Tenure Profiles and Efficient Separation in a Stochastic Productivity Model

I. Sebastian Buhai[1] and Coen N. Teulings[2,3]

[1]Department of Economics, Stockholm University, 106 91 Stockholm, Sweden; sbuhai@gmail.com

[2]Faculty of Economics, University of Cambridge, Cambridge CB3 9DD, United Kingdom; cnt23@cam.ac.uk

[3]Department of Economics, University of Amsterdam

This version: November 2013

Abstract

We develop a theoretical model based on efficient bargaining, where both log outside productivity and log productivity in the current job follow a random walk. This setting allows the application of real option theory. We derive the efficient worker-firm separation rule. We show that wage data from completed job spells are uninformative about the true tenure profile. The model is estimated on the PSID. It fits the observed distribution of job tenures well. Selection of favourable random walks can account for the concavity in tenure profiles. About 80% of the estimated wage returns to tenure is due to selectivity in the realized outside productivities.

Keywords: random productivity growth, efficient bargaining, job tenure, inverse gaussian, wage-tenure profiles, option theory

JEL codes: C33, C41, J31, J63

1 Introduction

A large empirical literature has looked at wage returns to job tenure, see Farber (1999) for a survey. The conclusions of this research still diverge, despite analyzing data from the same countries (mainly the USA) or even the same longitudinal datasets (mostly the PSID), see, e.g., Altonji and Shakotko (1987), Abraham and Farber (1987), Altonji and Williams (1997, 2005), Abowd et al (1999); Topel (1991), Dustmann and Meghir (2005), Buchinsky et al (2010). Here, we propose a new direction for this line of research. From a theoretical point of view, large "true" returns to tenure are problematic. Why would a worker separate when she loses her tenure profile by doing so? Hence, separation is likely to be induced by the firm, what we call a layoff. But why would the worker and the firm prefer separation above renegotiation? Although some models offer explanations for this, the size of the reported wage returns to tenure remains puzzling.

This paper addresses explicitly whether the data is consistent with efficient separations, by modelling simultaneously the evolution of wages and the distribution of job tenures. The model explains the correlation between wages and job tenure from the random evolution of both the job's *inside productivity* and the *outside productivity*, i.e. the productivity in the best alternative job. Separation occurs when the value of the inside productivity falls below the outside productivity. By some form of bargaining, log wages are a linear combination of the in- and outside log productivity. Then, wages and tenure are correlated because only those jobs survive for which the inside productivity remains above the outside productivity. There is no such thing as "the return to tenure" in this model. In some jobs wages go up because the inside productivity evolves favorably. In other jobs wages go down for *mutatis mutandis* the same reason. However, these jobs are gradually eliminated from the stock of ongoing job spells because there are no options left for mutually gainful renegotiation.

We assume both log inside and outside productivities to follow Brownian motions. Since log wage is a linear combination of them, it also follows a random walk. The evolution of an individual's log wage is indeed reasonably described by a random walk with transitory shocks, see, e.g., Abowd and Card (1989), or Meghir and Pistaferri (2004). What we call the return to tenure is then the difference between the drifts of the log wage and of the log outside productivity. Starting a job requires an irreversible specific investment, which is lost upon separation. The combination of irreversibility and productivity following a random walk implies that we can apply the theory of real options, see, e.g., Dixit (1989) and Dixit and Pyndick (1994), compare Teulings and Van der Ende (2000). The predicted hazard rates of this model are well in line with the empirical distribution of job exits. From the distribution of job tenures we are able to estimate the surplus of the inside over the outside productivity, and the drift of this surplus. We obtain a positive drift, indicating that some 10% of all jobs will end only by retirement. We use these results to compute the evolution of the expected surplus, conditioning on both the date of job start and of job termination, an empirical strategy explored earlier by Abraham and Farber (1987). Our model predicts this variable to be correlated to the evolution of wages.

