

Experience and worker flows

ASPEN GORRY

Department of Economics and Finance, Utah State University

This paper studies the role of worker learning in a labor market where workers have incomplete information about the quality of their employment match. The amount of information about the quality of a new match depends on a worker's past job experience. Allowing workers to learn from experience generates a decline in job finding probabilities with age that is consistent with patterns found in the data. Moreover, workers with more past experience will on average have less wage volatility on new jobs, which is also consistent with the data. In contrast to the fact that the cross-sectional wage distribution fans out with experience, this second result implies that individual wage changes become more predictable.

KEYWORDS. Learning, experience, wage volatility, worker flows, job finding probability.

JEL CLASSIFICATION. E24, J24, J31, J64.

1. INTRODUCTION

This paper modifies the classic [Jovanovic \(1984\)](#) random matching model with worker learning to allow past work experience to give unemployed workers more information about the quality of a new match before choosing to accept or reject a job. Specifically, when an unemployed worker matches with a firm, the amount of information about the match depends on her past work experience. Thus, experience allows workers to better distinguish between good and bad job offers. This simple mechanism generates new and empirically relevant predictions about job finding probabilities over the life cycle while maintaining the rich set of implications from previous models of worker learning. Moreover, the model predicts that wage volatility on new jobs decreases with past work experience when job finding probabilities decline. This prediction is then found to be consistent with life cycle wage patterns in data from the National Longitudinal Survey of Youth 1979 (NLSY79). Without worker learning from past experience, the model cannot account for declines in job finding probabilities or wage volatility over the life cycle.

In contrast to the literature that emphasizes employer learning, this paper focuses on worker learning by studying the role of experience in providing information about

Aspen Gorry: aspen.gorry@gmail.com

I thank Fernando Alvarez, Frank Caliendo, Carlos Dobkin, Rick Evans, Devon Gorry, Florian Hoffmann, David Lagakos, Giuseppe Moscarini, Derek Neal, Ezra Oberfield, Ryan Oprea, Robert Shimer, Ed Schlee, Nancy Stokey, Carl Walsh, Andy Zuppann, numerous seminar participants, the editor, and the anonymous referees for helpful comments. I also thank Justin Gardiner and Regan Kuchan for their excellent research assistance. All mistakes are my own. This paper was previously distributed as “Labor market experience and worker flows.”

the quality of new job matches. Existing models with worker learning emphasize the role of workers sorting themselves into better jobs and generate accurate predictions about unemployment, job separation probabilities, and wages over the life cycle. Allowing experience to provide more information about the quality of a new match also generates declining job finding probabilities with experience and implies that individual wage volatility on a new job decreases with experience. While models of employer learning can explain cross-sectional inequality over the life cycle, the novel prediction from the new model with worker learning implies that wage volatility for a worker on a new job decreases with experience. This reduction in wage volatilities has implications for the income risk that individuals face despite growing cross-sectional inequality.¹

Implications of the model are assessed numerically. First, the model is parameterized to match estimates of the average duration of first jobs, the fraction of new jobs that involve job-to-job transitions, the age profile of job finding probabilities, and the magnitude of individual wage changes. This calibration implies that the precision of the signal that the worker receives from past experience is nearly 10% as strong as if she had been working at that job for her entire career. That is, for a worker with 10 years of past job experience she would receive nearly a year's worth of information about the quality of her match about a prospective new job. Next, the implications of the novel feature of the model are assessed. The model with experience generates declining job finding probabilities with experience that are consistent with the decline in U.S. data. The model can even generate empirically relevant declines in job finding probabilities even when experience only provides small amounts of information about the quality of a new match. Moreover, the model predicts that when job finding probabilities are decreasing, the volatility of wages that a worker faces on a new job will also decline with experience. Wage volatility is measured as the unanticipated component of wage changes during the first year on a job. When workers have more information about a new match before accepting the job, the unanticipated wage changes are smaller because workers are more certain of the quality of the match before accepting a job.

To provide evidence of the importance of the experience mechanism, predictions from the model about the volatility of wages with experience are shown to be consistent with NLSY79 data. A fixed effects specification confirms individual level reductions in wage volatility with work experience both in data simulated from the model and in NLSY79 data. Labor market experience is associated with a large decrease in within-job wage volatility that can be an important source of income risk facing workers. It is interesting to contrast this finding with the well known fact that the variance of wage distributions increases with worker experience.² The new finding that more past experience reduces the volatility of wages on individual jobs implies that while unexplained fac-

¹Most recently [Kahn and Lange \(2013\)](#) study how much of the unexplained increase in the variance of residual wages with age can be accounted for by employer learning. Other studies on the role of employer learning include [Lange \(2007\)](#), who studies the speed of learning, and [Kahn \(2013\)](#), who studies the degree of asymmetry in information between different employers.

²The empirical regularity that the cross-sectional variation in wages increases with age or experience holds both for raw data and when measured as the residual of estimated wages. For a summary of the literature, see [Neal and Rosen \(2000\)](#). [Rubinstein and Weiss \(2006\)](#) summarize the known facts about life cycle wage variation. [Kahn and Lange \(2013\)](#) estimate the role of employer learning and the growth in worker human capital as potential explanations of this phenomenon.

tors increase the dispersion of wages across the population, individuals have increasing certainty about how their own wages will evolve while working on a particular job.

An empirical literature that examines the role of experience in the evolution of wages relates to the findings in this paper. [Altonji and Shakotko \(1987\)](#), [Topel \(1991\)](#), and [Altonji and Williams \(2005\)](#) examine the extent to which wages rise with tenure in a given job rather than through experience. Learning in this paper has both a firm specific and a general effect. While experience provides individuals with information about the quality of future matches, workers learn about the quality of their current job at a faster rate. A recent line of research studies the degree to which skills are transferrable across jobs. While [Kambourov and Manovskii \(2009\)](#) found that human capital is occupation specific, recent work by [Poletaev and Robinson \(2008\)](#) and [Gathmann and Schönberg \(2010\)](#) implies that human capital formation is task based. [McCall \(1990\)](#) explores whether human capital is job or occupation specific in a search model. While this paper does not study occupational choice, it can explain [McCall's \(1990\)](#) finding that longer tenure in the first job implies lower hazard rates in future employment as experience allows workers to reject poor second matches.

This paper also contributes to a literature that uses learning models to explain labor market outcomes. [Jovanovic \(1979\)](#) explains broad features of worker turnover behavior: hump-shaped hazards of separation from a job by tenure and declining separation probabilities with age. Recent models use learning to understand the persistence of unemployment rates, wage dispersion, wage growth, and occupational mobility (see, for example, [Pries \(2004\)](#), [Moscarini \(2005\)](#), [Farber and Gibbons \(1996\)](#), [Gibbons, Katz, Lemieux, and Parent \(2005\)](#), [Gorry, Devon, and Trachter \(2014\)](#), and [Papageorgiou \(2014\)](#)). While previous learning models explain worker behavior when employed, they have little to say about worker behavior while unemployed. The major exception to this gap in the literature is the work by [Gonzalez and Shi \(2010\)](#). In their model, workers learn from search outcomes while unemployed rather than by observing outcomes during employment. While their model predicts lower wages upon employment for workers with longer unemployment durations, they do not generate different durations of unemployment by age.

By assessing a mechanism that explains why employment outcomes vary by age, this paper contributes to a growing literature that uses life cycle models to understand employment outcomes including [Cheron, Hairault, and Langot \(2011\)](#), [Esteban-Pretel and Fujimoto \(2014\)](#), [Menzio, Telyukova, and Visschers \(2012\)](#), [Yedid-Levi, Jaimovic, Siu, and Gervais \(2015\)](#), and [Gorry \(2013\)](#). These papers seek to explain the patterns of labor market outcomes over the life cycle and evaluate the effect of labor market policies such as unemployment insurance and minimum wages. Most closely related to this paper, [Esteban-Pretel and Fujimoto \(2014\)](#) introduce age dependent information revelation in a general equilibrium life cycle model to generate a decline in job finding probabilities. The learning mechanism in this paper where experience provides new information adds an additional component to life cycle human capital investment. Such a motive enhances the simple information mechanism because workers have a greater incentive to become and remain employed early in their careers. This paper also contributes to the literature by providing empirical evidence for this mechanism.

Finally, this paper contributes to a literature that seeks to explain the decline in employment transitions as workers age. In matching models like [Jovanovic \(1979\)](#) all information is job specific and unemployment durations do not change over the life cycle. More generally, models with job specific human capital and firm productivity shocks successfully explain a decline in job separation probabilities, but job specific human capital has no implication for job finding probabilities. In contrast, sorting models such as [Gibbons et al. \(2005\)](#) take learning to be about a worker's ability. In these models, a worker's performance on one job generates an equivalent amount of knowledge about her performance on all other jobs. These models also have no prediction about job finding probabilities as workers direct their search to their most profitable job. Other common explanations for the decline in turnover focus specifically on job separations. For instance, [Neal \(1999\)](#) presents a model in which workers search for both a career and a job specific match. The empirical implications of career and job matches for job turnover are explored in [Pavan \(2010\)](#). The life cycle patterns of job finding and separation probabilities must be explored together, because job finding probabilities can have a direct effect on equilibrium job separation probabilities that depend on the worker's value of unemployment.

The paper proceeds as follows. Section 2 presents the model. Section 3 describes how the parameters of the model are chosen. Section 4 presents the results from the calibrated model concerning job finding and separation probabilities, unemployment, and wage growth. Section 5 shows that the model's predictions about wage volatility are consistent with data from the NLSY79. Section 6 concludes. Technical data are available in a supplementary file on the journal website, http://qeconomics.org/supp/363/code_and_data.zip.

2. A LEARNING MODEL WITH EXPERIENCE AND JOB-TO-JOB TRANSITIONS

This section describes the economic environment of an individual making optimal decisions when faced with uncertain job opportunities. The worker has incomplete information about the quality of any particular job match.³ When engaged in production she learns about the quality of her match, and when searching for a new job she is randomly matched to employment opportunities and gets an initial signal about the quality of any prospective match. The novel feature of the model is that the precision of the signal about prospective matches depends on her past job experience.

2.1 Preferences and production

The infinitely lived worker has preferences given by

$$\sum_{t=0}^{\infty} \beta^t c_t.$$

³While the focus of this paper is on the worker's decisions as a result of learning, in the setup of the problem workers and firms have common knowledge about the quality of the match. For simplicity the exposition will focus on the worker's decision about which jobs to accept following [Jovanovic \(1984\)](#). Moreover, this paper does not take a stance on whether information from experience is generated by the worker or the firm as long as it is common knowledge.

