growth over potential experience remains even after education level is controlled for (Table 2, Column 2) as already noticed by the previous studies (e.g., Farber and Gibbons, 1996; Altonji and Pierret, 2001). The AFQT score and earnings growth over potential experience are also positively correlative among high school graduates (Table 2, Column 3).[24] In addition, the positive correlation between the AFQT score and earnings is increasing but concave in potential experience (Table 2, Column 5) as noticed by Lange (2007).[25] That is, the AFQT score is associated with both higher earnings level and faster

[23]Other ability measures, such as Rotter's Locus of Control scale and Rosenberg's Self-Esteem Scale, are strongly positively correlated with earnings level but not significantly correlated with earnings growth. These self-reported measure are obviously measures of *known* ability.

[24]Unlike Arcidiacono et al. (2010), I find strong positive correlation between the AFQT score and earnings growth over potential experience among college graduates as well. The differences mainly arise from the empirical definition of "college graduates". Arcidiacono et al. (2010)'s definition is more conventional—exactly 16 years of education at the final survey. My definition is less conventional—the individuals who entered the labor market after 16 years of education with no extra education. That is, I remove all records from the individuals who worked first and then finished college. This is because (extra) educational decision is highly correlated with ability. Of course, the current way of controlling for such bias is not fully satisfactory.

[25]Lange (2007) uses this pattern for estimating the speed of employer learning.

earnings growth but the difference in earnings growth disappear over time.

Ability matters for entry level earnings as well as subsequent earnings growth, which is a well-established empirical pattern in the literature. Then, what is the mechanism(s)? Those correlation patterns are well matched with an (employer) learning hypothesis if the AFQT score has some information on unknown ability. For example, Altonji and Pierret (2001) assume that the AFQT score was not available to employers.[26] After then, they interpret the positive correlation between the AFQT score and earnings-experience interaction, along with the negative correlation between the AFQT score and education-experience interaction, as evidence of employer learning. This implies that the positive correlation between the AFQT score and earnings growth over experience, conditional on high school graduates, is also from employer learning. This implication (not the Altonji and Pierret (2001)'s original test), however, can be challenged by alternative explanations with similar predictions. For example, if more able workers accumulate general or job-specific skills faster than others, we will observe the same patterns (learning ability and differential skill production). Moreover, career decisions such as job change, movements into and out of nonemployment, can be all affected by ability.

Job change is important for wage growth (Topel and Ward, 1992), and ability may affect earnings through job changes. Table 3 shows that the AFQT score is negatively correlated with job change, and the negative correlation disappears over the life cycle (Column 1). A job change is empirically defined by working at interview date with job tenure less than 52 weeks, conditional on working 52 weeks before. This disappearing negative correlation over the life cycle is true even after education is controlled (Table 3, Column 2-4). This is not likely totally driven by negative selection into unemployment. When we compare job-to-job changes versus all job changes, the correlation between the AFQT score and job-to-job change (vs. all job changes) is increasing positive (Table 4, column 1), but the correlation becomes small and statistically not different from zero conditional on high school graduates (Table 4,

[26]Lange (2007) has a detailed discussion on this assumption.

Table 2: AFQT Score and Labor Earnings

Dependent Variable: Log Real Weekly Earnings

	(1) All	(2) All	(3) HSG	(4) CLG	(5) All
AFQT	0.092***	0.039***	0.041***	0.043	-0.007
	(0.008)	(0.009)	(0.011)	(0.025)	(0.013)
AFQT x P. Exper. /10	0.035***	0.035***	0.028***	0.067***	0.127***
	(0.005)	(0.006)	(0.007)	(0.017)	(0.018)
AFQT x P. Exper.^{2} /100					-0.029***
					(0.005)
Educ.		0.074***			0.075***
		(0.005)			(0.005)
Educ. x P. Exper. /10		-0.001			-0.002
		(0.003)			(0.003)
Observations	23,319	23,319	10,640	4,960	23,319

From the NLSY79, Round 1-25. Cross-Sectional, White Male Sample.
a. AFQT score is standardized within each birth year group. Mean 0 and SD 1.
b. All specifications control for a cubic in potential experience, year fixed effects, regional fixed effects, parents' education and the number of siblings.
Standard errors in parentheses.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

column 3). Job-to-job change here is defined by a job change without any non-employment spells within 52 weeks.

Table 3: AFQT Score and Job Change

Dependent Variable: Job Change (last 52 weeks) | Work

	(1) All	(2) All	(3) HSG	(4) CLG
AFQT	-0.032***	-0.029***	-0.027***	-0.034*
	(0.005)	(0.006)	(0.007)	(0.015)
AFQT x P. Exper. /10	0.013***	0.011**	0.009*	0.012
	(0.003)	(0.003)	(0.004)	(0.010)
Educ.		-0.004		
		(0.003)		
Educ. x P. Exper. /10		0.002		
		(0.002)		
Observations	23,028	23,028	10,332	4,944

From the NLSY79, Round 1-25. Cross-Sectional, White Male Sample.
a. AFQT score is standardized within each birth year group. Mean 0, SD 1.
b. All specifications control for a cubic in potential experience, year fixed effects, regional fixed effects, parents' education, the number of siblings.
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Since job changes includes both job-to-job changes (E-E') and job changes with a gap (E-N-E'), the correlation between the ability measure and (non)employment also needs to be

Table 4: AFQT Score and Job-to-Job Transitions

Dependent Variable: Job-to-Job | Job Change (last 52 weeks)

	(1) All	(2) All	(3) HSG	(4) CLG
AFQT	-0.005	0.001	0.014	-0.022
	(0.009)	(0.010)	(0.013)	(0.028)
AFQT x P. Exper. /10	0.025**	0.018	0.010	-0.004
	(0.009)	(0.010)	(0.012)	(0.032)
Educ.		-0.007		
		(0.006)		
Educ. x P. Exper. /10		0.008		
		(0.005)		
Observations	6,675	6,675	3,048	1,411

From the NLSY79, Round 1-25. Cross-Sectional, White Male Sample.
a. AFQT score is standardized within each birth year group. Mean 0, SD 1.
b. All specifications control for a cubic in potential experience, year fixed effects, regional fixed effects, parents' education, the number of siblings.
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

examined. Previous studies report that job loss has long-lasting negative impacts on labor earnings. The positive selection into employment gets clearer over time (Table 5, Column 1-2). Less able workers are less and less likely to work as time in the labor market goes on, especially those among high school graduates (Table 5, Column 3). Among college graduates, such tendency is less clear (Table 5, Column 4). These patterns clearly show that employment status is also affected by ability. Also, this selection is likely to arises at nonemployment to employment (N-to-E) transitions rather than at employment to nonemployment (E-to-N) transitions.[27]

Job change and employment status decisions are heterogeneous by an ability measure, at least within the sample used here.[28] These patterns show that we need a more flexible model of life-cycle career decisions. Allowing individual heterogeneity in career decisions is

[27]This result is not fully tested yet. A crude analysis using employment status changes between interview dates show that N-E transition is significantly different by ability (especially at later career), while E-N transition is not that different by ability. This result needs to be tested using fixed and shorter time intervals.

[28]This result needs further robustness checks because these correlation patterns are not fully established in the literature. For example, Schönberg (2007) finds the negative correlation between the AFQT score and job mobility among college graduates, but not among high school graduates. Her definition of education groups are very different from mine: high school graduates includes high school dropouts and G.E.D's; college graduates include those who attained additional education after first labor market entry.

Table 5: AFQT Score and Employment

Dependent Variable: Being employed (interview date)

	(1) All	(2) All	(3) HSG	(4) CLG
AFQT	0.009**	0.007*	0.013**	0.002
	(0.003)	(0.004)	(0.005)	(0.009)
AFQT x P. Exper. /10	0.007***	0.006**	0.008*	0.009
	(0.002)	(0.002)	(0.003)	(0.006)
Educ.		0.003		
		(0.002)		
Educ. x P. Exper. /10		0.000		
		(0.001)		
Observations	25,257	25,257	11,465	5,324

From the NLSY79, Round 1-25. Cross-Sectional, White Male Sample.
a. AFQT score is standardized within each birth year group. Mean 0, SD 1.
b. All specifications control for a cubic in potential experience, year fixed effects, regional fixed effects, parents' education, the number of siblings.
Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

important *per se*; also, it can produce useful insights to understand extremely heterogeneous earnings dynamics even if we cannot model all heterogeneity.

One possible way of modeling these heterogeneous job mobility and employment is to incorporate multi-dimensional learning about ability and job match quality, as presented in the model section. Information shocks on *perceived* productivity affect career decisions as in Jovanovic (1984). Information shocks on *misperceived* productivity explain the disappearing negative selection into job change.[29] The positive selection into employment is mainly explained by positive leisure utility and search frictions—especially, job search costs reinforce the positive selection into employment by raising the huddle at N-to-E transitions. The increasing positive selection, however, is explained by information shocks—less able workers work *more than they would do under full information* because of their incorrectly high belief on ability, and the negative selection into employment due to misperceived ability disappears over time.

[29]The observed correlation patterns between the AFQT score and job-to-job change are interpreted as evidence of asymmetric employer learning in the literature (e.g., Schönberg, 2007; Pinkston, 2009). This is not coincidence, and the difference between their asymmetric model and my symmetric model is probably smaller than it appears. For example, if I interpret a job-specific match quality, which is a *productivity* inapplicable to outside employers in this paper, as an *information on ability* unavailable to outside employers, a multidimensional learning story is then interpreted, very roughly, as a kind of asymmetric learning story.

In addition, putting all separate pictures together can give us a better idea about how to identify different mechanisms which are observationally-equivalent at single dimension. For example, a skill accumulation story (learning ability) and an information-updating story (learning about ability) predict the same thing on earnings growth—a positive association between ability and earnings growth over experience. In other words, more able workers can have faster earnings growth for two different reasons: because more able workers accumulate skills faster; or because their unknown ability gets revealed over time. It is not possible to know which story really have generated the observed pattern on earnings growth.