We obtain the following results. First, a closed form expression is computed for the expected surplus in completed job spells. We show this not to depend on the drift of the surplus. The evolution of wages in completed spells is thus uninformative on the return to tenure in this model, since the effect of the drift is exactly offset by the selection due to the elimination of bad matches. This is an unexpected conclusion, given that so many studies have tried to identify the return to tenure from this type of data. Second, we show that our model can easily explain the observed concavity in the tenure profile from the selection effect. Selection is much more important than the drift. Third, we show that the selection effect is driven by the selectivity in the outside, as opposed to the inside, productivity. Workers switch jobs mainly when the outside productivity is high, not so much when the inside productivity is low. Selectivity in the outside option accounts for 50 to 80% of the tenure profile. This source of selectivity usually receives little attention in the literature on wage-tenure profiles, though it figures in models of equilibrium on-the-job-search, e.g., Burdett and Mortensen (1998). In this search literature, closest to us are "persistent earnings dynamics" studies by Postel-Vinay and Turon (2010) and Low et al (2010); unlike them, our wage process does not require nesting within models of search, or classic ARMA decomposition. Finally, our estimation results suggest downward rigidity in wages, as discussed, e.g., by Beaudry and DiNardo (1991). This downward rigidity does not fit the efficient bargaining hypothesis. Our estimates also provide evidence of excess variance in wages for job movers, implying failure of our Walrasian market assumption for outside offers.

The paper is structured as follows: the model is discussed in Section 2, the identification and the estimation are set out in Section 3, and the empirical analysis is presented in Section 4.

2 The Random Productivity Growth Model

2.1 Model Assumption

Consider a labor market in continuous time, where both workers and firms are risk neutral. We focus on a single cohort of homogeneous workers. We normalize our measure of time t such that it is also equal to the workers' experience. There is no disutility of effort, so that the workers' utility depends on their expected lifetime income only. Each firm offers a single job, of which the job specific productivity P_t evolves according to a geometric Brownian motion with drift. At the moment a worker is hired for a vacant job, a specific investment has to be made which is partly paid by the firm and partly by the worker, and which is irreversibly lost upon separation between the worker and the firm. However, the firm retains the option value on the vacancy: it can hire a new worker at any future time, provided that the cost of the specific investment is paid again. The investment is verifiable. There is no search cost involved from either party in finding a new job: an unemployed worker can just pick the most attractive vacancy that is available at that time, at zero cost. Let R_t be the return on this vacancy, net of the cost of investment; R_t is exogenous in this model. Like P_t, it evolves according to a geometric Brownian motion with drift. Both workers and firms are perfectly informed about the current values of P_t and R_t, but their future evolution is unknown. The value of the specific investments for a job starting at time t is $R_t I$. One can think of I as the cost of investment measured in units of labor time and of R_t as the price of one unit at time t. Using lower cases to denote the logs of the corresponding upper cases, the law of motion of p_t and r_t, for $t > s$, is characterized by a bivariate normal distribution:

$$\begin{bmatrix} p_t - p_s \\ r_t - r_s \end{bmatrix} \sim N\left[(t-s)\underline{\mu}, (t-s)\Sigma\right]$$

where:

$$\underline{\mu} = \begin{bmatrix} \mu_p \\ \mu_r \end{bmatrix}, \Sigma = \begin{bmatrix} \sigma_p^2 & \sigma_{pr} \\ \sigma_{pr} & \sigma_r^2 \end{bmatrix} \tag{1}$$

Since μ_r is the drift in the log outside option of the worker, it can be interpreted as the sum of the return to experience and the secular growth in real wages due to technological progress. The worker and the firm bargain over the surplus of the productivity of the job above the shadow price of a worker, $P_t - R_t$. This bargaining is efficient: as long as there is a surplus, the worker and the firm will agree on a sharing rule. In the empirical application in Sections 3 and 4, I and μ_p will be allowed to depend on personal characteristics. For the derivation of the model this dependence on personal characteristics can be ignored.