There is no storage technology. The worker makes two decisions. When employed, she decides between quitting to search for a new opportunity and continuing to produce. When unemployed, she chooses to accept or reject prospective matches.

Production occurs when a worker is matched with a job. In each period, a match of type μ produces output

$$x_t = \mu + z_t,$$

where $z_t \sim N(0, \sigma^2)$ is independently and identically distributed noise on the output process. Therefore, $x_t \sim N(\mu, \sigma^2)$.

As in Moscarini (2005), the economy is composed of two types of jobs: $\mu \in \{\mu_h, \mu_l\}$. Let $\mu_h > \mu_l$ so that μ_h denotes the productivity of a good opportunity and μ_l denotes the productivity of a bad one. All jobs are drawn independently from the same distribution where a fraction p_0 of them are of type μ_h .

2.2 Learning

The worker is uncertain about the quality of her match and she learns about it in two ways. First, while employed, she observes her output and updates her beliefs about the quality of the match. Second, before accepting a job offer, she receives a signal about the quality of a prospective match that depends on her experience.

While employed, workers observe the output they produce in each period and update their beliefs. Given the normality of output noise, for any current belief, p , the expected density of output is given by

$$\psi(x|p) = p \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \mu_h}{\sigma}\right)^2\right) + (1 - p) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x - \mu_l}{\sigma}\right)^2\right).$$

This is the density of a mixture distribution where with probability p output is drawn from a normal distribution with mean μ_h and variance σ , and with probability $1 - p$ it is drawn from a normal with mean μ_l and the same variance.

Using this known distribution of output, the worker observes her output and uses it to update her belief about the probability that she has a good match. Given any current belief, p , and observed output for a given period, x , the updated belief, p' , is formed using Bayes' rule as follows:

$$\begin{aligned} f(p, x) &\equiv p' \\ &= \text{Prob}(\mu = \mu_h | p, x) \\ &= \frac{p \exp\left(-\frac{1}{2}\left(\frac{x - \mu_h}{\sigma}\right)^2\right)}{p \exp\left(-\frac{1}{2}\left(\frac{x - \mu_h}{\sigma}\right)^2\right) + (1 - p) \exp\left(-\frac{1}{2}\left(\frac{x - \mu_l}{\sigma}\right)^2\right)}. \end{aligned}$$

In the above expression, the numerator is proportional to the joint probability of observing output x and the match being good, while the denominator is proportional to the total probability of observing output x .

With this updating function, define the inverse function $f^{-1}(p'|p)$ to be the output, x , required to have posterior p' given prior p . This function is given by

$$f^{-1}(p'|p) = \frac{\sigma^2}{\mu_h - \mu_l} \ln\left(\frac{p'(1-p)}{(1-p')p}\right) + \frac{\mu_h + \mu_l}{2}.$$

Finally, the distribution $G(p'|p)$ can be defined as the distribution of updated beliefs after observing one period of output given a current belief p . Then the probability density function (p.d.f.) of the G distribution, g , is given by

$$\begin{aligned} g(p'|p) &= \psi(f^{-1}(p'|p)|p) \left| \frac{df^{-1}(p'|p)}{dp'} \right| \\ &= \psi(f^{-1}(p'|p)|p) \left(\frac{\sigma^2}{p'(1-p')(\mu_h - \mu_l)} \right). \end{aligned}$$

When meeting a new match the worker gets a signal about its quality that depends on her past experience. It is assumed that she receives a signal equivalent to observing $\alpha\tau$ observations from the output process. The variable τ is months of past work experience and $\alpha > 0$ determines how much information a unit of experience provides about the quality of a new match relative to month of tenure on a given job. The worker's initial belief about the quality of a match when she has no previous experience is p_0 . The normality assumption makes noninteger amounts of information well defined. Moreover, normality implies that to update beliefs after viewing $n \in \mathbb{R}$ observations the worker only needs to know her prior belief, p , the average value of the observations, \bar{x} , and the number of observations observed, t . The entire list of observations x_1, x_2, \dots, x_t is not needed. For a worker who observes n periods of output, the distribution of the average output per period, \bar{x} , is given by

$$\begin{aligned} \tilde{\psi}(\bar{x}; p, n) &= p \frac{1}{\frac{\sigma}{\sqrt{n}} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\bar{x} - \mu_h}{\frac{\sigma}{\sqrt{n}}}\right)^2\right) \\ &\quad + (1-p) \frac{1}{\frac{\sigma}{\sqrt{n}} \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\bar{x} - \mu_l}{\frac{\sigma}{\sqrt{n}}}\right)^2\right). \end{aligned}$$

This is again the mixture distribution of drawing output from good and bad jobs for n periods weighted by the current belief, p .

Using the same updating strategy, the posterior after observing the output from n periods is computed as

$$\tilde{f}(p, \bar{x}, n) = \frac{p \exp\left(-\frac{1}{2} \left(\frac{\bar{x} - \mu_h}{\frac{\sigma}{\sqrt{n}}}\right)^2\right)}{p \exp\left(-\frac{1}{2} \left(\frac{\bar{x} - \mu_h}{\frac{\sigma}{\sqrt{n}}}\right)^2\right) + (1-p) \exp\left(-\frac{1}{2} \left(\frac{\bar{x} - \mu_l}{\frac{\sigma}{\sqrt{n}}}\right)^2\right)}.$$

Inverting \tilde{f} gives the value of \bar{x} needed to generate posterior p' : $\tilde{f}^{-1}(p', p, t) = \bar{x}$. Define $H(p'|\tau)$ as the distribution of initial beliefs from a new production opportunity. Hence the p.d.f. of the H distribution, h , is given by

$$h(p'|\tau) = \tilde{\psi}(\tilde{f}^{-1}(p', p_0, \alpha\tau); p_0, \alpha\tau) \left(\frac{\sigma^2}{p'(1-p')(\mu_h - \mu_l)} \right),$$

where α and p_0 are parameters. Note that for any new match the prior belief is the unconditional probability that the match is good, p_0 , and the worker updates her belief with $\alpha\tau$ observations from the distribution of output.⁴

2.3 Wages

As in Jovanovic (1979, 1984), when firms have zero cost of entry, any wage process that pays the worker her expected output is an equilibrium. For this to be the case, the assumption that firms and workers are randomly matched is important so that workers of different levels of experience are matched with the same jobs. In such a pooling equilibrium, firms and workers are randomly matched together and then employers separate workers using available information to set wages in each period. This equilibrium does not uniquely determine the wage process as long as workers get all of the rents from the match. In this paper, it is assumed that the worker is paid her expected marginal product on the job. In particular, it is assumed that the worker's wage is given by

$$w(p) = p\mu_h + (1-p)\mu_l.$$

Under this assumption, firms make zero expected profits and any matching rate, λ , between workers and production opportunities can be sustained as an equilibrium outcome.⁵ While the wage process is not unique, the worker's matching processes with firms is.⁶

⁴The process of on the job learning can be generalized beyond the specified output process to be arbitrary distributions G and H . The distribution G must depend on the value of the current belief, p , so the distribution of updated beliefs, p' , is given by $G(p'|p)$. For a general learning process, two restrictions are made on G . First, G is nondegenerate so that the signal conveys some information about p . Second, G is restricted so that p is a martingale. This is a natural restriction since G is used to update an individual's current beliefs.

The distribution $H(p'|\tau)$ can be generalized beyond the specific normality assumptions described above. In general, for H to provide more information about jobs it must be weakly increasing in τ in terms of second order stochastic dominance. This means that for $\tau_1 > \tau_2$,

$$\int_0^x H(p'|\tau_1) - H(p'|\tau_2) dp' \geq 0 \quad \forall x \in [0, 1].$$

For higher values of τ workers get more initial information about the quality of a job. This increasing information for experienced unemployed workers is the novel feature of the model. A sufficient condition for second order stochastic dominance is that if $\tau_1 > \tau_2$, then $H(p'|\tau_1)$ is a mean preserving spread of $H(p'|\tau_2)$.

⁵One potential issue with this setup is that such a wage process is not self-enforcing as the firm will want to lower the worker's wage after a match. As discussed in Jovanovic (1984), this problem can be overcome either through accounting for the firm's reputation or assuming that there are enforceable contracts.

⁶While this simple equilibrium is used to analyze the model, it is also possible to enrich the model into a standard search and matching framework by allowing firms to post vacancies. In that case, the model

2.4 Value functions

This section defines the value functions for the worker's problem. When employed, the worker consumes her wage, $c_t = w(p)$, that depends on the expected probability that her job is good. Workers separate from their current job for three reasons. First, in each period, with probability $s > 0$ employed workers are exogenously separated from their jobs. The variable s captures reasons for job separations that are not captured by the endogenous quits that arise from learning. Possible examples include plant closures or geographic relocation by the worker. Second, they could receive an unfavorable signal about the job quality and decide to quit. Finally, when remaining employed they receive an outside job offer with probability δ . If their outside offer is better than their current job, they will quit and transition directly to the new job.

In the discrete time model, the timing of events must be specified, but alternate timing assumptions have little effect on the overall results. The following timing of events for employed workers is used. First, the worker finds out if she receives an exogenous separation shock. If still employed, she updates her belief about the quality of her current match based on output produced in the current period using $G(dp'|p)$ and decides if she will quit to unemployment or remain employed. If she remains employed, she receives an outside job offer with probability δ . The outside offer is of type μ_h with probability p_0 . Based on her current experience τ , the worker updates her belief about the quality of the outside offer. Her posterior belief is given by $H(q|\tau)$ and she changes jobs if her belief about the quality of the new job is higher than her updated belief about the quality of her current job.