The extended versions of the two stories, have different predictions on job mobility. A combination of general and job-specific skill accumulation (or more generally, occupation-specific skill accumulation) predicts that more able workers are less likely to move and this difference in job hazard between more able and less able workers increases over time. While multidimensional learning about ability and job match quality also predicts more able workers are less likely to move, the difference in job hazard decreases over time. That is, the coefficient on (unknown) ability-experience interaction term in a (voluntary) job hazard equation have different signs in the two stories.

This last point can provide a source of identification between skill- and information-based mechanisms. This is important for the identification of the structural model presented in the previous section.

3.3 A Data Assumption: AFQT score is a measure of unknown ability

One last thing that should be mentioned in this data section is the measure of *unknown* ability. The identification argument mentioned above depends on the assumption that the econometrician has a measure of *unknown* ability—from the perspective of the agents. In this paper, I make a data assumption which is about the nature of the AFQT score: the AFQT score carries some information unknown to both the workers and employers.

This is similar to Altonji and Pierret (2001)'s assumption on the AFQT score, but I modify their assumption in two ways. First, there exists a test score (θ^{TS}), a measure of ability, that is available to all workers and employers but not for the econometrician (for example, the SAT/ACT score or high school GPA). Second, the AFQT score is another independent measure of ability that is available to the econometrician but not available to employers and unnoticed by (non-military) workers. As is well known, the AFQT score is highly correlated with the SAT/ACT score, but both scores measure ability with error; both carry some information on ability not contained in the other one.

The first modification is certainly relaxing Altonji and Pierret (2001)'s, but the second one is probably stronger. The AFQT score was in fact sent to the NLSY respondents although not directly to employers. I assume that non-military workers do not know how to interpret their AFQT score or just ignore it because the score is basically for the Military Services and they have another good measure of ability—this is plausibly true in reality. Also, this is why I restrict the sample for only civilians without any military experience throughout all surveys.

4 Identification and Estimation

4.1 Identification

The number of model parameters are not many; it is not very difficult to identify the structural model from the observed data. Nonetheless, not all parameters can be identified, so some are fixed: for example, relative risk-aversion(=0 or 1), time discounting factor (=0.95) and the real interest rates (=0.05)[30], the (annual) probability of a recall offer conditional on unemployed (=0.2)[31] are fixed; the variances of measurement error in test scores (σ^2_{ζ}, $\sigma^2_{\zeta'}$)

[30] $0.95 \times 1/1.05 = 0.9975 \approx 1$.

[31] Fujita and Moscarini (2013) report the probability of returning back to a previous employer after a permanent separation is about 20%. This number may be different from the recall offer arrival rate, but it must be very close to the arrival rate provided that returning to the previous employer must be the best option for workers who are currently unemployed.

are normalized at 1, which is equal to the normalized variance of ability in population.[32] The number of parameters to be identified after this is only 16.

Since there are many competing mechanisms, the remaining parameters are identified simultaneously rather than one by one. Here, I explain the identification by starting from a previous strategy on a related identification problem and extending the strategy.

If there were no learning in this model, Topel (1991)'s (or Altonji and Shakotko (1987)'s) identification strategy can be used to identify skill accumulation functions. Because of the (log-)linearity assumption, we can write $w_{ijt} = (1+\omega_{g,\theta}\theta_{it})g(E_{it})+(1+\omega_{s,\theta}\theta_{it})s(T_{ijt})+\omega_\theta\theta_{it}+\omega_\epsilon\epsilon_{ijt}+\nu_t$. ($\theta_{it} = E[\theta_i|\Omega_{it}]$, $\epsilon_{ijt} = E[\epsilon_{ij}|\Omega_{it}]$), Ω_{it} is all available information for individual i at the beginning of t, and ν_t is an measurement error.) If there were no learning ($\theta_{it} = \theta_i$ and $\epsilon_{ijt} = \epsilon_{ij}, \forall i$), g() and s() functions could be separately identified from within-job wage growth of job-stayers ($E_i[w_{ijt} - w_{ijt-1}]$) and the first wages of job-movers after job change ($E_i[w_{ijt}|j(i,t) \neq j(i,t-1)]$), following the strategy of Topel (1991) to identify s(). Even if skill accumulation is different by ability (g() and s() have θ_i as an argument), adding interacted terms between time and ability in the regression equation of within-job wage growth and the first wages of job-movers will identify the coefficients of complementarity between ability and skill production ($\omega_{\theta,g}, \omega_{\theta,s}$).

Because of multidimensional learning in the model, θ_{it} ($= E[\theta_i|\Omega_t]$) and ϵ_{ijt} ($= E[\epsilon_{ij}|\Omega_t]$) can be different from θ_i and ϵ_{ij}. In this case, within-job wage growth can come from information shock and selection into job change and nonemployment ($E_i[\omega_\theta(\theta_{it}-\theta_{it-1})+\omega_\epsilon(\epsilon_{ijt}-\epsilon_{ijt-1})] \neq 0$). With endogenous job change and individual employment choices, the expected change in the unobserved productivity within a job spell will not be zero because of selection effects. The first wages of job movers will be also affected by information shock and selection into nonemployment ($E_i[\omega_\theta\theta_{it}|j(i,t) \neq j(i,t-1)] \neq 0$)—this is in fact a source of potential bias in the Topel (1991)'s identification strategy.

The structural model describes the choices, and choice probabilities are important for

[32] This normalization is innocuous because the productivity effect (ω_θ) is adjusted corresponding to the population variance of perceived ability.

identification. Observed choice probabilities can help us to identify learning parameters (ω_θ, ω_ϵ, σ_η^2) along with search technology (δ,c) and preference (b) parameters. In particular, two signal-to-noise ratios are identified from over-time variation in job change probability and the difference in that over-time variation by the AFQT score ($Pr[D_{it} = 1|\theta_i^{AFQT}]$). Over-time variation in the variance of job movers' first wages after job change together with the signal-to-noise ratios can identify the three learning parameters. Employment-to-employment (E-to-E) transition ($Pr[P_{it} = 1|P_{it-1} = 1]$) identifies job destruction probability δ. Nonemployment-to-employment (N-to-E) transition ($Pr[P_{it} = 1|P_{it-1} = 0]$), along with E-to-E transition, identifies the leisure utility b. Although we do not observe job search choice probabilities, the difference between E-to-N and N-to-E transitions, especially by the AFQT score, reveals information on the search technology (c).

Some parameters have qualitatively the same predictions on both log earnings and choice probabilities, which makes it difficult to identify them separately. Note that both learning about job-specific match quality ($E_i[\epsilon_{ijt}-\epsilon_{ijt-1}]$) and job specific human capital accumulation ($s()$) predict increasing log earnings and decreasing job mobility in time on the job (and in the labor market). Hence, although over-time variation in job change probability has useful information on both s() and signal-to-noise ratios in the belief-updating process, the variation alone cannot separately identify the two mechanisms. With the aforementioned assumption on the AFQT score in the data section, I focus on the difference in the over-time variation by the AFQT score. The two mechanisms have different predictions as discussed in the previous section, so I can separately identify each mechanism. This endows the counterfactual analyses in the next section with empirical content.

In addition to the conditional means in log earnings for job- stayers and movers, conditional variances in log earnings also provide useful information for the variance of wage measurement error (σ_ν^2) and the uncertainty in the belief on ability.

4.2 Estimation: Indirect Inference

The estimation method is indirect inference, which is broadly a kind of the Method of Simulated Moments (MSM). Indirect inference is characterized by the use of an auxiliary model. The indirect inference estimator is the minimizer of the (optimally-weighted) distance between auxiliary models estimated from real data and simulated data. The estimator is consistent and as efficient as the Maximum Likelihood (ML) estimator if the auxiliary model contains all information of the true model (Gourieroux et al., 1993).

$$\hat{\Theta}_{II} = \underset{\Theta}{\operatorname{argmin}}\ [\hat{\beta} - \hat{\beta}(\Theta)]'W[\hat{\beta} - \hat{\beta}(\Theta)]$$
$$\hat{\beta}(\Theta) = \frac{1}{H}\sum_{h=1}^{H}\hat{\beta}_h(\Theta)$$

where W is the weighting matrix, $\hat{\beta}$ is the estimates of auxiliary model parameters from the real data, $\hat{\beta}(\Theta)$ is the estimates of auxiliary model parameters from the simulated data given Θ, H is the number of replications in simulation.

If the weighting matrix, W, is equal to the inverse of the covariance of true moments, the indirect inference estimator is:

$$\sqrt{n}(\hat{\Theta}_{II} - \Theta) \to^d N\left(0, (1 + 1/H)\left(\frac{\partial\hat{\beta}(\Theta)}{\partial\Theta}W\frac{\partial\hat{\beta}(\Theta)}{\partial\Theta}'\right)^{-1}\right)$$

The optimal distance between the estimated auxiliary models is calculated by using the inverse of the data moment covariance matrix as the weighting matrix (W). The covariance matrix is obtained by a block bootstrapping method ($b = 1,000$).

The auxiliary model in this paper is a system of seven regression equations that describes log earnings and log earnings squared for job-movers and job-stayers, respectively, and job change (E-to-E'), employment-to-employment (E-to-E) and nonemployment-to-employment (N-to-E) transition probabilities. All equations have the same explanatory variables: a constant, a quartic in potential experience, the AFQT score and its interaction terms with the quartic in potential experience, and lagged log earnings.

All seven equations are about changes, so it is difficult to match the levels. Hence, I use additional moments describing initial earnings and employment: cross-sectional mean and variance of initial earnings and initial employment rate at age 19. The regression coefficients from the seven equations and the additional data moments are in total $80(= 7 \times 11 + 3)$. Following Altonji et al. (2013), I use the same set of control variables for all equations as this equation system is a Seemingly Unrelated Regression (SUR) system.

The auxiliary model is estimated from both real and simulated data. Since the structural model does not have any yearly or regional change, I use predicted log earnings (obtained by first regressing log earnings on a quartic in regional and yearly fixed effects—the base is an urban area, an SMSA, Northeast region in Census and year 1979—and eliminating all yearly and regional fixed effects) in the estimation of the auxiliary model from real data. Also, I make simulated data equal to real data in all other respects. Whenever an observation is missing in real data, I consider the corresponding observation in simulated data as missing. In particular, the NLSY79 surveys are collected annually until 1994 and biannually afterwards; the same structure is imposed on the simulated sample.