2.2 Value of a Job and a Vacancy

Three assumptions made above greatly simplify the analysis. (i) The risk neutrality of both players implies that the allocation of risk is irrelevant: only expected values matter. (ii) The verifiability of investment implies that there are no hold up problems: the distribution of future surplusses $P_t - R_t$, $t > s$, is irrelevant for the timing of the investment decision, since the cost of the specific investment $R_s I$ can always be shared between the worker and the firm according to their share in future surplusses. Hence, the investment decision will maximize the joint expected surplus of the worker and the firm. (iii) Efficient bargaining implies that separation decisions will also maximize the joint expected surplus. Hence, separation occurs at mutual consent when there are no gains from trade left. Quits and layoffs are therefore observationally equivalent, as in McLaughlin (1991). For convenience, we shall refer to a separation as the firm firing the worker, though it can be both a quit or a layoff. Given these assumptions, wage setting and separation decisions can be analyzed separately: in the spirit of the Coase theorem, hiring and firing decisions maximize the joint expected surplus, regardless of its distribution.

First, we analyze hiring and firing. Since hiring requires an irreversible investment, while firing is an irreversible disinvestment, both can be analysed using real option theory, see, e.g., Dixit and Pindyck (1994). The easiest way to analyze this problem is to assume that workers always get paid their shadow price R_t. Then, hiring and firing simply maximize the expected value of the firm. Let $V(p_t, r_t)$ and $J(p_t, r_t)$ be the expected present value of a vacancy and respectively of a job, as functions of p_t and r_t. Applying Ito's lemma, the Bellman equations for both value functions read, compare Dixit and Pindyck (1994: pp.140-141):

$$\begin{aligned}\rho J &= \exp(p_t) - \exp(r_t) + \mu_p J_p + \mu_r J_r + \frac{1}{2}\sigma_p^2 J_{pp} + \sigma_{pr} J_{pr} + \frac{1}{2}\sigma_r^2 J_{rr} \qquad (2)\\ \rho V &= \mu_p V_p + \mu_r V_r + \frac{1}{2}\sigma_p^2 V_{pp} + \sigma_{pr} V_{pr} + \frac{1}{2}\sigma_r^2 V_{rr}\end{aligned}$$

where we leave out the arguments of $J(\cdot)$ and $V(\cdot)$ for convenience, and where ρ denotes the interest rate. The term $\exp(p_t) - \exp(r_t)$ in the first equation is the value of current output minus the wage of the worker; the other terms capture the wealth effects due to changes in the state variables p_t and r_t: the first order derivatives capture the effect of the drift in both state variables, the second order derivatives capture the effect of their variance. For optimal hiring and firing, value matching and smooth pasting conditions should be satisfied:

$$\begin{aligned} J(p_S, r_S) &= V(p_S, r_S) + \exp(r_S)I, \quad V(p_T, r_T) = J(p_T, r_T) \\ J_p(p_S, r_S) &= V_p(p_S, r_S), \quad V_p(p_T, r_T) = J_p(p_T, r_T) \end{aligned} \tag{3}$$

where S is the moment of hiring and T is the moment of firing. The first two conditions are the value matching conditions for hiring and firing respectively, the second pair of conditions are the smooth pasting conditions for p_t; the smooth pasting conditions for r_t are redundant. Value matching conditions impose value equality at the moment of hiring and firing; on top of that, smooth pasting conditions impose that slight variations in p_t should not affect the value equality, since hiring and firing decisions are irreversible. Hence, a decision maker should not regret her decision after slight variations in p_t. The above conditions and the Bellman equations (2) jointly determine $J(\cdot)$ and $V(\cdot)$.