Let $V(p, \tau)$ be the value function for an employed worker with belief p and experience τ . The value function is written as

$$\begin{aligned} V(p, \tau) = & w(p) + \beta s U(\tau + 1) \\ & + \beta(1 - s) \int_0^1 \max \left\{ U(\tau + 1), (1 - \delta)V(p', \tau + 1) \right. \\ & \left. + \delta \int_0^1 \max \{ V(p', \tau + 1), V(q, \tau + 1) \} H(dq|\tau) \right\} G(dp'|p). \end{aligned} \quad (1)$$

A worker with belief p and experience τ gets her expected output, $w(p)$. She discounts the future at rate β . In the next period, she is separated from her job with probability s , becoming unemployed with experience $\tau + 1$. With probability $1 - s$ she is not separated from her job and receives her updated belief from the distribution G . Depending on the realization of her updated belief she can choose to remain employed

would include a parameter for the flow cost of posting a vacancy that could be calibrated to generate a given match rate λ . The key assumption for such an equilibrium is that matching between firms and workers is random so that workers of different experience levels cannot be separated ex ante. This means that the job search process is a pooling equilibrium where all workers are competing for the same jobs. After matching, firms pay wages based on the firm's and worker's shared beliefs about the worker's expected output. This ex post separation is necessary for the model to generate predictions about how wages vary with worker experience. Alternately separating equilibria, where directed search allows firms to separate workers before matching, are possible and would likely have different implications. However, such analysis is beyond the scope of this paper.

with belief p' and experience $\tau + 1$ or quit to become unemployed with experience $\tau + 1$. If she remains employed she gets a job offer with probability δ and decides to stay or transfer depending on the quality of the new offer. With on the job search, higher values of δ give the worker a greater chance of meeting a new firm while employed, and hence lower the worker's propensity to separate from her job to unemployment. Moreover, because workers get all of the surplus in this environment, no special assumptions need to be made about the wage bargain at the point that the worker gets an outside offer. She simply chooses the job that she believes is the best and is paid her expected output. The model with job-to-job search is parameterized along with the baseline model in the next section.

Unemployed workers consume unemployment benefits, $c_t = b$. When unemployed, the worker matches with a firm with probability λ . The job is of type μ_h with probability p_0 and the worker updates her belief about the quality of the match based on her experience τ before choosing to accept the offer. The posterior distribution of beliefs about the quality of the match is given by $H(p'|\tau)$. She must choose between remaining unemployed and becoming employed with belief p' . If she does not receive a job offer she remains unemployed with the same experience.

Let $U(\tau)$ be the value function for an unemployed worker with experience τ . The value function is given by

$$U(\tau) = b + \beta(1 - \lambda)U(\tau) + \beta\lambda \int_0^1 \max\{U(\tau), V(p', \tau)\}H(dp'|\tau). \quad (2)$$

Finally, it is assumed that the maximum experience that can be accumulated by a worker is T . While the infinite horizon model would in principle allow a worker to accumulated more experience, including a finite maximum experience attained simplifies the computation of the model.⁷ It can be justified on two separate grounds. First, T can be chosen to be large enough so that workers already have nearly perfect information about new production opportunities after T periods of past experience. Second, the finite nature of individual working lives means that workers only accumulate a finite amount of experience before retirement. The assumption implies that the marginal value of additional periods of experience is 0 once a worker reaches T .

2.5 Model characterization

The general learning framework described above embeds the learning models of Jovanovic (1984) and Gibbons et al. (2005) into a matching framework so the implications for worker flows and wage volatility can be explored. When $\alpha = 0$ there is no learning across jobs and the model is identical to that of Jovanovic (1984), where workers search and get an initial signal about the quality of a match. In the case of $\alpha = 1$,

⁷An alternate assumption would be to use a finite horizon model. The finite horizon model expands the state space, as the worker's age becomes an additional state variable. Solving and simulating a finite horizon version of the model with the same parameters generates results that are nearly identical beyond minor changes in the last few years before retirement. Since the novel implications of learning are concentrated at the beginning of the workers career and important factors determining labor supply around retirement are not modeled, the paper uses the more streamlined infinite horizon model.

a worker gets a signal about the quality of an initial opportunity of equal strength to her entire past job experience. This case is similar to the model of [Gibbons et al. \(2005\)](#) where workers learn about their productivity in all jobs. The case of $\alpha = 1$ is still different from [Gibbons et al. \(2005\)](#) as in their model workers direct their search to the most productive job, so experience does not interact with job finding probabilities. Allowing experience to influence worker's information about the quality of new matches generates a theory of how job finding probabilities evolve over the life cycle. Different choices of α parameterize how much information transfers from one job to another.

The solution consists of a reservation belief about the quality of the match that depends on experience, $\bar{p}(\tau)$, such that workers will accept job offers or continue working as long as $p \geq \bar{p}(\tau)$ and reject offers or quit otherwise. The reservation belief is solved for by setting the value function of a matched worker with the reservation productivity equal to the value of an unmatched worker for each level of experience; that is, $\bar{p}(\tau)$ solves

$$V(\bar{p}(\tau), \tau) = U(\tau). \quad (3)$$

To make precise predictions about how additional experience influences the job finding probability and wage volatility of a worker, it is first important to know how additional experience influences the reservation value $\bar{p}(\tau)$. Unfortunately, it is not possible to characterize how $\bar{p}(\tau)$ changes with experience in terms of model primitives. The challenge in finding such a characterization is related to the result of [Radner and Stiglitz \(1984\)](#) about the nonconcavity of the value of information. In general, small amounts of information may not be enough to change individual decisions and hence have little value. [Chade and Schlee \(2002\)](#) emphasize the open question involved in generating global concavity. In this particular environment, to see why $\bar{p}(\tau)$ might be decreasing in τ consider an example where a worker gets no extra information about the quality of jobs until she gains $\bar{\tau}$ units of experience and then once she gets $\bar{\tau}$ units of experience the quality of any new match is perfectly revealed. Then the only value of experience before $\bar{\tau}$ is that it moves her closer to the threshold. Therefore, there will be a space of experience just before $\bar{\tau}$ that the worker will be willing to accept worse and worse opportunities just to get the payoff from getting $\bar{\tau}$ units of experience. In this case, the option value of experience outweighs the current value to the worker and can generate decreasing reservation values. The general model is rich enough to allow for different relationships between experience, the reservation belief, the job finding rate, and wages. The implications of the model are explored quantitatively in the next sections. See the [Appendix](#) for a discussion of what can be shown by making assumptions about the value of information in a simple learning model with no job-to-job transitions.

While the remainder of the paper will quantitatively explore the implications of the model, we can still understand the model better by discussing a few special cases. First, consider a model where the precision of the worker's signal about the quality of a new job is independent of her experience, as in [Jovanovic \(1984\)](#). In this case, job finding probabilities will remain constant over the life cycle. However, we can also characterize how the precision of the signal influences the worker's probability of accepting any given

job. As the precision of the signal increases, she will generally become pickier about which jobs she is willing to accept. To see this consider two cases: one where she receives no information about the quality of new jobs and a second where she receives perfect information. In the no information case, jobs are pure experience goods, so she will accept any job offer from unemployment so as to learn about its quality. However, in the second case she will only accept good jobs as they are inspection goods, lowering her job finding rate. In the current paper, work experience provides workers with a more precise signal about the quality of a prospective job. This modifies the problem by adding additional option value to remaining employed, so workers would lower their reservation belief $\bar{p}(\tau)$ relative to a model where experience had no value. Therefore, for job finding probabilities to decline over the life cycle the information effect of rejecting bad jobs has to outweigh *changes* in the option value. This will often be true if the option value declines as individuals have more information.

Finally, the mechanism described above that generates a decline in job finding probabilities with age also generates a decline in wage volatility on a new job. This result occurs as long as individuals becoming pickier about new jobs corresponds with a more precise signal about the quality of their new match. When they have a more precise signal about their new job it means that additional information that they gain while working on that job will imply smaller changes in their belief about the quality of the match. Therefore, wages will change less for new information that they receive. Again, this mechanism is more formally explored for a simple model in the [Appendix](#). The remaining sections of the paper will quantitatively evaluate these mechanisms using simulations of the model.

3. PARAMETERIZATION

This section parameterizes the model to match key features of the U.S. labor market. The period length is 1 month. Eleven parameters must be chosen to compute the model: the discount factor, β , the maximum experience, T , the expected output from a good match, μ_h , the expected output from a bad match, μ_l , the probability that a match is good, p_0 , the standard deviation of output noise, σ , the flow value of unemployment, b , the exogenous separation probability, s , the job offer probability, λ , the job offer probability when employed, δ , and the proportion of experience used for new matches, α .

First, a number of parameters are set directly either as a normalization or to match values from the data. Because the model period is 1 month, β is set to 0.9966, which corresponds to an annual interest rate of 4%. The parameter T is set to 480, corresponding to a maximum level of experience of 40 years. Increasing the maximum level of experience has no effect on the results. Finally, μ_h is normalized to 1 and μ_l is normalized to 0. While this normalization does not change observed worker transitions in the model, the size of μ_h relative to μ_l will determine how much wage growth the model generates.

The remaining parameters are calibrated together to match targets about worker's early career experiences. All remaining parameters are chosen simultaneously to minimize the sum of the square percent deviation between the chosen moments and moments simulated from the model. The remainder of this section will describe the parameters of the model along with the moments that they are most closely associated with.

First, p_0 , the probability that any new job is good, is chosen to match data on the average duration of first jobs from Farber (1994). Higher values of p_0 mean that new jobs are more likely to be good and hence workers stay in them for longer on average. Using data from NLSY79, Farber finds that the average duration of jobs is 13.7 months. As this sample focuses on young workers entering the labor market and the longest that a job could last in his sample is 10 years, we target an average job duration of 13.7 months. Before computing the mean, jobs with durations of longer than 10 years are dropped from the simulation sample.

For any value of p_0 the standard deviation of output noise, σ , and the flow value of unemployment, b , are directly computed. In the model, given the normalization of μ_h and μ_l , the evolution of p is fully determined by the signal to noise ratio: $\frac{\mu_h - \mu_l}{\sigma}$. Therefore, σ determines the evolution of beliefs. Hence, we target the standard deviation of annual changes in log wages during workers' first year on their first job. In particular, using data from the NLSY79, for each job started with experience less than 1 year that lasts longer than 1 year, the difference in log wages between the initial wage and first year wage is computed. Then the standard deviation of this measure is taken across individuals. With 1668 first job observations, the standard deviation of first year log wages is 0.3555. Then σ is computed to match the standard deviation of wage changes for 100,000 workers who start new jobs with belief p_0 . Different values of σ are computed for each value of p_0 used in the calibration and b is set to be half the value of p_0 . This sets the value of unemployment to be half of wages earned on a new job since p_0 is the wage on a new job where the worker has no past job experience. In this environment, b must be high enough that if a worker knows for certain that a job is bad, it is optimal to quit, and low enough that if the worker knows that the job is good that she will work. These assumptions ensure that the worker's search problem is nontrivial. This parameter determines the relative desirability of being employed in a bad job compared to searching for a new job. Higher values of b make unemployment more attractive. The exogenous job separation probability, s , is set to match the average job separation probability of 51–55 year old workers in the data.⁸ The average job separation probability for these workers in the data is 0.015. This moment is targeted for the fraction of separations directly to unemployment in the model (not counting job-to-job transitions). The data show that both job separation and job finding probabilities decline with age. Therefore, this is a good target for exogenous separations in the model as older workers will not separate for endogenous reasons late in life when they are sure about the quality of their match.