The auxiliary model includes several discrete variables, which makes auxiliary model parameters (regression coefficients) discontinuous in true model parameters. To use a gradient-based minimization algorithm in the estimation procedure, the discrete variables need to be smoothed. Although the auxiliary model I use in this paper is similar to Altonji et al. (2013)'s, the structural model I use heavily depends on past decisions including employment status and savings.[33] I introduce importance sampling weights (e.g., McFadden, 1989) to the auxiliary model in this indirect inference procedure, following the way of Sauer and Taber (2013).

The procedure is as follows:

[33] Keane and Smith (2003)'s smoothing method in this case requires me to go back two periods and describe all conditional probabilities between the two periods, which is an extremely difficult task—especially when there are continuous state variables such as posterior beliefs and savings. I tried this method first, but I could not achieve enough smoothing regardless of the choice of smoothing parameters.

1. Given an initial set of structural parameters, $\Theta^{(1)}$, simulate a data set, $\Upsilon^{(1)}$. This simulated data set includes all unobservables, as well as all observables.

2. Evaluate the likelihood of the structural parameters given the simulated data: $\ell_i(\Theta^{(1)}; \Upsilon^{(1)})$.

3. Evaluate the likelihood of a new set of structural parameters given the same simulated data set, $\ell_i(\Theta^{(2)}; \Upsilon^{(1)})$.

4. Estimate a new set of auxiliary model parameters, $\beta(\Theta^{(2)})$, by using the ratio of the two likelihoods, $\frac{\ell_i(\Theta^{(2)};\Upsilon^{(1)})}{\ell_i(\Theta^{(1)};\Upsilon^{(1)})}$, as sampling weights.

5. Minimize the (optimized) distance between the two sets of auxiliary model estimates ($\beta(\Theta^{(2)})$ and β), repeating 3-4.

6. Update the initial set of parameters with the new minimizer ($\Theta^{(2)}$) and simulate a new data set, $\Upsilon^{(2)}$. Repeat 2-6 if necessary.

Although I need to specify a likelihood function in this case, the function in this setup is much easier to write down because I can directly use unobserved variables as noted by Sauer and Taber (2013)—I do not need the integration over unobserved variable, which is the main difficulty in using the Maximum Likelihood (ML) approach. In the evaluation of the likelihood, I include all state and decision variables—unobserved variables such as true ability, job search and savings decisions[34] and productivity signals as well as observed variables such as work decisions and labor earnings.

The algorithm used for the minimization is a quasi-Newton's method. The derivative-free simplex method is also used to verify the estimates.

4.3 Estimation Results

The estimation results are summarized in Table 6 and 7. Figure 1 and 2 show how the model fits the data. The risk-neutral workers case ($\gamma = 0$) is not meaningfully different from the risk-averse workers with savings case ($\gamma = 1$) in earnings growth and job mobility;

[34] Although savings are observed in the NLSY data, the variable is very noisy and available only for limited time periods.

however, employment levels are quite different between the two cases.

Table 6: Estimates: Risk-Neutral Case ($\gamma = 0$)

Parameter	$\hat{\theta}$	S.E.($\hat{\theta}$)	
ω_0	6.4552	(0.0231)	Initial Average Productivity
ω_1	0.0528	(0.0041)	General Skill Accumulation (1st order)
ω_2	-0.0017	(0.0001)	General Skill Accumulation (2nd order)
ω_3	0.0116	(0.0068)	Job-Specific Skill Accumulation (1st order)
ω_4	-0.0004	(0.0003)	Job-Specific Skill Accumulation (2nd order)
ω_θ	0.1783	(0.0169)	Productivity Effect of Ability (+1SD)
$\omega_{g,\theta}$	0.0457	(0.0346)	(Learning) Ability (+1SD) on General Skill Production
$\omega_{s,\theta}$	0.0962	(0.0813)	(Learning) Ability (+1SD) on Specific Skill Production
ω_ϵ	0.3227	(0.0147)	Productivity Effect of Job Match Quality (+1SD)
σ^2_ν	0.4886	(0.0343)	Variance: Wage Measurement Error
κ	0.4899	(0.0960)	Scale Parameter: Second Stage Preference Shock
b	13,493	(3,269)	Leisure Utility
σ^2_η	0.0721	(0.0056)	Variance: Noise in Productivity Signals
λ	0.8804	(0.0358)	Job Offer Arrival Probability
δ	0.0501	(0.0080)	Job Destruction Probability
c	10,351	(3,312)	Job Search Costs

Asymptotic standard errors are in parentheses.

Table 7: Estimates: Risk-Averse Case ($\gamma = 1$)

Parameter	$\hat{\theta}$	S.E.($\hat{\theta}$)	
ω_0	6.4556	(0.0225)	Initial Average Productivity
ω_1	0.0455	(0.0043)	General Skill Accumulation (1st order)
ω_2	-0.0015	(0.0002)	General Skill Accumulation (2nd order)
ω_3	0.0193	(0.0049)	Job-Specific Skill Accumulation (1st order)
ω_4	-0.0004	(0.0002)	Job-Specific Skill Accumulation (2nd order)
ω_θ	0.1691	(0.0109)	Productivity Effect of Ability (+1SD)
$\omega_{g,\theta}$	0.0321	(0.0152)	(Learning) Ability (+1SD) on General Skill Production
$\omega_{s,\theta}$	0.0319	(0.0176)	(Learning) Ability (+1SD) on Specific Skill Production
ω_ϵ	0.3504	(0.0196)	Productivity Effect of Job Match Quality (+1SD)
σ^2_ν	0.4783	(0.0345)	Variance: Wage Measurement Error
κ	0.6393	(0.0680)	Scale Parameter: Second Stage Preference Shock
b	0.4602	(0.0485)	Leisure Utility
σ^2_η	0.1154	(0.0123)	Variance: Noise in Productivity Signals
λ	0.8909	(0.0358)	Job Offer Arrival Probability
δ	0.0240	(0.0053)	Job Destruction Probability
c	0.2817	(0.1440)	Job Search Costs

Asymptotic standard errors are in parentheses.

Figure 1 and 2 compare log real weekly earnings, job change and employment status by observed ability between the real data (left panels) and the simulated data (right panels). These data moments are not directly used in the estimation; the auxiliary model comprises a set of equations describing year-to-year changes. The sample is split into two groups by observed ability—high AFQT (≥median) and low AFQT (<median) groups. In the graphs from the real data, high AFQT group shows not only higher levels but also faster growth of labor earnings than low AFQT group. Also, the former group moves less and works more than the other group throughout their working lives. It is well matched with the regression results in the previous section.

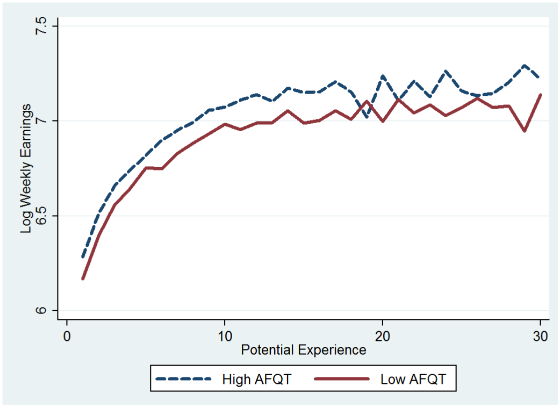
The estimated model (in both cases) fits the most important qualitative features with a relatively small number of parameters (16). First, average labor earnings increase over time at a deceasing rate, and the growth rate is significantly different across ability groups, up to almost .2 log points. Second, job mobility decreases over time at a decreasing rate, and less able workers move relatively more frequently. Third, employment ratio fluctuates around .9-.95. Employment ratio initially rises after labor market entry and then decreases afterwards—especially for low ability workers in the data.

There are several patterns in the data not very well explained by the model, which leave room for improvement. First, the curvature in the simulated age-earnings profile is not flexible enough to capture the rapid earnings growth during early career in the data. This is due to the functional form restriction on skill accumulation (quadratic). Second, the differences in employment ratio across ability groups, especially in late career are not very well matched. One reason is the large variances of employment-related moments in the data; as a result, employment ratio has only very small weights in the estimation procedure. Another reason is complicated incentives near retirement age. Relatively high employment ratio of senior workers with high ability (or relatively high nonemployment ratio of senior workers with low ability) is likely related to life expectancy, bequest motives and medical expenses (e.g., French and Jones, 2011) or disability insurance, which are all omitted in the current model.

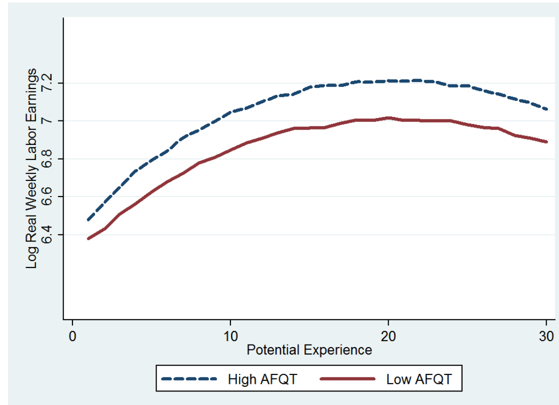
It is noteworthy that both job mobility and employment in the risk-neutral case are affected by *misperceived* productivity as expected. Less able workers move more because of their incorrectly low belief on job match quality. They work less because of their belief on the ability is on average low; however, they work more compared to the full information situation—that is, they are more likely to search and more willing to work because of their incorrectly high belief on ability, according to the model. As time passes on in the labor market, these incentives based on *misperceived* productivity disappear. In the risk-averse case with savings, employment levels are not very different by ability. One explanation would be life-cycle savings.