Define $b_t \equiv p_t - r_t$; b_t is the log relative surplus P_t/R_t. By (1), $b_t - b_s \sim \mathrm{N}\left[(t-s)\mu, (t-s)\sigma^2\right]$, with

$$\mu \equiv \mu_p - \mu_r, \quad \sigma^2 \equiv \sigma_p^2 + \sigma_r^2 - 2\sigma_{pr} \tag{4}$$

Proposition 1 *The value functions $J(\cdot)$ and $V(\cdot)$ can be written as $J(p_t, r_t) = \exp(r_t)\, j(p_t - r_t)$ and $V(p_t, r_t) = \exp(r_t)\, v(p_t - r_t)$, where $j(\cdot)$ and $v(\cdot)$ satisfy:*

$$\begin{aligned} \left(\rho - \mu_r - \frac{1}{2}\sigma_r^2\right) j &= \exp(b_t) - 1 + \left(\mu + \sigma_{pr} - \sigma_r^2\right) j' + \frac{1}{2}\sigma^2 j'' \\ \left(\rho - \mu_r - \frac{1}{2}\sigma_r^2\right) v &= \left(\mu + \sigma_{pr} - \sigma_r^2\right) v' + \frac{1}{2}\sigma^2 v'' \end{aligned} \tag{5}$$

leaving out the argument of $j(\cdot)$ and $v(\cdot)$ for convenience.

Proof: The definition of $j(\cdot)$ and $v(\cdot)$ implies $J_p = \exp(r_t) j'$, $J_{pp} = \exp(r_t) j''$, $J_r = \exp(r_t)(j - j')$, $J_{rr} = \exp(r_t)(j - 2j' + j'')$, and $J_{pr} = \exp(r_t)(j' - j'')$, and likewise for $V(\cdot)$. The proposition follows directly from substitution of these expressions in equation (2) and (3).■

The factor $\rho - \mu_r - \frac{1}{2}\sigma_r^2$ is a modified discount rate, which accounts for the fact that future revenues are discounted at rate ρ, but increase in expectation at rate $\mu_r + \frac{1}{2}\sigma_r^2$ due to the drift and the variance of R_t. The hiring and separation rules depend therefore purely on b_t: a vacancy should be filled at the first time t that $b_t = b^S$, a worker should be fired from the job at the first time t that $b_t = b^T$. This proposition characterizes the decision problem of the firm by two second order differential equations, four boundary conditions and two decision parameters, b^S and b^T, see Dixit and Pindyck (1994; ch. 5.1-5.2), to whom we refer for the subsequent arguments. The two differential equations have an analytical solution. These solutions yield four constants of integration. Two of these constants have to be zero due to transversality conditions. The constants of integration reflect the option value for the firm of hiring and firing a worker. The option value of hiring converges to zero when $b_t \to 0$, while the option value of firing converges to zero when $b_t \to \infty$. These constraints can only be satisfied by setting two constants of integration equal to zero. Hence, the four boundary conditions determine four unknown parameters: b^S, b^T, and the two remaining constants of integration. One can prove $b^T < 0 < b^S$. Hence, at the moment of hiring, $P_S > R_S$ because the firm has to recoup the cost of investment and because the investment is irreversible, so that the firm loses the option value of delaying hires, while in the meantime b_t might fall below b^S at a later point in time. Similarly, at the moment of firing, $P_T < R_T$ because the firm accepts some losses before firing the worker, since by doing so it loses the option value of firing the worker at a later point in time.

2.3 Job Tenure Distribution

The duration of a job spell is a stochastic variable, equal to the time it takes the random variable b_t to travel down from b^S to b^T. The standard deviation of b_t is unidentified in this model because, for any time t, we observe only whether the spell is still incomplete, implying $b_t - b^T > 0$ ever since the start of the job spell. We can therefore normalize all parameters by σ. For each job spell, we define $\tau \equiv t - S$, with $\tau \geq 0$, and respectively $\Theta \equiv T - S$, with $\Theta > 0$; τ is the incomplete tenure, while Θ is the completed tenure of that job spell. Define:

$$\Omega_\tau \equiv \frac{b_t - b^T}{\sigma}, \quad \Omega \equiv \frac{b^S - b^T}{\sigma} > 0, \quad \pi \equiv \frac{\mu}{\sigma} \tag{6}$$