The parameter λ is the probability that an unemployed worker is matched with a firm in each month. This value determines the maximum job finding probability for unemployed workers if they accept any possible match. In the model, workers with no experience will accept all offers as their belief about any match is p_0 , which must be above

⁸In this paper, average monthly job finding and job separation probabilities are constructed from Current Population Survey (CPS) monthly microdata between the years 1976 and 2005 using the same procedure as Shimer (2012). The job finding probability is the monthly hazard of a worker moving from unemployment to employment and the job separation probability is the hazard of an employed working moving to unemployment in a three state model where workers can also be out of the labor force. These measures correct for the possibility of multiple switches within a month. Finally, the time series of each measure is averaged to generate average job finding and job separation probabilities by age.

their reservation belief in any interesting parameterization else they would never become employed. Hence, we target the job finding probability of 18–20 year old workers of 0.403. Note that the entire profile of job finding rates is targeted in the parameterization, λ is the highest job finding rate possible in the model and hence closely corresponds to the highest job finding rates observed in the data.⁹

The parameter δ is the probability at which employed workers get an outside job offer. Higher values of δ give the worker more frequent contact with outside jobs and hence increase the likelihood of switching directly to a new job rather than having a spell of unemployment. This is true both because they directly get more outside offers and because the fact that they get offers while employed makes unemployment less attractive. We target that 58% of new jobs are found without an unemployment spell as Farber (1994) finds that this fraction of workers in the NLSY79 data switch directly to new jobs.¹⁰

Finally, α determines by how much experience increases the amount of information a worker receives about a potential new job. With $\alpha = 0$ the model generates no decline in job finding rates with age while values of $\alpha > 0$ generate different rates of decline with age. Hence, 5-year average job finding rates from the United States are used as targets so that the model matches life cycle patterns of job finding rates from the data. Seven 5-year targets are used starting with the job finding probability of 21–25 year old workers of 0.364 and ending with the probability for 51–55 year old workers of 0.270. Higher values of α mean that past experience generates more information about prospective jobs. Hence, one feature of the calibration is to provide an estimate of how much information workers are able to transfer when assessing a new prospective match relative to the amount that they learn while working on a particular job.

Table 1 summarizes the values of each of the parameters from this calibration. In particular, $\alpha = 0.098$ implies that nearly 10% of the information learned on a given job is transferred as new information about a prospective match. The next section assesses the results of the simulated model relative to models where learning across jobs and job-to-job transitions are shut down.

4. SIMULATED RESULTS

This section documents the implications of the calibrated model. When $\alpha > 0$ past experience now has implications for worker's job search behavior. Results from the calibrated model are compared to two counterfactual models where $\alpha = 0$ and $\delta = 0$. These

⁹In previous versions of the paper λ and s have been set to directly match the job finding probability of 18 year old workers and the job separation probability of 51–55 year old workers before simulating moments of the model. The results from those calibrations were similar. Results are also robust to setting λ to higher values and s to lower values than these direct targets.

¹⁰Fallick and Fleischman (2004) and Moscarini and Thomsson (2007) also show that job-to-job transitions are extremely important in the U.S. economy. The number of employed workers directly changing employers each month is at least 2.6%, which is about twice as large as the percent of employed workers who move into unemployment. Given the magnitude of these job-to-job flows, it is possible that they could alter the selection of workers into jobs that drives the implications of the model. Hence, it is important to include these transitions when assessing the implications of the model for worker flows and wage volatility.

TABLE 1. Calibrated values of the model parameters for the baseline model and the model with job-to-job transitions.

Parameter	Value	Target Moment	Simulated Moment	Target
β	0.9966	NA	NA	4% interest rate
T	480	NA	NA	Max 40 years experience
μ_h	1	NA	NA	Normalization
μ_l	0	NA	NA	Normalization
p_0	0.647	13.7	14.36	First job duration
σ	4.40	0.3555	0.3552	Standard deviation of wage changes
b	0.319	NA	NA	Half of initial wage (p_0)
s	0.014	0.0154	0.155	51–55 year old separation probability
λ	0.392	0.403	0.397	18–20 year old finding probability
δ	0.225	0.58	0.597	Fraction of job-to-job transitions
α	0.098	0.364	0.378	21–25 year old finding probability
		0.334	0.345	26–30 year old finding probability
		0.318	0.314	31–35 year old finding probability
		0.306	0.299	36–40 year old finding probability
		0.290	0.288	41–45 year old finding probability
		0.278	0.275	46–50 year old finding probability
		0.270	0.271	51–55 year old finding probability

alternate models are the cases where either experience does not influence worker's information about the quality of a new match or where there are no job-to-job transitions. Simulations help shed light on how each of these mechanisms influences worker flows over the life cycle.

To simulate the model, the value functions are computed to generate reservation probabilities for workers at each experience level. Using these decision rules, employment outcomes are simulated for individual workers. The outcomes for 20,000 simulated workers are computed from the date that workers enter the labor force, shown as age 18 in the figures. Monthly employment, job finding probabilities, job separation probabilities, wages, tenure, and total experience are recorded. Outcomes from the model can be compared to the data by assuming that there are a large number of workers facing identical decision problems. Each worker faces a different history of idiosyncratic shocks. Hence, aggregate data are constructed from the model by averaging outcomes across workers and monthly observations to generate average annual outcomes. Simulated data over a 40-year career are generated to see if they match general patterns observed in U.S. labor market data.

4.1 Job finding probabilities

First, simulated job finding probabilities by age are shown in Figure 1. The figure plots average monthly job finding probabilities for each age group targeted in the calibration (18–20, 21–25, 26–30, 31–35, 36–40, 41–45, 46–50, and 51–55) along with lines that show average job finding rates by each age from the data. The simulated lines still show some variation even with 20,000 individual simulations since only a small fraction of individuals are unemployed in any given month. The figure shows that the baseline calibration

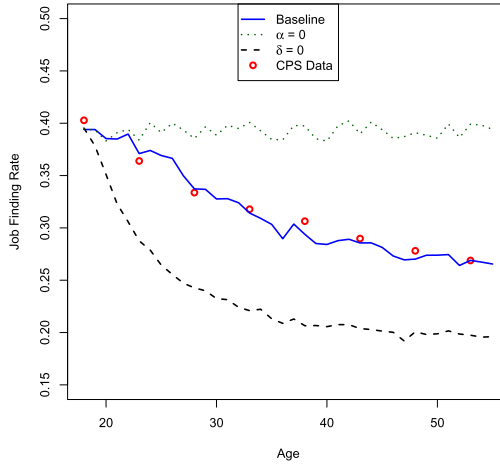


FIGURE 1. Monthly job finding probabilities by age for the U.S. economy along with simulations of the baseline calibrated model, the model without experience ($\alpha = 0$), and the model without job-to-job transitions ($\delta = 0$). Data are denoted by dots; calibrated models are denoted by lines.

closely matches the decline in job finding probabilities as workers age. In contrast, when $\alpha = 0$ the model does not generate any decline in job finding rates with age. The parameter α determines how much information past job experience provides about new job opportunities: $\alpha = 0$ is analogous to the standard Jovanovic (1979) model where individuals learn nothing about future jobs and the employment is a pure renewal process. Higher values of α imply that workers learn faster about future jobs. This can generate steeper declines in both job finding and separation probabilities as workers can choose to reject bad jobs before accepting them and reject bad jobs more quickly when employed.

Figure 1 also plots results for the model with no job-to-job transitions when $\delta = 0$. The ability of workers to transfer directly from one job to another reduces the option value of unemployment, making even bad jobs more attractive. This implies that workers are willing to accept a larger fraction of jobs, so information will have less of an effect in reducing job finding probabilities. This can be seen from the fact that the rate of decline in the job finding rate is much steeper for a given value of α when $\delta = 0$. When interpreting the results it should also be emphasized that even small values of α can generate substantial declines in job finding rates over the life cycle. While the calibrated value implies that for every 10 years of work experience nearly 1 year of information about a new job is generated, even smaller values of information can generate declines in job finding rates, modifying the implications of standard learning models.

4.2 Job separation probabilities

Next, we consider the model's predictions about job separation probabilities over the life cycle. In the data, the job separation probability has a sharp initial decline during the first three average points from age 18 to 30. To compare the model predictions to the data it should be noted that job separations can be measured in two ways from the model.

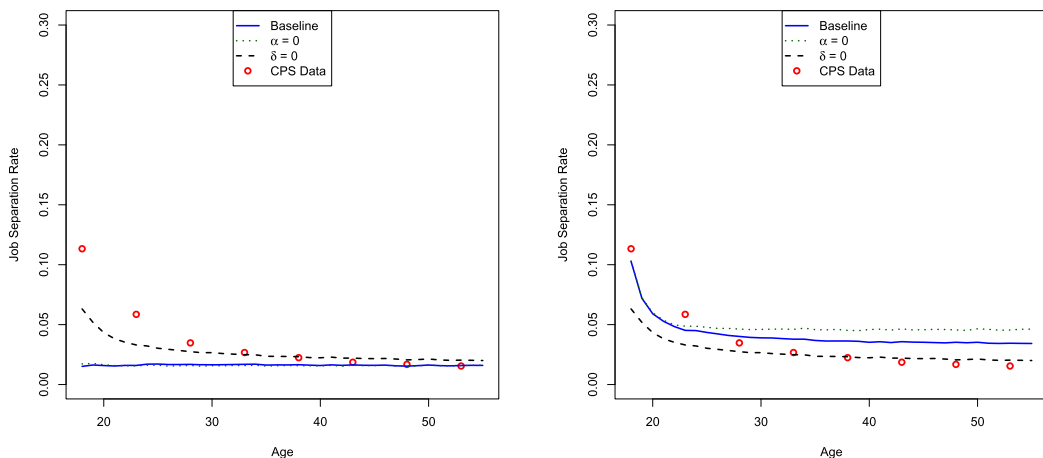


FIGURE 2. Monthly job separation probabilities by age for the U.S. economy along with simulations of the baseline calibrated model, the model without experience ($\alpha = 0$), and the model without job-to-job transitions ($\delta = 0$). Left panel: job-to-job transitions excluded. Right panel: all separations. Data are denoted by dots; calibrated models are denoted by lines.