Figure 1: Model Fit (Risk-neutral case, $\gamma = 0$) (Left: Data, Right: Model)

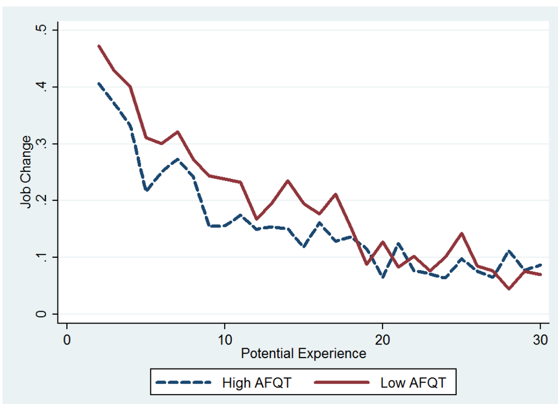
(a) Data: Log Earnings



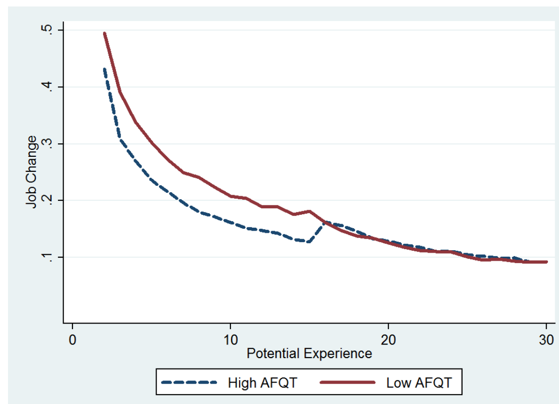
(b) Model: Log Earnings



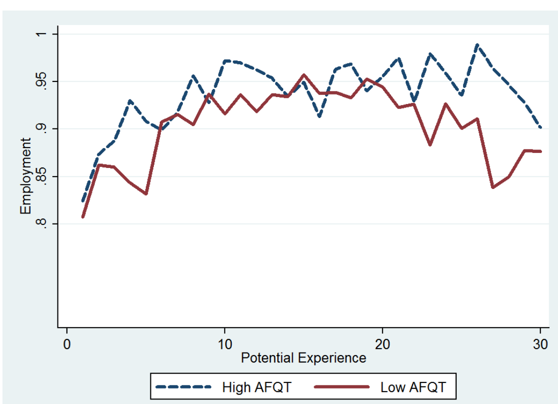
(c) Data: Job Change



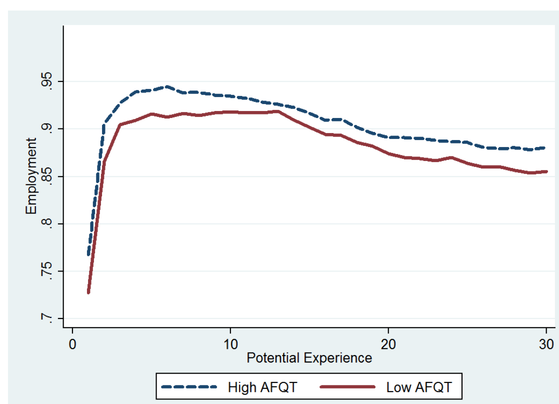
(d) Model: Job Change



(e) Data: Employment



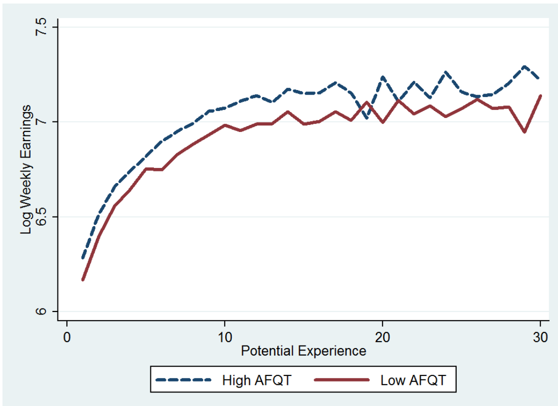
(f) Model: Employment



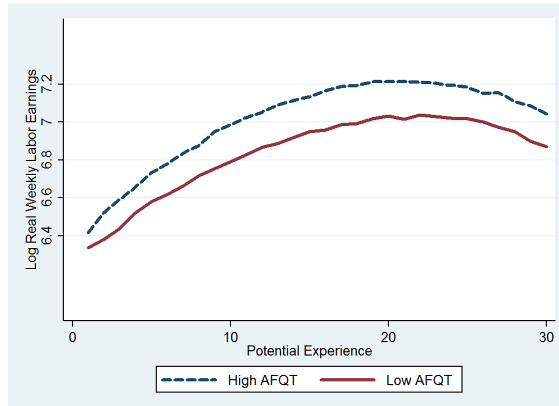
Note: Left panels are from real data. 1. log real weekly earnings is first regressed on a quartic equation in potential experience and its interaction terms with the AFQT score with other controls such as regional dummies (urban, SMSA, and Census regions) and yearly fixed effects. Then the estimated regional and yearly fixed effects are subtracted from the log earnings variable. 2. job change is empirically defined by working on the interview date and having changed to a new employer within 52 weeks. 3. employment variable is an indicator of working on the interview date. Right panels are from simulated data. The graphs show cross-sectional averages over potential experience. High AFQT: ≥med., Low AFQT: <med.

Figure 2: Model Fit (Risk-averse case with savings, $\gamma = 1$) (Left: Data, Right: Model)

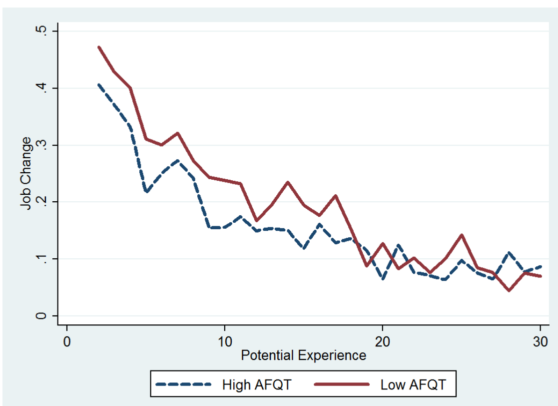
(a) Data: Log Earnings



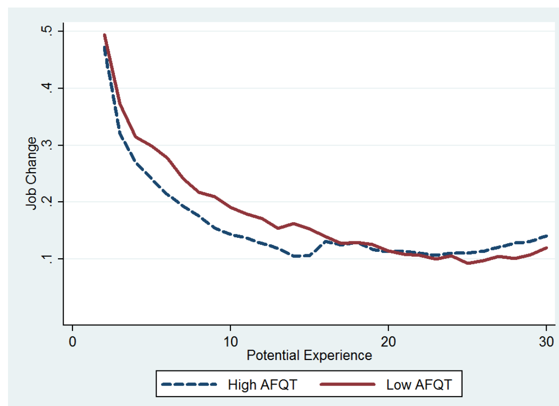
(b) Model: Log Earnings



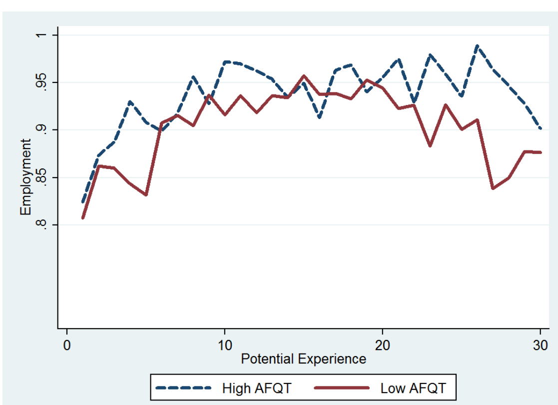
(c) Data: Job Change



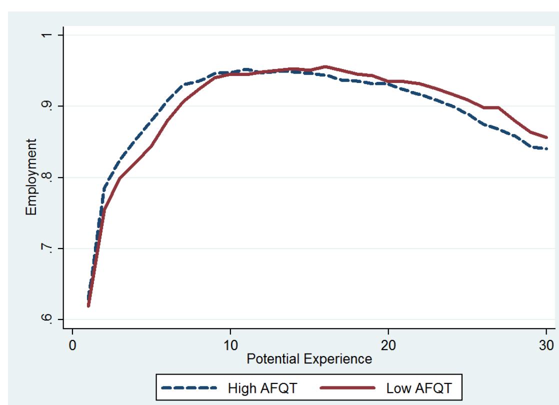
(d) Model: Job Change



(e) Data: Employment



(f) Model: Employment



Note: Left panels are from real data. 1. log real weekly earnings is first regressed on a quartic equation in potential experience and its interaction terms with the AFQT score with other controls such as regional dummies (urban, SMSA, and Census regions) and yearly fixed effects. Then the estimated regional and yearly fixed effects are subtracted from the log earnings variable. 2. job change is empirically defined by working on the interview date and having changed to a new employer within 52 weeks. 3. employment variable is an indicator of working on the interview date. Right panels are from simulated data. The graphs show cross-sectional averages over potential experience. High AFQT: ≥med., Low AFQT: <med.

5 Counterfactual Analyses: Earnings Dynamics

The estimated structural model is useful in many ways. One important use of the estimated model is to analyze earnings dynamics over the life cycle. By shutting down each channel, we can find out the importance of each channel on how earnings change over the life cycle.

Similar exercises has been done by many previous works, but two things are new here. First, the structural model in this paper can address both issues of average earnings growth and individual heterogeneity in earnings dynamics—especially in relation to individual heterogeneity in job mobility (job floundering). Understanding the mechanisms of earnings dynamics, especially individual heterogeneity, is clearly a key issue in labor economics, macroeconomics, and public policy debates. Second, the model can empirically investigate the importance of learning on earnings dynamics. Of course, this is possible only because of the assumption on the AFQT score. To the best knowledge of the author, this paper is new in both aspects.

Another important use of the model is to analyze policies from the perspective of the social planner. This is not done in this version of this paper and left for future research.

In this section, I perform several simulations on the sources of average earnings growth and heterogeneity in earnings growth. I only report the simulation results from the risk-averse case because the risk-neutral case is not very different.