Thus Ω_τ is a Brownian motion with drift π and unit variance per unit time. By construction, $\Omega_0 = \Omega$ and $\Omega_\Theta = 0$. Ω can be shown to be an implicit function of the model's structural parameters, $I \equiv H\left(\Omega, \underline{\mu}, \Sigma\right)$, with $H_\Omega(\cdot) > 0$. Hence, we treat the parameter Ω as the parameter of interest. If desired, the underlying structural parameter I can be recovered via the function $H\left(\Omega, \underline{\mu}, \Sigma\right)$.

The completed job spell Θ is determined by the time it takes Ω_τ to pass the barrier $\Omega_\tau = 0$ for the first time. This process satisfies the "First Passage Time" distribution, which has been applied previously by Lancaster (1972) for modelling strike durations, and by Whitmore (1979) for job spells. The unconditional density of $\Omega_\tau = \omega$ reads:

$$\frac{1}{\sqrt{\tau}} \phi\left(\frac{\omega - \Omega - \pi\tau}{\sqrt{\tau}}\right)$$

where $\phi(\cdot)$ is the standard normal PDF. However, a job spell is completed if and only if Ω_t has not been negative for any $t \in [0, \tau]$. Hence, we are interested in the density of Ω_τ conditional on $\Omega_t > 0, \forall t \in [0, \tau]$. For this conditioning, we can apply the Reflection Principle, first discussed by Feller (1968): there is a one-to-one correspondence between trajectories of Ω_τ from Ω to ω which have crossed the barrier $\Omega_\tau = 0$ at least once, and trajectories of Ω_τ from $-\Omega$ to ω. These trajectories should therefore be subtracted to obtain the conditional density of Ω_τ. Define: $g(\omega, \tau) \equiv \Pr(\Omega_\tau = \omega \wedge \Theta > \tau)$. It satisfies, see, e.g., Kijima (2003, p.185-187):

$$g(\omega, \tau) = \frac{1}{\sqrt{\tau}} \left[\phi\left(\frac{\omega - \Omega - \pi\tau}{\sqrt{\tau}}\right) - e^{-2\Omega\pi} \phi\left(\frac{\omega + \Omega - \pi\tau}{\sqrt{\tau}}\right) \right] \tag{7}$$

where $\phi(.)$ is the standard normal density function. The first term in square brackets is the unconditional density; the second is the effect of the conditioning. By the Reflection Principle, the latter is the density of trajectories of Ω_τ from $-\Omega$ to ω. The factor $e^{-2\Omega\pi}$ corrects for the differential effect of the drift on the density for upward and downward trajectories. By integrating out ω we get the cumulative distribution of jobs surviving at τ, $\overline{F}(\tau) = \Pr(\Theta > \tau)$:

$$\overline{F}(\tau) \equiv \Phi\left(\frac{\Omega + \pi\tau}{\sqrt{\tau}}\right) - e^{-2\Omega\pi} \Phi\left(\frac{-\Omega + \pi\tau}{\sqrt{\tau}}\right) \tag{8}$$

where $\Phi(.)$ is the standard normal CDF. This expression is identical to Whitmore (1979: eq. 2). The distribution of Θ is therefore fully specified by two parameters, the initial distance from the separation threshold Ω, and the drift π. Hence, Ω and π can be identified from data on job tenures, while the parameter σ cannot. The corresponding density function is minus the derivative of $\overline{F}(\tau)$ with respect to τ:

$$f(\tau) = \frac{\Omega}{\tau\sqrt{\tau}} \phi\left(\frac{\Omega + \pi\tau}{\sqrt{\tau}}\right) \tag{9}$$