The first measure only looks at separations to unemployment while the second measure includes both separations to unemployment and those generated by job-to-job transitions. The simulated results are presented in Figure 2. The left panel shows results from each model for separations directly to unemployment. Recall that the baseline calibration targets the model to match the data for 51–55 year old workers. This target roughly corresponds to the rate of exogenous job separations in the model. Endogenous separations arise in the model as workers quit when their belief about the quality of their current match falls below an endogenous threshold. The rate of these separations early in life will depend on the fraction of jobs that are good, the speed of learning, and the relative rate of job offers in employment and unemployment. The figure shows that the baseline model with job-to-job transitions closely matches the target, but does not generate many increased separations to unemployment early in the worker's career. This is because the job offer rate for employed workers is high enough that unemployment is no longer a very attractive option. Hence, the job separation profile is flat for both the baseline model and the model when $\alpha = 0$. In contrast, in the case where $\delta = 0$ and there are no job-to-job transitions the model generates a modest amount of endogenous separations to unemployment early in life.

The right panel of Figure 2 shows simulations for each specification of the model for all separations. The separation rate includes both separations to unemployment and job-to-job transitions. Including job-to-job transitions implies that both the baseline model and the model with $\alpha = 0$ now have declining job separation rates with age. This is because job-to-job transitions occur at a higher rate early in life as workers sort themselves into good jobs. Once a worker is fairly certain that her job is a good match she is unlikely to receive an outside offer that is better so the rate of job-to-job transitions declines with age. Note that in the baseline model the decline in job separations continues for longer as increased learning about the quality of new jobs allows new job offers

to improve over the life cycle, while when $\alpha = 0$ all new jobs have belief p_0 so the rate of separations levels off much sooner. In this case, the model better matches the CPS data on the job separation probability by age. It is worth noting that the CPS data are measured as separations to unemployment in each month, but the procedure in [Shimer \(2012\)](#) imputes the number of unemployment spells lasting less than a month. Hence, the measure could be capturing some direct job-to-job transitions depending on the definitions used.

4.3 Wages

Finally, [Figure 3](#) shows the average annual wages by age from each model. All three models generate rapid wage growth during the first 10 years of experience and then level off. The amount of wage growth generated by the models is arbitrary from the normalization of the levels of μ_h and μ_l in the parameterization. It is worth noting that α plays an important role in continued increases in wages after the first 10 years or the worker's career. In the model with $\alpha = 0$ wages level off as workers do not learn about future prospective matches, while when $\alpha > 0$ wages continue to rise throughout the life cycle as workers select into better and better matches if they ever become unemployed.

While the models match the general pattern of wage growth, they can generate an arbitrary amount of wage growth over the life cycle. These patterns are qualitatively consistent with the behavior of wages over the life cycle. [Flinn \(1986\)](#) argues that wage growth and turnover are related for young workers. The model presents a theory that accounts for both phenomena. In particular, [Topel and Ward \(1992\)](#) document a number of features of wage profiles during a worker's first 10 years of experience. They document that the first 10 years of the career account for two-thirds of lifetime wage growth. Job changes explain about one-third of wage growth. Moreover, wages on the job approximate a random walk.

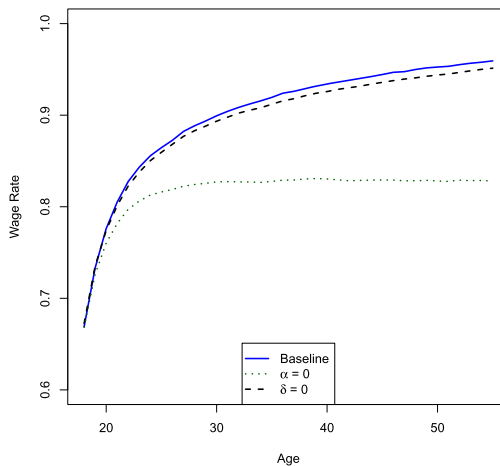


FIGURE 3. Wages by age in the baseline calibrated model, the model without experience ($\alpha = 0$), and the model without job-to-job transitions ($\delta = 0$).

5. WAGE VOLATILITY

This section quantitatively evaluates the model's predictions on wage volatility using data on wages and job changes from the NLSY79. The novel feature of the model is that workers who start jobs with more experience have better information about the quality of their new job. This information means that on average workers with more experience start with a higher p . Therefore, when starting a new job, the change in wages over the first year would display less volatility than a worker starting out with less experience.¹¹ These predictions are documented using observations simulated from the model. Moreover, as evidence that the learning mechanism proposed is important for workers, the same results are documented using data from the NLSY79. As robustness, the implications of the model with no learning across jobs ($\alpha = 0$) and no job-to-job transitions ($\delta = 0$) are also explored. The model with $\alpha = 0$ does not generate any decline in wage volatility with experience even with job-to-job transitions while the results are robust in the model with $\delta = 0$. This implies that selection into better matches generated by job-to-job transitions alone cannot generate the decline in wage volatility with experience observed in the data.

The results contrast with the well known fact that the cross-sectional variation in residual wages increases as workers age as discussed in [Kahn and Lange \(2013\)](#) among others. While pay distributions fan out with experience in the cross section, the model in this paper predicts that the unanticipated changes in individual wages within a given job shrink with experience. Therefore, individuals are more certain about their future wages as their careers progress. While the effect of experience on wages has been explored by a large theoretical literature (see [Neal and Rosen \(2000\)](#), [Gibbons et al. \(2005\)](#)), previous work has not explored the impact of experience on individuals' within job wage volatility. Understanding the features of the individual income process is important to explain a wide array of individual behavior (see [Meghir and Pistaferri \(2004\)](#)).

5.1 *Data and measurement of wage volatility*

The NLSY79 is a nationally representative longitudinal survey conducted by the Bureau of Labor Statistics that samples 12,686 individuals who were between the ages of 14 and 22 years old when first surveyed in 1979. The NLSY79 provides a rich set of panel data for tracking workers' career outcomes. Individuals were surveyed every year until 1994 and every 2 years thereafter. The sample is restricted to the 1979–1994 period for the analysis as yearly differences in wages are needed to construct measures of wage volatility within jobs.

Key variables used in the analysis are worker experience and wage volatility on new jobs. Experience is constructed by taking a cumulative sum of the weeks worked in the

¹¹As shown in the [Appendix](#), a worker with more experience and hence more information about their quality of the match will experience lower wage volatility in the case where $\bar{p}(\tau) > 0.5$ and is increasing in τ . The mechanism then is that workers who start with a higher belief p will have smaller changes in future wages on a job because p is also a sufficient statistic for the variance of the worker's belief. The same result would hold more generally for the case where the productivity of jobs were drawn from a normal distribution rather than being of just two types.

past year variable. This number is divided by 52 to create years of experience. To avoid miscalculation of past experience, the sample is limited to workers who are 17 years old or younger at the time of the first interview so that all experience after finishing high school is counted. For the initial job, the greater value of weeks worked during the past year and the tenure on that job are used. This variable provides a measure of the worker's cumulative experience and can be used as a measure of how much actual (rather than potential) experience the worker has in any given year when starting a new job.

Wage volatility is measured as the unanticipated change in workers' wages during their first year on a new job. To construct job variables the data are broken into job observations using NLSY79 data on the total number of past jobs that the respondent has held. The NLSY79 defines a job as a relationship between an individual employer and the worker so that changes in position within a firm are not considered new jobs. A new job observation is created if the total number of jobs in year t is greater than in year $t - 1$. For each job observation the wage in each year is given by the CPS wage variable.¹² Wage observations in the top and bottom 1% are dropped to correct for misreporting in the wage data and to make sure that outliers are not driving the results. The final data set includes 10,246 job observations.

Each worker's employment history is broken into jobs that are characterized by a wage for each year of tenure on the job and the initial experience level when starting the job. To do this the job number variable from the NLSY79 is used that counts only new jobs, excluding recalls to the same employer. To look at an individual's wage volatility on a particular job we focus on wage changes during the first year on that job. Therefore, jobs with fewer than two annual wage observations are dropped from the sample. Wage volatility is measured as the absolute deviation from the worker's expected wage growth path. The wage volatility measure used is

$$v = |\log(w_1) - \log(w_0) - \overline{\Delta w_e}|,$$

where v is the volatility of wages on a given job. The variable w_t is the wage observed at tenure t , so the measure computes the log difference between the wage after 1 year and the initial wage. Finally, since wages grow with experience and also grow more rapidly early in a worker's career, the expected log wage growth for the worker's level of experience when starting the job, $\overline{\Delta w_e}$, is subtracted from the observed wage change. The experience level $\overline{\Delta w_e}$ is measured as the predicted value of the log wage change estimated from a quintic polynomial of annual experience. Taking the absolute value of this difference provides the magnitude of the difference in wages from the expected wage path after 1 year on a particular job. This maps directly into the predictions of the model that higher experience should result in smaller changes in the worker's belief p and wage $w(p)$.

¹²Since wages are only recorded each year it is possible that the wages could have already changed from the initial wage at the time of first observation. Despite this measurement issue, the same issue arises when annual data are taken from the simulated model. In the simulated model workers' wages change every month based on their updated beliefs. Treating the simulated data the same as the NLSY79 observations should yield similar biases.

A nice feature of this measure is that looking at only individual wage differences abstracts from a number of issues that arise when using cross-sectional wage variation. Higher volatility implies that a given individual experiences larger changes in her wages on a given job. By subtracting Δw_e the measure of volatility used in this paper controls for expected wage gains in each year of tenure at a particular job and level of experience. While most workers get wage increases from year to year, subtracting the expected wage growth means that many workers are both above and below the expectation.

To compare observed outcomes from the NLSY79 with the model, 25 years of annual observations from the model are simulated for the worker's employment status, past experience, accumulated job number, and wage.