5.1 Average earnings growth

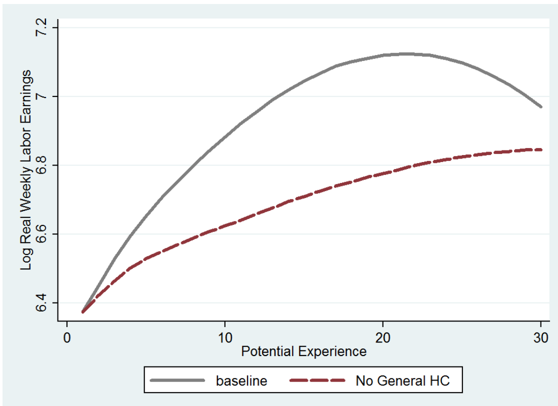
Earnings growth can arise from various channels, according to the model. Earnings may increase because of direct productivity improvement by general and/or job-specific skill accumulation. Earnings can also grow simply due to improved job match quality—on average, job match gets better over the life cycle when workers and employers learn something about their matches and terminate ones perceived as unproductive (job shopping). Individual earnings become more dispersed over time as ability is gradually reassessed after productivity signals in the labor market—individual earnings will diverge to the levels associated with their true ability (learning about ability). Of course, all components are interconnected as well.

Figure 3 shows how much each channel contributes to average earnings growth, by shutting down each channel one by one. Skill accumulation is the most important contributor to life-cycle earnings growth. General skill accumulation accounts for .12 log points increase in average earnings over the first 5 years and .28 and .34 log point increases over the first 10 and 20 years in the labor market (Figure 3 (a)). Job-specific skill accumulation explains .04, .10 and .14 log point increases over the first 5, 10 and 20 potential experience, respectively (Figure 3 (b)). General skill accumulation is surely more important, but the relative importance of job-specific skill accumulation increases over time. This is partly due to the possibility of recall to a previous employer.

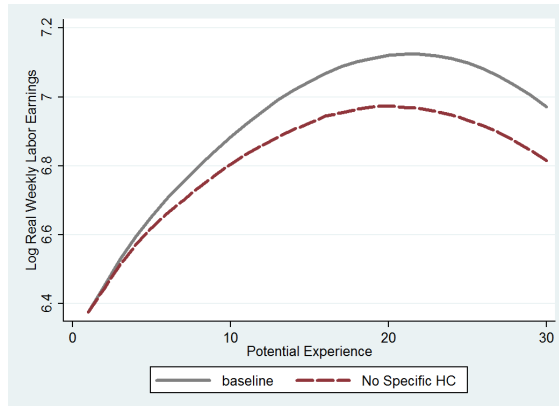
Job shopping also contributes considerably. Average job match quality rises quickly up to .32, .53 and .80 SD during the first 5, 10 and 20 years, respectively, and this is clearly associated with job mobility (Figure 4). That is, young workers select into stable employment relations through job changes as described in Topel and Ward (1992). This average match enhancement explains .10, .22 and .32 log point increases in earnings for 5, 10 and 20 years of potential experience (Figure 3 (c)). The contribution of job shopping to earnings growth is almost two times higher than other recent estimates (e.g., Altonji et al., 2013). The matching gains are partly associated with the possibility of recall, especially after 10 years of labor market experience (Figure 3 (d)). That is, good matches are saved thanks to recalls. The skill- and match- preserving effect of recall is very small during the first 10 years, but increasingly important, about .01 and .05 after 20 and 30 years in the labor market, respectively.

Information frictions significantly restrict average earnings growth. If there were no information friction, earnings would be up to .07 log points higher according to panel (f) in

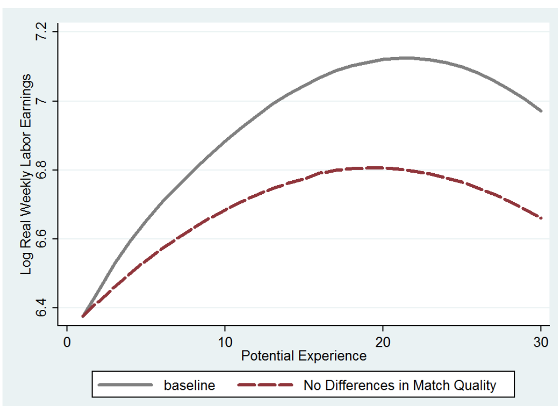
Figure 3: Various Reasons for Earnings Growth: Counterfactuals



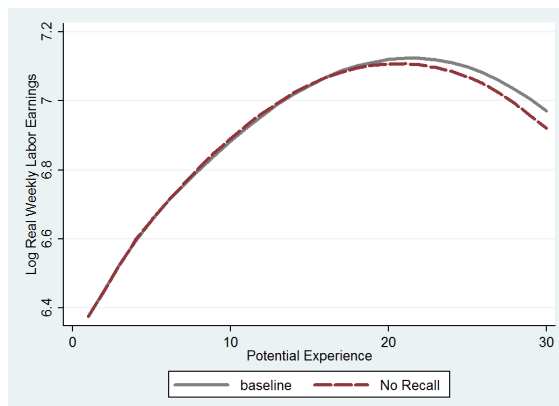
(a) No General Human Capital



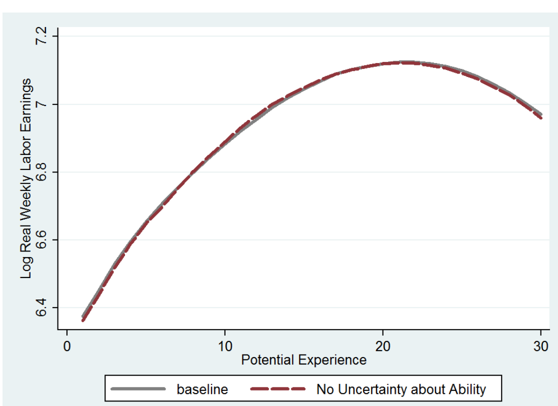
(b) No Job-Specific Human Capital



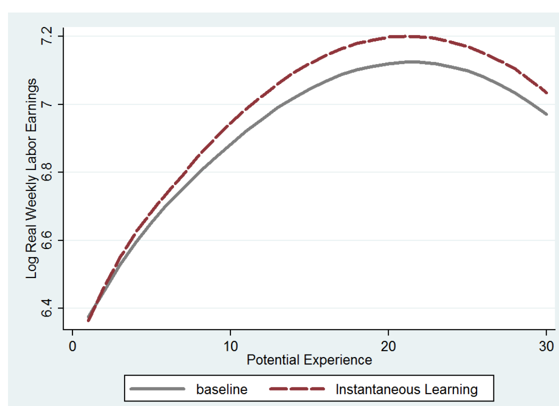
(c) No Job Shopping



(d) No Recall



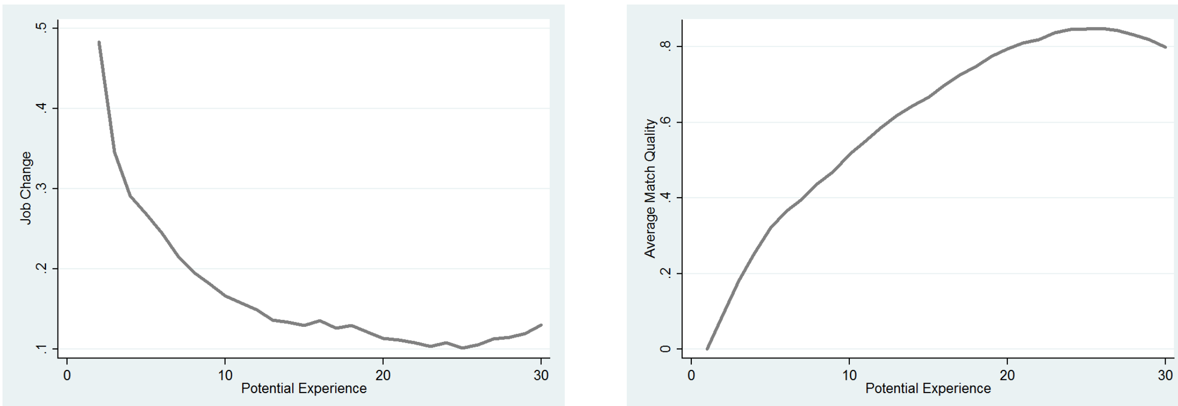
(e) No Uncertainty about Ability



(f) No Info. Friction (Full Info.)

Figure 3. On the contrary, the uncertainty in ability has almost no effect on average earnings growth (Figure 3 (e)), but it clearly affects the distribution of earnings growth—the next subsection explains how it affects individual heterogeneity in earnings growth.

Figure 4: Job Mobility and Average Job Match Quality over the Life Cycle



(a) Job Mobility

(b) Average Job Match Quality

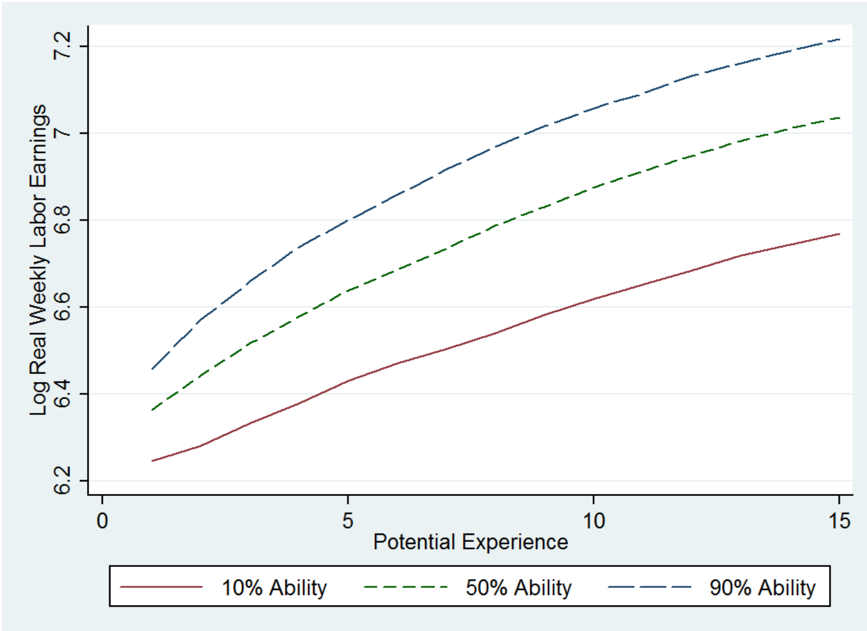
5.2 Individual heterogeneity in earnings growth

Decomposing average earnings growth into various sources provides useful information, but it can be even more interesting for policymakers to understand the reasons for heterogeneous earnings growth. If skill accumulation and information-updating processes are heterogeneous across individuals, the implications on labor market policies and income transfer programs can be surprisingly different from the policy implications of the previous analysis on average earnings growth. For example, job mobility contributes much to young workers' average earnings growth, but the effect can also be widely different across individuals. Labor market flexibility might have much lower or even negative impact on the earnings growth of less able workers.