where we use $\phi\left(\frac{\Omega+\pi\tau}{\sqrt{\tau}}\right) = e^{-2\Omega\pi}\phi\left(\frac{-\Omega+\pi\tau}{\sqrt{\tau}}\right)$. The exit rate $f(\tau)/\overline{F}(\tau)$ can be shown to be hump shaped, starting from 0, reaching a peak at τ^*, $0 < \tau^* < \frac{2}{3}\Omega^2$, and afterwards either declining monotonically to 0 for a positive drift $\pi > 0$, or to ${}^1/{}_2\pi^2$ for $\pi < 0$. Farber (1994) and Horowitz and Lee (2002) have documented this hump shaped pattern using NLSY data. A positive drift $\pi > 0$ implies that some fraction of the started jobs will never end. This fraction is equal to the survivor function (8) evaluated at $\tau \to \infty$, hence to $1 - e^{-2\Omega\pi}$.

In Figure 1, we plot the exit rates for the pairs $\Omega = e^{-1.20} \simeq 0.30$, $\pi = 0.14$, and respectively $\Omega = e^{-1.24} \simeq 0.29$, $\pi = 0.23$, which are the two estimates for Ω and π for the mean values of the observed and unobserved worker characteristics, see Section 3, Table 2. In both cases the peak is reached at $\tau \simeq 0.04$ years. Since $\pi > 0$, the hazard rate converges to zero and a positive fraction of the jobs, about 10%, will never end.

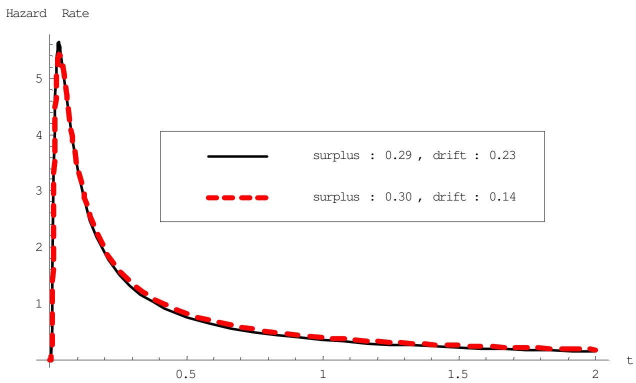

Figure 1: Predicted Job Hazards

2.4 Tenure Profile in Wages

2.4.1 Sharing Rule of Surpluses and Wages

We extend this model with an explicit sharing rule of surpluses during the course of the job spell. Ideally, we would derive this sharing rule from an explicit bargaining game, such as Nash bargaining. For the sake of convenience, we use a simpler approach, imposing the log linearity of the sharing rule *a priori*, and deriving the intercept of that rule from the assumption of efficient bargaining. According to this rule, the worker's log wage w_t satisfies:

$$w_t = r_t + \beta\left(b_t - b^T\right) + u_t = r_t + \overline{\sigma}\Omega_\tau + u_t \tag{10}$$

where $\overline{\sigma} \equiv \beta\sigma$. We assume u_t is an i.i.d. random variable distributed $N\left(0, \sigma_u^2\right)$: this specification of wages as following a random walk, Ω_τ, with a transitory shock MA(1), u_t, is broadly consistent with a large number of studies on the dynamics of wages, see, e.g., MaCurdy (1982), Abowd and Card (1989), Topel and Ward (1992), and Meghir and Pistaferri (2004). The parameter β can be interpreted as the worker's bargaining power. If $\beta = 0$, the wage is equal to the worker's outside productivity R_t, while if $\beta = 1$, the wage is her inside productivity P_t. The transitory error can be interpreted as either measurement error in wages, see, e.g., Meghir and Pistaferri (2004), or as short run fluctuations that do not affect the long run payoff of the specific investment I in the current job. In either interpretation, these shocks do not affect the optimizing behavior of agents regarding job change.