5.2 Results

The novel prediction of the model is that wage volatility declines with job experience. To explore this prediction, measured volatility for each job is plotted by previous experience. The left panel of Figure 4 shows the scatter plot for the calibrated model. The model generates a decline in wage volatility for workers with more past work experience. The right panel of Figure 4 shows the scatter plot of the median wage volatility by experience for the NLSY79 data. Just as in the simulation the data show a decline in wage volatility for workers with more past experience. The patterns of volatility with experience are consistent with those predicted by the model. Note that there is a difference in scale between the model and the data with a higher average level of wage volatility in the data. Since the magnitudes of wage volatility are higher in the data, the model does not capture all of the observed volatility. However, ignoring the differences in levels, the model and data have a similar rate of decline in volatility over the first 15 years of work experience.

Figure 5 presents plots of wage volatility with experience for the other two specifications of the model. The left panel shows the case where $\alpha = 0$. In this case the figure suggests that there is no decline in wage volatility with experience. The right panel plots the case of $\delta = 0$. Here the figure shows some decline, but the magnitude appears to be smaller than the baseline model.

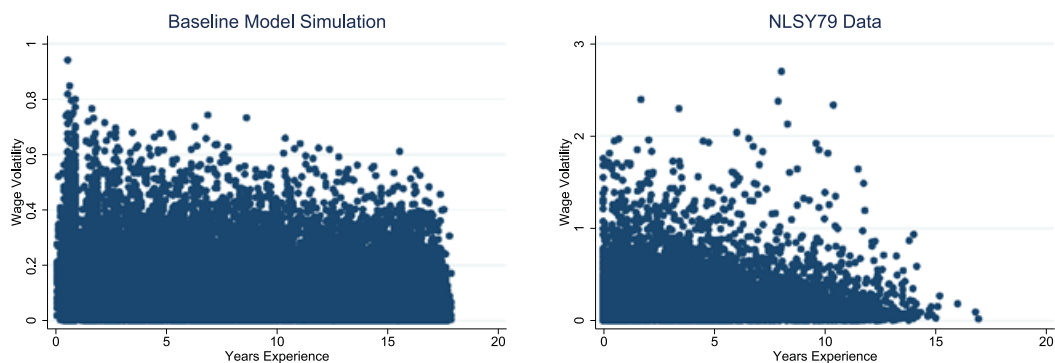


FIGURE 4. Wage volatility (v) by years of past experience. Baseline simulated model in the left panel; NLSY79 data in the right panel.

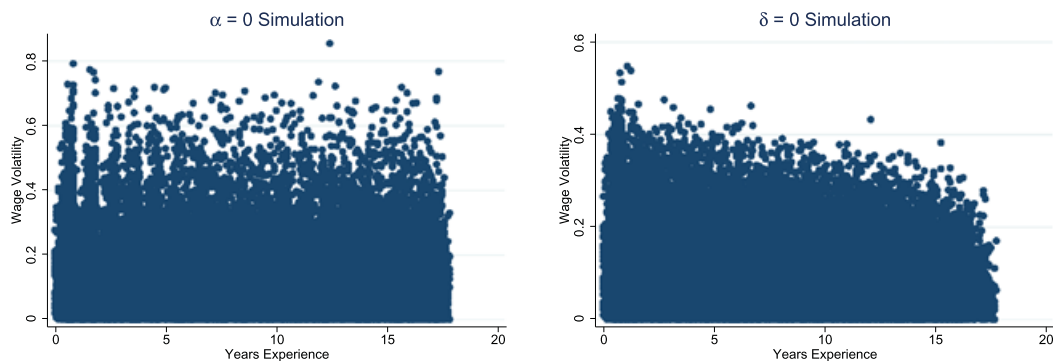


FIGURE 5. Wage volatility (v) by years of past experience: simulated model without experience ($\alpha = 0$) in the left panel and without job-to-job transitions ($\delta = 0$) in the right panel.

TABLE 2. Regression results of wage volatility on years of experience for simulated data from each of the three model specifications.

Variable	Wage Volatility					
	Baseline		$\alpha = 0$		$\delta = 0$	
Experience	-0.0037*** (0.0001)	-0.0039*** (0.0001)	0.0001 (0.0001)	-0.0001 (0.0001)	-0.0032*** (0.0001)	0.0032*** (0.0001)
Intercept	0.1143*** (0.0007)	0.1154*** (0.0006)	0.1160*** (0.0007)	0.1169*** (0.0007)	0.1126*** (0.0006)	0.1130*** (0.0005)
Fixed effects	No	Yes	No	Yes	No	Yes
Observations	66,538	66,538	62,289	62,289	61,766	61,766

Note: Standard errors clustered by individual in OLS specifications. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

Next, the model predictions about the decline in wage volatility with experience are empirically evaluated. Table 2 shows the results of regressions of wage volatility on years of experience for simulated data from each of the three models. For each model, results are shown for a standard ordinary least squares (OLS) regression and an empirical specification with fixed effects to show that the patterns hold at the individual level. The first two columns of the table report results from the baseline model. The results imply that an additional year of experience reduces wage volatility by about 37 basis points. The final four columns present the results for the model with no learning from experience ($\alpha = 0$) and no job-to-job transitions ($\delta = 0$). When $\alpha = 0$ and workers do not learn anything from past work experience, the model does not generate any reduction in wage volatility with experience. Both the OLS and fixed effects specifications generate precise zero estimates of the coefficient on experience. This is the case because for any new job that a worker starts she has the same belief about its quality, p_0 . Therefore, even with job-to-job transitions there is no reduction in wage volatility on new jobs. However, in the specification with no job-to-job transitions there is still a strongly negative effect of experience on wage volatility. The effect is slightly smaller than the baseline model, demonstrating that job-to-job transitions can amplify the effects from the model. This

TABLE 3. Regression results of wage volatility on years of experience for NLSY79 data with controls.

Variable	Wage Volatility			
Experience	-0.0048*** (0.0008)	-0.0048*** (0.0008)	-0.0068*** (0.0014)	-0.0038*** (0.0010)
Intercept	0.2189*** (0.0044)	0.2248*** (0.0121)	0.1960*** (0.0196)	0.2145*** (0.0049)
Individual controls	No	Yes	Yes	NA
Age and time dummies	No	No	Yes	NA
Individual fixed effects	No	No	No	Yes
Observations	10,246	10,246	10,246	10,246

Note: Robust standard errors clustered by individual in OLS specifications. *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

occurs for two reasons. First, when workers learn, job-to-job transitions allow workers to move from their current match to an even better match. This combination of learning is strengthened by the restriction that workers will only choose to switch if they are more certain that the new job is good. Second, as shown in the results the model with $\delta = 0$ generates an even more rapid decline in job finding rates with age because unemployment is the only route to find a new job. This reduction in job finding probabilities corresponds with increasing pickiness with regard to which jobs to accept as the individual cannot switch jobs when employed, enhancing the effect of learning.

Finally, Table 3 empirically evaluates the relationship between wage volatility and experience in the NLSY data. The first column of the table shows the raw regression of wage volatility on work experience and finds a significant negative coefficient of similar magnitude found in the model with job-to-job transitions with a 1 year increase in experience corresponding to a 48 basis point reduction in wage volatility. The next two columns add individual controls and age and time dummies to the regression. The individual controls include gender, education, and race dummies. The inclusion of additional controls does not change the main estimate. Most of the dummies are insignificant with the exception that females have significantly lower wage volatility than males. The education dummies are important as the impact of experience on individual wage volatility remains the same even though different education groups are known to have different wage experience profiles.¹³ The third column includes age and time dummies to control for the possibility that the year that individuals entered the labor market or other cyclical factors could influence the outcomes. With these additional controls the relationship between experience and wage volatility becomes stronger with an additional year of experience corresponding to a 68 basis point reduction in wage volatility. Finally, the fourth column shows the specification with fixed effects to document that the predicted pattern from the model holds within individual observations. In this specification, an additional year of experience corresponds to a 38 basis point reduction in wage volatility that is significant at the 1% level.

¹³See for instance Farber and Gibbons (1996) and Lange (2007) for a more recent discussion.

The results in this section confirm a robust negative relationship between past work experience and observed wage volatility in the first year of a new job. This relationship is surprising in light of the well known fact that cross-sectional wage variation increases with worker's experience. The cross-sectional variation of residual wages is measured across workers, which contrasts with the measure of wage volatility of workers in this paper measured as individual wage changes on a particular job. Such a cross-sectional increase in the dispersion of wages is generated by learning models because beliefs become more diffuse as workers learn about the quality of their jobs. However, the results in this paper suggest that even though differences in wages across workers grow, experience allows workers to better predict how their own wages will evolve. Hence, cross-sectional wage variation may not be a good measure of the uncertainty that individual workers face about their wages.

6. CONCLUSION

This paper presents a model of learning that can explain changes in workers' job finding probabilities over their life cycle. Workers' learning about the quality of their match is important for both observed outcomes while employed, such as wages and employment durations, and outcomes while unemployed. This insight motivates the model where experience gives workers both knowledge about the quality of their current job and the ability to distinguish between good and bad job offers.

A model with learning about both the quality of the current match and future matches has rich implications for labor market outcomes. It is consistent with the age profiles of unemployment, job finding probabilities, job separation probabilities, hazard rates of separation with tenure, wage dispersion, and wage growth. The model is used to generate new predictions about workers' wage volatility on jobs based on their experience. The prediction of lower volatility with more past experience holds in the NLSY79 data. While job-to-job transitions enhance the reduction in wage volatility generated by the model, the model cannot generate any decline in volatility without learning about the quality of new matches from experience. Finally, the fact that past experience has an impact on wage levels and volatility on new jobs has implications for how to identify returns to experience.

APPENDIX: MODEL CHARACTERIZATION

This section seeks to characterize the model by making additional assumptions about the value of information. To simplify the analysis consider a simplified model with no job-to-job transitions. Let $V(p, \tau)$ be the value function for an employed worker with belief p and experience τ . The value function is written as

$$V(p, \tau) = w(p) + \beta s U(\tau + 1) + \beta(1 - s) \int_0^1 \max\{U(\tau + 1), V(p', \tau + 1)\} G(dp' | p).$$

The unemployed value function remains the same.

So as to characterize how additional experience influences the optimal reservation value, define the notation

$$\Delta m(y) \equiv m(y+1) - m(y)$$

for any function $m(y)$. For functions of more than one variable, Δ will always denote changes in experience, τ .