Figure 5 shows how earnings growth is different by ability. The differences in ability are associated with not only entry-level earnings but also earnings growth over the life cycle.

The model decomposes the distribution of earnings growth into various channels (Figure 6). Ability affects earnings growth mainly through information-updating channels.

Figure 5: Heterogeneous Earnings Growth by Worker Ability

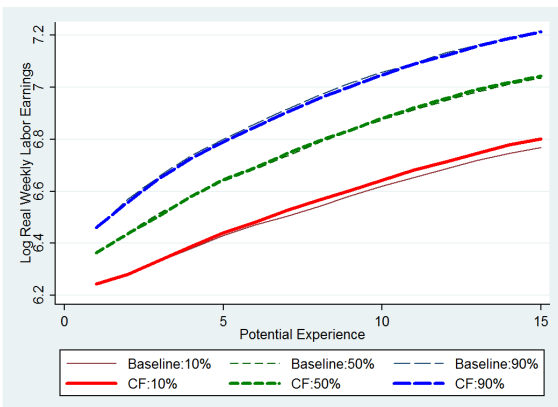


The first two panels, (a) and (b), in Figure 6 show two counterfactual experiments on skill accumulation. In the first counter-factual situation that everyone accumulates general skill at the same median speed (50 percentile in ability), we see only slight differences between the baseline and the counterfactual age-earnings profiles (Figure 6, (a)). In a similar experiment regarding job-specific skill accumulation (Figure 6, (b)), we also observe almost no difference between the baseline and the counterfactual profiles.

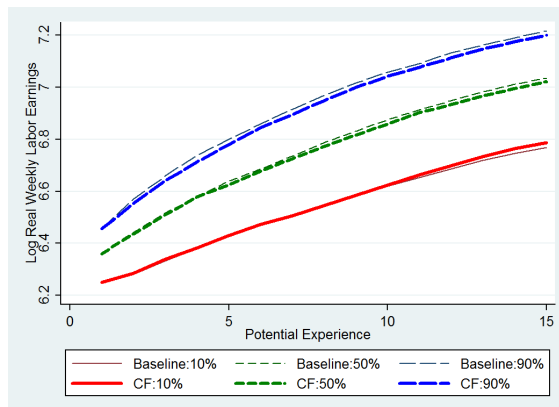
The third panel (Figure 6 (c)) shows another counterfactual experiment on uncertainty about ability. If ability were certain from the beginning of labor market experience, labor earnings would grow at almost the same speed for all workers. It is noteworthy that this would be achieved by increased earnings inequality among young workers.

Two different effects contribute to the increased earnings inequality among young workers with full information on ability. The first effect is a direct reassessment effect. With only partial information on ability, individual earnings level gradually converges to the level associated with true ability, which is similar to the story in employer learning (e.g., Altonji and Pierret, 2001). With full information on ability, earnings level would jump to the true

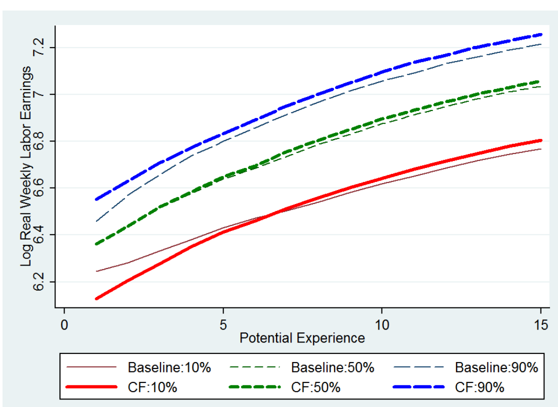
Figure 6: Various Reasons for Differential Earnings Growth by Ability: Counterfactuals



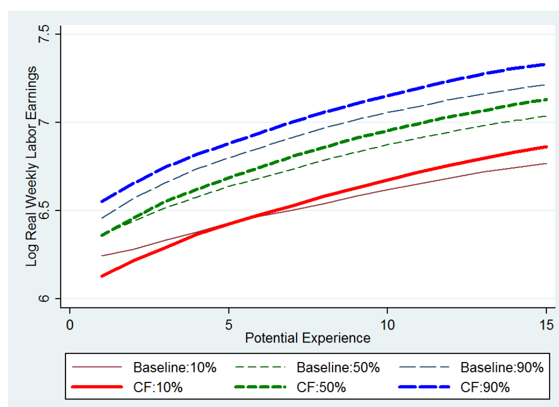
(a) No Differences in General Skill Accum.



(b) No Differences in Job Specific Skill Accum.



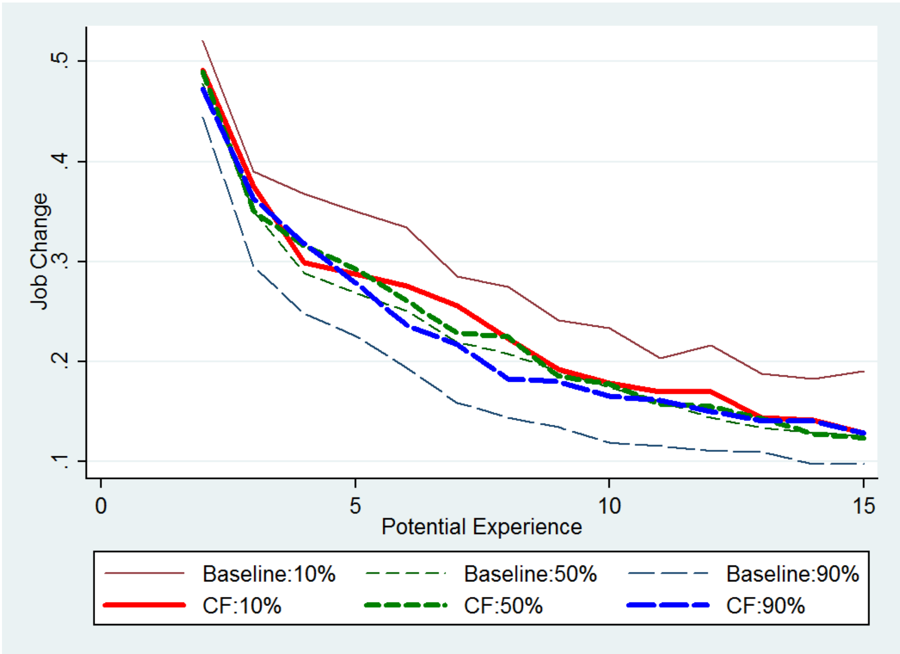
(c) No Uncertainty about Ability



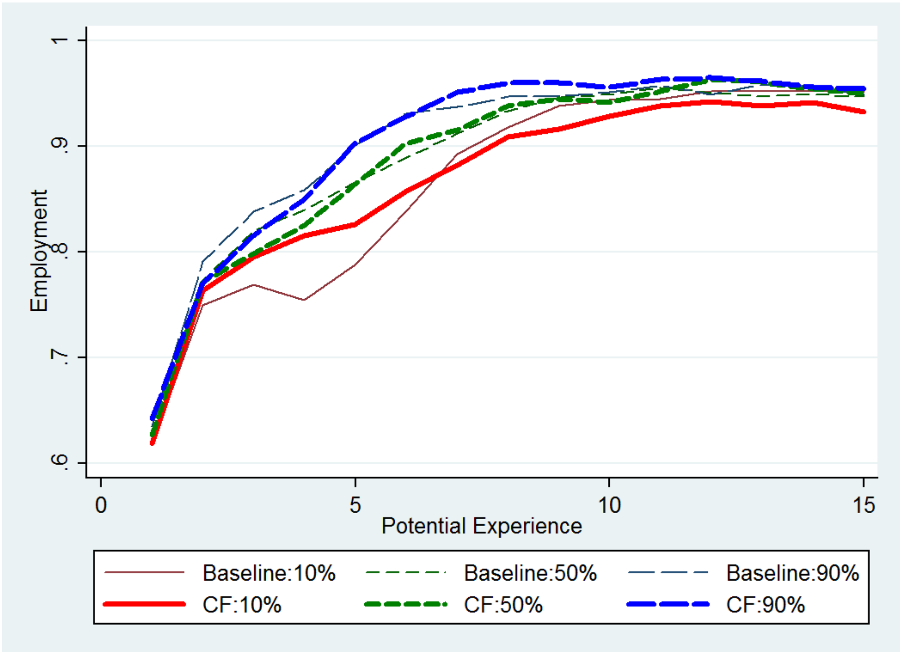
(d) No Info. Friction (Full Info.)

Figure 7: Job Mobility and Employment by Ability (CF: No Uncertainty about Ability)

(a) Job Mobility



(b) Employment



Note: Job change indicates working at a new job conditional on working.

level from the beginning.

The second effect is an indirect matching efficiency effect. When workers (and employers) cannot directly distinguish between ability and job match effects, negative selection into job mobility can arise as a result. For example, high ability workers misinterpret their good productivity signals in favor of their job match because they perceive themselves as mediocre workers (conditional on certain characteristics such as a test score). As a result, they tend to stay at a job even when it is desirable to move to another job. On the contrary, low ability workers tend to move to another job even when it is actually better for them to stay there. They misinterpret their bad productivity signals as signs of a wrong match because of their (partially) incorrect prior belief. This matching inefficiency gradually disappears over time as workers (and employers) learn more about their ability with more signals.