By making assumptions about the value of information, the sign of $\Delta \bar{p}(\tau)$ can be determined. The reservation productivity level will be increasing in experience if the marginal value of information while employed at the reservation belief is less than the marginal value of information when unemployed, that is, when $\Delta U(\tau+1) \leq \Delta U(\tau)$. Without job-to-job transitions, extra experience only impacts a worker when she becomes unemployed and seeks a new job. Similar results can be obtained in the full model with additional work to consider changes in the value of information for job offers received while employed. This condition can be interpreted as requiring the marginal value of experience for unmatched workers to be decreasing in τ . Another way to understand the result is that the direct benefit from the additional unit of experience has to be greater than the option value of the unit of experience for getting more experience later in life. Moscarini and Smith (2002) provide evidence that this is not a very restrictive assumption by showing that the value of information is asymptotically decreasing in the number of signals.

This intuition is formalized in the following proposition.

PROPOSITION 1. *If $\Delta U(\tau+1) \leq \Delta U(\tau)$, then $\Delta \bar{p}(\tau) > 0$ for all $\tau \in \{0, 1, \dots, T\}$.*

PROOF. Differencing (3) with respect to τ gives

$$\begin{aligned} \Delta \bar{p}(\tau) V_p(\bar{p}(\tau), \tau) + \Delta V(\bar{p}(\tau), \tau) &= \Delta U(\tau), \\ \Delta \bar{p}(\tau) &= \frac{\Delta U(\tau) - \Delta V(\bar{p}(\tau), \tau)}{V_p(\bar{p}(\tau), \tau)}. \end{aligned}$$

Then $\Delta \bar{p}(\tau) > 0$ if $\Delta V(\bar{p}(\tau), \tau) \leq \Delta U(\tau)$. It suffices to show that $\Delta V(p, \tau) \leq \Delta U(\tau+1)$ for all $\tau \in \{0, 1, \dots, T\}$ and $p \in [0, 1]$. We will proceed by backward induction starting from $\tau = T$. For $\tau = T$,

$$\Delta V(p, T) = \Delta U(T+1) = \Delta U(T) = 0.$$

For $T-1$,

$$\begin{aligned} \Delta V(p, T-1) &= [\beta s + \beta(1-s)G(\bar{p}(T)|p)]\Delta U(T) \\ &\quad + \beta(1-s) \int_{\bar{p}(T)}^1 \Delta V(p', T)G(dp'|p) \\ &= 0 = \Delta U(T). \end{aligned}$$

For $T - 2$,

$$\begin{aligned}\Delta V(p, T - 2) &= [\beta s + \beta(1 - s)G(\bar{p}(T - 1)|p)]\Delta U(T - 1) \\ &\quad + \beta(1 - s) \int_{\bar{p}(T-1)}^1 \Delta V(p', T - 1)G(dp'|p) \\ &= [\beta s + \beta(1 - s)G(\bar{p}(T - 1)|p)]\Delta U(T - 1) \\ &= \beta[s + (1 - s)G(\bar{p}(T - 1)|p)]\Delta U(T - 1) < \Delta U(T - 1).\end{aligned}$$

Finally, assuming $\Delta V(p, T - n) \leq \Delta U(T - n + 1)$, we can solve for $T - n - 1$:

$$\begin{aligned}\Delta V(p, T - n - 1) &= [\beta s + \beta(1 - s)G(\bar{p}(T - n)|p)]\Delta U(T - n) \\ &\quad + \beta(1 - s) \int_{\bar{p}(T-n)}^1 \Delta V(p', T - n)G(dp'|p) \\ &\leq [\beta s + \beta(1 - s)G(\bar{p}(T)|p)]\Delta U(T - n) \\ &\quad + \beta(1 - s)(1 - G(\bar{p}(T - n)|p))\Delta U(T - n + 1) \\ &\leq [\beta s + \beta(1 - s)G(\bar{p}(T)|p)]\Delta U(T - n) \\ &\quad + \beta(1 - s)(1 - G(\bar{p}(T - n)|p))\Delta U(T - n) \\ &= \Delta U(T - n).\end{aligned}$$

The first inequality comes from the induction and the second comes from the hypothesis that $\Delta U(\tau + 1) \leq \Delta U(\tau)$ for all $\tau \in \{0, 1, \dots, T\}$. \square

Although, Proposition 1 does not characterize the sign of $\Delta \bar{p}(\tau)$ in terms of model parameters, it provides clear intuition for when the reservation belief will be increasing in experience.

Using the results on the worker's reservation decision above, it is useful to consider the behavior of the job finding rate as a function of experience, $f(\tau)$. The job finding rate is determined by the exogenous rate of matches combined with the workers willingness to accept production opportunities:

$$f(\tau) = \lambda(1 - H(\bar{p}(\tau)|\tau)).$$

PROPOSITION 2. *If $\Delta U(\tau + 1) \leq \Delta U(\tau)$ and $\bar{p}(\tau) \leq p_0$, then $\Delta f(\tau) < 0$.*

PROOF. Taking the derivative of $f(\tau)$ with respect to τ gives

$$\Delta f(\tau) = -\lambda h(\bar{p}(\tau)|\tau)\Delta \bar{p}(\tau) - \lambda \Delta H(\bar{p}(\tau)|\tau),$$

where $h(\bar{p}(\tau)|\tau)$ is the p.d.f. of H . By Proposition 1, $\Delta U(\tau + 1) \leq \Delta U(\tau)$ implies that $\Delta \bar{p}(\tau) > 0$. This condition guarantees that the first term is negative. The second term is also negative when $\bar{p}(\tau) \leq p_0$ because $H(p|\tau)$ is a mean preserving spread around p_0 . \square

The conditions in Proposition 2 generate declining job finding probabilities early in workers lives where $\bar{p}(\tau) \leq p_0$ as workers accept most jobs to gain experience.¹⁴ There are three effects that change the job finding rate in the model. First, more information shifts the distribution of possible jobs. When higher values of τ produce a mean preserving spread on $H(\bar{p}(\tau)|\tau)$, the direction of the effect depends on whether $\bar{p}(\tau)$ is greater than or less than p_0 . For $\bar{p}(\tau) < p_0$ more information shifts more of the mass of the distribution below $\bar{p}(\tau)$. Second, the worker's optimal choice shifts in response to the known distribution. Under assumptions about the value of information, more information shifts $\bar{p}(\tau)$ upward. Finally, the value of information drives a wedge between the value of employment and the value of unemployment as workers only gain experience when employed. This can increase or decrease $\bar{p}(\tau)$, depending on the marginal value of an additional unit of experience. While generating a decline in job finding probabilities does not provide a test of the model, the strength of the model is that it generates a theory about when job finding probabilities will be increasing or decreasing. Job finding probabilities are likely to decline early in workers careers when additional information has diminishing returns.

The final implications of the model involve the process of the worker's current belief p while employed. Given the binary structure of productive opportunities in the model, the density of output $\psi(x|p)$ is a mixture density of two normal distributions with means μ_h and μ_l and weights of p and $(1-p)$. The properties of the mixture density are characterized in the following proposition.

PROPOSITION 3. *The distribution characterized by the mixture density $\psi(x|p)$ has mean $\mu_{\text{mix}} = p\mu_h + (1-p)\mu_l$ and variance $\sigma_{\text{mix}}^2 = p(1-p)(\mu_h - \mu_l)^2 + \sigma^2$ where σ^2 is the variance of each of the two distributions in the mixture.*

PROOF. The density $\psi(x|p)$ is the mixture of two normal distributions with means μ_h and μ_l and the same variance σ^2 . Letting w_i denote the weights in each distribution, we have that the mean of the mixture distribution μ_{mix} is given by

$$\mu_{\text{mix}} = \sum_{i \in \{h,l\}} w_i \mu_i = p\mu_h + (1-p)\mu_l.$$

The variance σ_{mix}^2 is given by

$$\begin{aligned} \sigma_{\text{mix}}^2 &= \sum_{i \in \{h,l\}} [w_i(\mu_i^2 + \sigma^2)] - \mu_{\text{mix}}^2 \\ &= p(\mu_h^2 + \sigma^2) + (1-p)(\mu_l^2 + \sigma^2) - (p\mu_h + (1-p)\mu_l)^2 \\ &= p(1-p)\mu_h^2 + (1-p)(1-(1-p))\mu_l^2 - 2p(1-p)\mu_h\mu_l + \sigma^2 \\ &= p(1-p)(\mu_h - \mu_l)^2 + \sigma^2. \quad \square \end{aligned}$$

¹⁴The focus of this paper is to model the initial decline in job finding probabilities for young workers. In certain parameterizations, job finding probabilities can be increasing later in life when $\bar{p}(\tau) > p_0$; however for most parameter values the increase in job finding probabilities occurs after longer horizons than the typical length of a worker's career.

The standard deviation of expected output depends on both the standard deviation of the output process σ^2 and the uncertainty about the worker's job type. While the noise from the output process is constant the uncertainty depends on the current belief p . This implies that the standard deviation of expected output peaks at $p = 0.5$ and decreases as p converges to 0 or 1. While the model does not provide a closed form solution for the standard deviation of $G(p'|p)$, the above intuition shows that it is decreasing in p if $p > 0.5$.

This result is important to understand how workers learn. Unlike in Jovanovic (1979) the standard deviation of output does not decrease monotonically as the worker learns more. Instead, the standard deviation depends on the current value of p and will on average decrease for workers as their current belief converges to either 0 or 1.¹⁵

Next, the expected initial belief based on initial experience τ is characterized in the following proposition.

PROPOSITION 4. *If $\tilde{\tau} > \tau$, $\bar{p}(\tau) \leq \bar{p}(\tilde{\tau})$, and $\bar{p}(\tau) \leq p_0$, then the expected value of the initial beliefs for an accepted offer is higher for the worker with more experience. That is,*

$$E[H(p|\tau)|p \geq \bar{p}(\tau)] \leq E[H(p|\tilde{\tau})|p \geq \bar{p}(\tilde{\tau})].$$

PROOF. First, note that since $\tilde{\tau} > \tau$, $H(p|\tilde{\tau})$ is a mean preserving spread of $H(p|\tau)$. By definition of a mean preserving spread we have

$$\int_0^x H(p|\tau) dp \leq \int_0^x H(p|\tilde{\tau}) dp$$

for any value of $x \in (0, 1]$. Note that the lower bound on the integrals is 0 as the support of the distribution H is from 0 to 1. Integrating the above equation by parts on each side gives

$$xH(x|\tau) - \int_0^x ph(p|\tau) dp \leq xH(x|\tilde{\tau}) - \int_0^x ph(p|\tilde{\tau}) dp.$$

Subtracting the bound x from each side and adding the mean of the distribution H , p_0 , to each side of the equation gives

$$\begin{aligned} xH(x|\tau) - x + p_0 - \int_0^x ph(p|\tau) dp &\leq xH(x|\tilde{\tau}) - x + p_0 - \int_0^x ph(p|\tilde{\tau}) dp, \\ -x[1 - H(x|\tau)] + \int_x^1 ph(p|\tau) dp &\leq -x[1 - H(x|\tilde{\tau})] + \int_x^1 ph(p|\tilde{\tau}) dp. \end{aligned}$$

Factoring and simplifying each side gives

$$[1 - H(x|\tau)] \left[\frac{\int_x^1 ph(p|\tau) dp}{1 - H(x|\tau)} - x \right] \leq [1 - H(x|\tilde{\tau})] \left[\frac{\int_x^1 ph(p|\tilde{\tau}) dp}{1 - H(x|\tilde{\tau})} - x \right],$$

¹⁵If the job types are drawn from a normal distribution instead of one with two types, the variance of beliefs about the worker's type would decrease monotonically in information.

$$\frac{\int_x^1 ph(p|\tau) dp}{1 - H(x|\tau)} \leq \frac{1 - H(x|\tilde{\tau})}{1 - H(x|\tau)} \left[\frac{\int_x^1 ph(p|\tilde{\tau}) dp}{1 - H(x|\tilde{\tau})} - x \right] + x,$$

$$E[H(p|\tau)|p \geq x] \leq \frac{1 - H(x|\tilde{\tau})}{1 - H(x|\tau)} [E[H(p|\tilde{\tau})|p \geq x] - x] + x.$$

The above equation holds for any $x \in (0, 1]$. Now if $x = \bar{p}(\tau)$ we have

$$\begin{aligned} E[H(p|\tau)|p \geq \bar{p}(\tau)] &\leq \frac{1 - H(\bar{p}(\tau)|\tilde{\tau})}{1 - H(\bar{p}(\tau)|\tau)} [E[H(p|\tilde{\tau})|p \geq \bar{p}(\tau)] - \bar{p}(\tau)] + \bar{p}(\tau) \\ &\leq [E[H(p|\tilde{\tau})|p \geq \bar{p}(\tau)] - \bar{p}(\tau)] + \bar{p}(\tau) \\ &\leq E[H(p|\tilde{\tau})|p \geq \bar{p}(\tau)] \\ &\leq E[H(p|\tilde{\tau})|p \geq \bar{p}(\tilde{\tau})]. \end{aligned}$$

The second inequality follows from $\bar{p}(\tau) \leq p_0$ and the definition of a mean preserving spread; the fourth inequality comes from $\bar{p}(\tau) \leq \bar{p}(\tilde{\tau})$. \square

Given the wage process, the behavior of Proposition 3 and Proposition 4 can be used to make predictions about volatility of wages on new jobs. A worker with more experience who starts a new match will have more information about the quality of that match than a worker with less experience. In the case where $\bar{p}(\tau) > 0.5$ and is increasing in τ , the model predicts that more experience translates, on average, to a higher value of p at the start of a new job and hence smaller changes in future wages. These implications are quantitatively evaluated with simulations of the model.

While the above propositions deal with the specific learning mechanism in this paper with two types of jobs, the mechanism is more general. In particular, any setup where the variance of beliefs is declining in experience should generate declining wage volatility on new jobs based on the amount of past experience that the worker has. Also, in the model with two types of jobs, job-to-job transitions should strengthen the result as workers will only accept jobs that they believe to be better than their current one and this means that they are more certain of their type. Generally, with more than two types of jobs the variance of the belief can be separated from the level of the current belief potentially breaking this connection.

REFERENCES

- Altonji, J. G. and R. A. Shakotko (1987), “Do wages rise with job seniority?” *Review of Economic Studies*, 54 (3), 437–459. [227]
- Altonji, J. G. and N. Williams (2005), “Do wages rise with job seniority? A reassessment.” *Industrial & Labor Relations Review*, 58 (3), 370–397. [227]
- Chade, H. and E. Schlee (2002), “Another look at the Radner–Stiglitz nonconcavity in the value of information.” *Journal of Economic Theory*, 107 (2), 421–452. [234]

- Cheron, A., J.-O. Hairault, and F. Langot (2011), “Age-dependent employment protection.” *The Economic Journal*, 121 (557), 1477–1504. [227]
- Esteban-Pretel, J. and J. Fujimoto (2014), “Life-cycle labor search with stochastic match quality.” *International Economic Review*, 55, 575–599. [227]
- Fallick, B. and C. A. Fleischman (2004), “Employer-to-employer flows in the U.S. labor market: The complete picture of gross worker flows.” Finance and Economics Discussion Series 2004-34, Board of Governors of the Federal Reserve System (U.S.). [237]
- Farber, H. S. (1994), “The analysis of interfirm worker mobility.” *Journal of Labor Economics*, 12 (4), 554–593. [236, 237]
- Farber, H. S. and R. Gibbons (1996), “Learning and wage dynamics.” *Quarterly Journal of Economics*, 111, 1007–1048. [227, 246]
- Flinn, C. J. (1986), “Wages and job mobility of young workers.” *Journal of Political Economy*, 94 (3), S88–S110. [241]
- Gathmann, C. and U. Schönberg (2010), “How general is human capital? A task-based approach.” *Journal of Labor Economics*, 28 (1), 1–49. [227]
- Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2005), “Comparative advantage, learning, and sectoral wage determination.” *Journal of Labor Economics*, 23 (4), 681–724. [227, 228, 233, 234, 242]
- Gonzalez, F. M. and S. Shi (2010), “An equilibrium theory of learning, search, and wages.” *Econometrica*, 78 (2), 509–537. [227]
- Gorry, A. (2013), “Minimum wages and youth unemployment.” *European Economic Review*, 64 (C), 57–75. [227]
- Gorry, A., G. Devon, and N. Trachter (2014), “Learning and life cycle patterns of occupational transitions.” Working Paper 14-15, Federal Reserve Bank of Richmond. [227]
- Jovanovic, B. (1979), “Job matching and the theory of turnover.” *Journal of Political Economy*, 87 (5), 972–990. [227, 228, 231, 239, 251]
- Jovanovic, B. (1984), “Matching, turnover, and unemployment.” *Journal of Political Economy*, 92 (1), 108–122. [225, 228, 231, 233, 234]
- Kahn, L. B. (2013), “Asymmetric information between employers.” *American Economic Journal: Applied Economics*, 5 (4), 165–205. [226]
- Kahn, L. B. and F. Lange (2013), “Employer learning, productivity and the earnings distribution: Evidence from performance measures.” Working paper, Yale School of Management. [226, 242]
- Kambourov, G. and I. Manovskii (2009), “Occupational mobility and wage inequality.” *Review of Economic Studies*, 76 (2), 731–759. [227]
- Lange, F. (2007), “The speed of employer learning.” *Journal of Labor Economics*, 25, 1–35. [226, 246]

- McCall, B. P. (1990), "Occupational matching: A test of sorts." *Journal of Political Economy*, 98 (1), 45–69. [227]
- Meghir, C. and L. Pistaferri (2004), "Income variance dynamics and heterogeneity." *Econometrica*, 72 (1), 1–32. [242]
- Menzio, G., I. A. Telyukova, and L. Visschers (2012), "Directed search over the life cycle." Working Paper 12-002, Penn Institute for Economic Research, Department of Economics, University of Pennsylvania. [227]
- Moscarini, G. (2005), "Job matching and the wage distribution." *Econometrica*, 73 (2), 481–516. [227, 229]
- Moscarini, G. and L. Smith (2002), "The law of large demand for information." *Econometrica*, 70 (6), 2351–2366. [248]
- Moscarini, G. and K. Thomsson (2007), "Occupational and job mobility in the US." *Scandinavian Journal of Economics*, 109 (4), 807–836. [237]
- Neal, D. (1999), "The complexity of job mobility among young men." *Journal of Labor Economics*, 17 (2), 237–261. [228]
- Neal, D. and S. Rosen (2000), "Theories of the distribution of earnings." In *Handbook of Income Distribution*, Vol. 1 (A. B. Atkinson and F. Bourguignon, eds.), Chapter 7, 379–427, Elsevier, Amsterdam. [226, 242]
- Papageorgiou, T. (2014), "Learning your comparative advantage." *Review of Economic Studies*, 81 (3), 1263–1295. [227]
- Pavan, R. (2010), "The role of career choice in understanding job mobility." *Labour*, 24 (2), 107–127. [228]
- Poletaev, M. and C. Robinson (2008), "Human capital specificity: Evidence from the dictionary of occupational titles and Displaced Worker Surveys, 1984–2000." *Journal of Labor Economics*, 26 (3), 387–420. [227]
- Pries, M. J. (2004), "Persistence of employment fluctuations: A model of recurring job loss." *Review of Economic Studies*, 71 (1), 193–215. [227]
- Radner, R. and J. Stiglitz (1984), "A nonconcavity in the value of information." In *Bayesian Models in Economic Theory* (M. Boyer and R. E. Kihlstrom, eds.), 33–52, Elsevier, Amsterdam. [234]
- Rubinstein, Y. and Y. Weiss (2006), "Post schooling wage growth: Investment, search and learning." In *Handbook of the Economics of Education*, Vol. 1 (E. Hanushek and F. Welch, eds.), Chapter 1, 1–67, Elsevier, Amsterdam. [226]
- Shimer, R. (2012), "Reassessing the ins and outs of unemployment." *Review of Economic Dynamics*, 15 (2), 127–148. [236, 241]
- Topel, R. H. (1991), "Specific capital, mobility, and wages: Wages rise with job seniority." *Journal of Political Economy*, 99 (1), 145–176. [227]

Topel, R. H. and M. P. Ward (1992), “Job mobility and the careers of young men.” *Quarterly Journal of Economics*, 107 (2), 439–479. [241]

Yedid-Levi, Y., N. Jaimovic, H. Siu, and M. Gervais (2015), “What should I be when I grow up? Occupations and employment over the life cycle and business cycle.” Technical report. [227]

Co-editor Jose-Victor Rios-Rull handled this manuscript.

Submitted May, 2013. Final version accepted May, 2015.