Figure 7 shows that the differences in job mobility across ability groups would disappear if initial uncertainty about ability were not there. More information would increase the overall efficiency in terms of utility for all workers, but the realized job match quality improvement could be highly heterogeneous across ability groups. If there were no uncertainty about ability, high ability workers would improve job match quality by more aggressive job shopping; low ability workers would either save moving costs or improve job match quality by reducing unnecessary job moves.[35] The median ability group would not show any differences in job match quality.

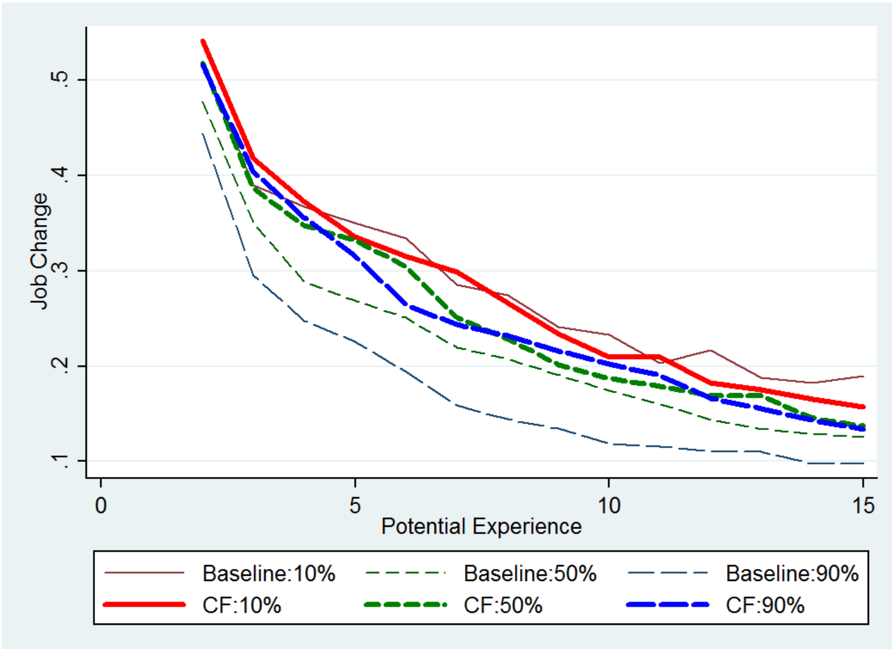
The panel (d) in Figure 6 suggests that full information would initially increase earnings inequality among young workers but improved job matching efficiency would eventually raise labor earnings of all workers above the baseline level. As predicted, job mobility would not be different across ability groups, but the average job mobility would be initially higher than in the previous case, by almost 10 percent point (Figure 8 (a)). This is related to increased amount of information in the labor market, and it leads to increases in average job match

[35]The risk-neutral case is meaningfully different from the risk-averse case only in this part. Less able workers in risk-neutral case do not gain much from more information on their ability while the same workers in risk-averse case do benefit from better information.

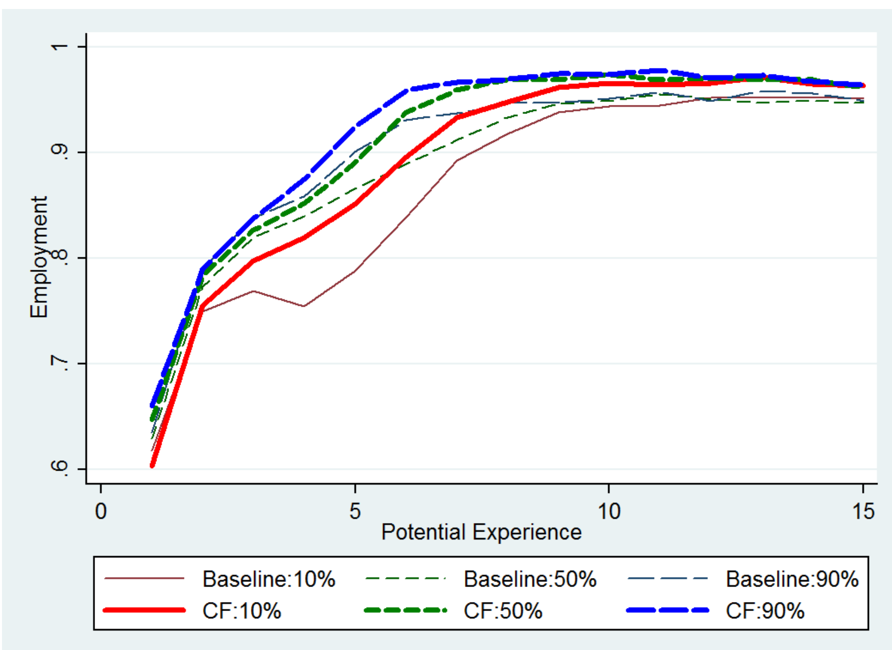
quality.

Figure 8: Job Mobility and Employment by Ability (CF: Full Information

(a) Job Mobility



(b) Employment



Note: Job change indicates working at a new job conditional on working.

6 Conclusion

This paper develops a model of life-cycle career decisions under various types of uncertainty, focusing on the roles of skill accumulation and information-updating in earnings dynamics. Workers accumulate general and job-specific skills over the life cycle, possibly at different speeds according to their ability; workers learn about both work ability and job match quality by trial and error. In this setup, skill accumulation and information updating processes can be interconnected; more importantly, various uncertainties can jointly affect work decisions, producing distinctive predictions such as a negative but disappearing correlation between ability and job change.

With an additional assumption that the AFQT score carries over some information on unknown ability, I estimated the model from a sample of white male high school graduates who did not attain any post-secondary education nor have any military experience throughout the NLSY79 surveys.

The estimated model shows that average life-cycle earnings growth develops from various sources: general skill accumulation accounts for approximately 33 percent points of earnings growth over the first 10 years; job-specific skill accumulation, 10 percent points; and job shopping, 24 percent points during the same period. The contribution of job shopping to earnings growth is almost two times higher than previous estimates, which suggests the importance of individual heterogeneity in understanding job mobility and life-cycle earnings growth. Recall to a previous employer plays a role during the later career. Information friction restricts average earnings growth considerably, about 7 percent points.

The model also provides an opportunity to look at individual heterogeneity in earnings growth. Information and uncertainty explain most of the difference in earnings growth across individuals: first, individual earnings converge to the level associated with true ability; second, job mobility is heterogeneous, and the improvement in job match quality through job shopping is different across ability groups. Skill accumulation is fairly homogeneous across individuals and explains little about individual heterogeneity in earnings growth.

References

Altonji, Joseph G and Charles R Pierret (2001), "Employer learning and statistical discrimination." *The Quarterly Journal of Economics*, 116, 313–350.

Altonji, Joseph G and Robert A Shakotko (1987), "Do wages rise with job seniority?" *The Review of Economic Studies*, 54, 437–459.

Altonji, Joseph G, Anthony A Smith, and Ivan Vidangos (2013), "Modeling earnings dynamics." *Econometrica*, 81, 1395–1454.

Antonovics, Kate and Limor Golan (2012), "Experimentation and job choice." *Journal of Labor Economics*, 30, 333–366.

Arcidiacono, Peter (2004), "Ability sorting and the returns to college major." *Journal of Econometrics*, 121, 343–375.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo (2010), "Beyond signaling and human capital: Education and the revelation of ability." *American Economic Journal: Applied Economics*, 2, 76–104.

Farber, Henry S (1999), "Mobility and stability: The dynamics of job change in labor markets." *Handbook of labor economics*, 3, 2439–2483.

Farber, HS and R Gibbons (1996), "Learning and wage dynamics." *Quarterly Journal of Economics*, 111, 1007–1047.

French, Eric and John Bailey Jones (2011), "The effects of health insurance and self-insurance on retirement behavior." *Econometrica*, 693–732.

Fujita, Shigeru and Giuseppe Moscarini (2013), "Recall and unemployment." Technical report, National Bureau of Economic Research.

Gladden, Tricia and Christopher Taber (2009), "Turnover and wage growth for low skilled young men." *Unpublished manuscript*.

Gourieroux, Christian, Alain Monfort, and Eric Renault (1993), "Indirect inference." *Journal of applied econometrics*, 8, S85–S85.

Greenwald, Bruce C (1986), "Adverse selection in the labour market." *The Review of Economic Studies*, 325–347.

Heckman, James J, Jora Stixrud, and Sergio Urzua (2006), "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." *Journal of Labor Economics*, 24, 411–482.

James, Jonathan (2011), "Ability matching and occupational choice." *FRB of Cleveland Working Paper*.

Jovanovic, Boyan (1979), "Job matching and the theory of turnover." *The Journal of Political Economy*, 87, 972–990.

Jovanovic, Boyan (1984), "Matching, turnover, and unemployment." *The Journal of Political Economy*, 108–122.

Keane, Michael and Anthony A Smith (2003), "Generalized indirect inference for discrete choice models."

Kennan, John (2006), "A note on discrete approximations of continuous distributions." *University of Wisconsin Madison*, 211–251.

Kennan, John (2015), "Spatial variation in higher education financing and the supply of college graduates." Technical report, National Bureau of Economic Research.

Kennan, John and James R Walker (2011), "The effect of expected income on individual migration decisions." *Econometrica*, 79, 211–251.

Lange, Fabian (2007), "The speed of employer learning." *Journal of Labor Economics*, 25, 1–35.

McFadden, Daniel (1989), "A method of simulated moments for estimation of discrete response models without numerical integration." *Econometrica: Journal of the Econometric Society*, 995–1026.

Miller, Robert A (1984), "Job matching and occupational choice." *The Journal of Political Economy*, 1086–1120.

Moscarini, Giuseppe (2005), "Job matching and the wage distribution." *Econometrica*, 73, 481–516.

Nagypál, Éva (2007), "Learning by doing vs. learning about match quality: Can we tell them apart?" *The Review of Economic Studies*, 74, 537–566.

Neal, Derek (1999), "The complexity of job mobility among young men." *Journal of Labor Economics*, 17, pp. 237–261.

Papageorgiou, Theodore (2014), "Learning your comparative advantages." *The Review of Economic Studies*, 81, 1263–1295.

Pinkston, Joshua C (2009), "A model of asymmetric employer learning with testable implications." *The Review of Economic Studies*, 76, 367–394.

Rubinstein, Yona and Yoram Weiss (2006), "Post schooling wage growth: Investment, search and learning." *Handbook of the Economics of Education*, 1, 1–67.

Sauer, Robert M. and Christopher Taber (2013), "Importance sampling in indirect inference."

Schönberg, Uta (2007), "Testing for asymmetric employer learning." *Journal of Labor Economics*, 25, 651–691.

Topel, Robert H (1991), "Specific capital, mobility, and wages: Wages rise with job seniority." *Journal of Political Economy*, 99, 145–76.

Topel, Robert H and Michael P Ward (1992), "Job mobility and the careers of young men." *The Quarterly Journal of Economics*, 107, 439–479.

Appendix

Appendix 1: NLSY79

For cognitive ability measures, I use age-adjusted Armed Forces Qualification Test (AFQT) and Armed Services Vocational Aptitude Battery (ASVAB) scores.[36] The AFQT score is a weighted sum of four section scores (1981), which are Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension and Numerical Operations with a half weight on the last one. I follow Lange (2007)'s construction of age-adjusted AFQT score. The ASVAB Math and Verbal scores are constructed from the original answer sheets by the NLS in 2010. These scores are maximum likelihood estimates of true abilities by the Item-Response Theory (IRT). The ASVAB scores are standardized within 4-month age intervals by the NLS, and I re-standardize each score to have mean 0 and s.d. 1 in each birth year group.

For non-cognitive ability measures, I use the Rotter's Locus of Control Scale and the Rosenberg Self-Esteem Scale in 1980. The Rotter's Locus of Control scale measures the extent that individuals attributes events in their life to uncontrollable factors. A low score (a negative z-score) indicates more internal, and a high score (a positive z-score) signifies more external locus of control. The Rosenberg Self-Esteem scale is a measure of self-worth based on a series of questions on how one feels about the self. Each item has four alternatives—from strongly agree to strongly disagree. I re-standardize each score to have mean 0 and s.d. 1 in each birth year group.

For education and work histories, I mostly follow the conventions developed in the previous literature.(e.g., Altonji and Pierret, 2001; Lange, 2007; Arcidiacono et al., 2010). Some measures (education groups, employment, job change) are redefined for the purpose of this paper.

[36] The ASVAB was administered to 11,914 NLS respondents (94 percent) during July through October of 1980 for the purpose of establishing a new national norm of the test (NLSY Attachment 106). ASVAB score are used to determine eligibility and assignment qualifications for specific military jobs for new enlistees, and the AFQT score, the sum of four subsection scores (word knowledge, paragraph comprehension, arithmetic reasoning and numeric operations), is a general measure of trainability and a primary criteria of enlisted eligibility for the Armed Forces (NLSY Attachment 106).

First, I use three schooling measures: *hgc0, hgc and hgcf.* *hgc* is the highest grade completed on the interview date, and *hgc0* is the highest grade completed on May 1 in the first graduation year, and *hgcf* is the highest grade completed ever throughout the surveys. There are very few cases that breaks the relationship $hgc0 \leq hgc \leq hgcf$, and $hgcf$ is on average 0.65 year higher than $hgc0$—almost 30 percent in the sample enrolled in college again after they had once stopped enrollment because of completion or graduation. Since extra college education choice can bring biases into estimation results, I use only a sample of high school graduates ($hgc0 = hgcf$) who are without a G.E.D. or any post-secondary education in the estimation of the model.

Second, the first entry into the labor market is defined by their first graduation year in the surveys, which is the earliest year that the respondent answers as a graduation year. The first graduation year is equal to the year of high school graduation for high school graduates ($hgc0 = hgcf$). The first graduation year is almost always the same with the last year before enrollment status (May 1) changes to completion/graduation for the first time in the surveys—the latter definition is not applicable to some respondents who had already graduated by their first interview in 1979.

Third, potential experience is simply the number of years after the first entry into the labor market. This is roughly equal to age minus 18 for high school graduates ($hgc0 = hgcf$).

Fourth, I focus on the main job which is the current or most recent job (job 1/CPS job) to determine one's employment status. A person is employed if and only if he answers as working at his main job on the interview date. Since multiple jobs are not uncommon, nonemployment by this definition may not correspond to actual unemployment or non-participation.

Fifth, a job change is also defined by the main job. A job is equivalent to an employer in the NLSY data, and I empirically define a job change (=employer change) as working on the interview date with less than 52 weeks of job tenure—that is, the starting point of the current job spell is within 52 weeks from the interview date.

Sixth, I use weekly earnings from the main job as the measure of labor earnings. It is constructed from the interview responses about the rate of pay and the time unit. Then the number is converted to the equivalent value in 2010 dollars using the annual average Consumer Price Index (CPI). Wages and earnings are used inter-changeably throughout this paper, and they always mean weekly earnings from the main job.

Appendix 2: Solution Concepts

Although all state variables are obviously relevant for work decisions, I focus on the unobserved individual beliefs $(\theta_t, \epsilon_{jt})$. Other variables such as the uncertainty of the beliefs (Σ_t), experience, job tenure, and assets are fixed. I here consider the case of $\gamma = 1$ (log utility) without savings. This section does not provide analytical solutions; it only provides at best a reasonable sketch of the solutions, which are numerical approximated.

First of all, job change choice can be characterized by a cutoff strategy in perceived job match quality $(\epsilon_{jt} < \epsilon^*_{jt})$. The cutoff value (ϵ^*_{jt}) exists within a reasonable range of parameters. The choice-specific value functions for staying and moving $(V^{Old}(\theta_t, \epsilon_{jt}; ...)$, $V^{New}(\theta_t, 0; ...))$ are additively separable over perceived ability and job match quality. After crossing out a common component, only one of the value functions changes—monotonically increasing in perceived job match quality. The cutoff value exists unless parameter values are extreme. This feature of a cutoff strategy is shared across many other optimal stopping problems (e.g., Jovanovic, 1979; Neal, 1999).

Then, employment status choice is also characterized by a cutoff strategy using both perceived values of ability and job match quality. This strategy depends on the previously-found cutoff value of job change. If the current belief on job match quality is below the cutoff level $(\epsilon_{jt} < \epsilon^*_{jt})$, there exists a cutoff value $(\theta_t < \theta^*_t(\epsilon^*_{jt}))$ below which the person does not take up a new offer, regardless of the perceived job match quality. If the current perceived job match quality exceeds the cutoff level $(\epsilon_{jt} > \epsilon^*_{jt})$, there exist cutoff values as a function of perceived job match quality $(\theta_t < \theta^*_t(\epsilon_{jt}))$, which must be decreasing in ϵ_{jt}. Interestingly, this is conceptually similar to Neal (1999)'s solution to a two-stage search problem although the problems are totally different.

Finally, job search under various employment-related shocks is also characterized by a cutoff strategy. The basic principle is that those who want to work at a new job choose to search, whether they are employed or unemployed. The two dimensional cutoffs for job search ($\epsilon_{jt} < \epsilon^{+}_{jt}(\theta_t), \theta_t < \theta^{+}_{t}(\epsilon_{jt})$) are basically similar to the aforementioned cutoffs for job-movers ($\epsilon_{jt} < \epsilon^{*}_{jt}, \theta_t < \theta^{*}_{t}(\epsilon^{*}_{jt})$). People whose value is the highest when they are not working would never search unless there are preference shocks—put differently, the positive selection into employment is reinforced by the existence of job search choice. Also, some people whose value is the highest when they stay at their current job might search if their perceived job match quality is only marginally higher than the job change cutoff (ϵ^{*}_{jt}) and their perceived ability is high—this is an insurance motive against job destruction. This incentive is small, however, and arises only because of the assumption on choice-shock sequence, so it is not very interesting. A more interesting feature of these search cutoffs is how they move with the presence of search costs and employment-related shocks. With search costs, some people who want a new job give up job search if their perceived ability is only marginally higher than the work cutoff $\theta^{*}_{t}(\epsilon_{jt})$ or if their perceived job match quality is only marginally higher than the job change cutoff. In general, it is likely that the positive selection into employment would get stronger with more costly job search.

Solving the finite horizon Bellman equations in the previous subsection is done by backward recursion. The expected values over the predicted distribution of the next period's beliefs (or evaluations) are calculated by simple averages over equally-probable support points. This method offers the best approximation given a fixed number of grid points (Kennan, 2006).

Appendix 3: A Note on Deterministic Covariance Matrices (Σ_t)

The beliefs system follows a first-order Markov process, which makes it easy to solve the aforementioned Bellman equations. The size of state space in that case grows linearly not exponentially, so computational burden should not be heavy, in principle.

The current version, however, does not fully take such advantage of this Markov structure. It is because of the covariance matrices (Σ_t). It is acceptable to (coarsely) discretize posterior mean space (Θ_t) and find the solutions given exact Σ_t; however, doing the same thing given (coarsely) discretized Σ_t can bring non-negligible approximation biases into the solutions.

Because of this possibility, I calculate exact Σ_t in this version, using full job change history (but not signal history) of each individual. Given each experience level E_t, not only the number of job changes but also the timing of each job change matter. If the weights were time-invariant, only the number of job changes would matter. This is equivalent to think about all job change possibilities for each experience level, and the size of state space grows exponentially in time in the labor market ($= K_t \times \Sigma_{E_t=1}^{t} 2^{E_t}$, K_t is fixed or only linearly growing in time) although the speed is not too fast during the first 15 periods.

To keep the state space manageable in this case, I additionally assume that the worker receives an exact signal on their ability after spending $T_0 (= 15)$ years in the labor market. Hence, the worker essentially switches from multidimensional learning to single-dimensional learning about job match quality after T_0 years of labor market experience. The posterior beliefs after this special signal is obtained by finding the limit of the beliefs as the noise in the signal approaches zero. That number is an arbitrary choice, but the number will not matter if the speed of learning is quite fast. I have tested other numbers up to 22, but found no evidence of significant changes in estimates.