2.4.2 Selectivity in Tenure Profiles

Equation (10) implies that log wages within a job follow a Brownian motion with drift $\mu_r + \overline{\sigma}\pi$; μ_r is the sum of the return to experience and the secular growth of real wages due to technological progress; $\overline{\sigma}\pi$ is the deterministic part of the tenure profile. Were the realizations of Ω_τ independent of the completed job tenure Θ, $\overline{\sigma}\pi$ could be estimated easily. However, in completed job spells, Ω_τ is correlated to Θ for three reasons: (i) $\Omega_0 = \Omega$, (ii) $\Omega_\Theta = 0$, and (iii) $\Omega_t > 0, \forall t, 0 \leq t < \Theta$. For the sake of brevity, we refer to this information set as $A(\Theta)$. *Mutatis mutandis,* the same applies to incomplete spells. Let Ψ be the incomplete tenure at the last date for which data are available. Again, there are three pieces of information: (i) $\Omega_0 = \Omega$, (ii) $\Theta > \Psi$, and hence (iii) $\Omega_t > 0, \forall t, 0 \leq t \leq \Psi$. We refer to this second information set as $B(\Psi)$.

Proposition 2 *Let* $m(\tau) \equiv \frac{\Theta-\tau}{\Theta}$. $E[\Omega_\tau|A(\Theta)]$ *and its derivatives satisfy*:

$$\begin{aligned} E[\Omega_\tau|A(\Theta)] &= 2\sqrt{m(\tau)\tau}\phi\left(\sqrt{m(\tau)/\tau}\Omega\right) - \left(\frac{\tau}{\Omega} + m(\tau)\Omega\right)\left[1 - 2\Phi\left(\sqrt{m(\tau)/\tau}\Omega\right)\right] \\ \lim_{\tau\to 0}\frac{dE[\Omega_\tau|A(\Theta)]}{d\tau} &= \frac{1}{\Omega} - \frac{\Omega}{\Theta}, \quad \lim_{\tau\to\Theta}\frac{dE[\Omega_\tau|A(\Theta)]}{d\tau} = -\infty, \quad \frac{d^2E[\Omega_\tau|A(\Theta)]}{d\tau^2} < 0 \end{aligned}$$

Proof See Appendix.■

This proposition implies that E$[\Omega_\tau|A(\Theta)]$ does not depend on the tenure profile in wages, $\overline{\sigma}\pi$; see also Van der Ende (1997). Hence, conditional on the model that we specified, the evolution of wages in completed job spells does not provide any information whatsoever on the tenure profile in wages. Given the many papers that have tried to estimate tenure profiles from data on completed job spells, this is a very surprising conclusion. The intuition for this result is that an increase in $\overline{\sigma}\pi$ has two offsetting effects on ΔE$[\Omega_\tau|A(\Theta)]$. On the one hand, it raises the deterministic part of the tenure profile, so that the change in the unconditional expectation ΔE$[\Omega_\tau]$ goes up. On the other hand, it makes separation a less likely event, so that the condition $A(\Theta)$ becomes more informative: the higher is $\overline{\sigma}\pi$, the more unfavourable the evolution of the non-deterministic part of $\Delta\Omega_\tau$ must have been to warrant a separation. Hence, the deterministic part of the tenure profile does affect the job separation rate, but it does not affect the evolution of wages, conditional on the moment of separation Θ.

The conclusion above depends crucially on the assumption of efficient bargaining. Ignoring the impact of temporary shocks u_t, this assumption dictates that the evolution of wages over a job spell satisfies

$$(w_S - r_S) - (w_T - r_T) = \overline{\sigma}\Omega$$

see equation (10). The difference between the starting and the terminal value of this log relative wage is equal to the worker's share in the surplus due the specific investment in the job, $\overline{\sigma}\Omega$. Hence, irrespective of the steepness of the tenure profile $\overline{\sigma}\pi$, or the job spell length Θ, log relative wages decline by $\overline{\sigma}\Omega$ over the duration of a completed job spell. However, π can be estimated from the tenure distribution. Efficient bargaining implies that this distribution is informative on the tenure profile, since under efficient bargaining a higher tenure profile means that jobs will survive longer. From this perspective, data on the tenure distribution are more informative on the return to tenure than data on wages.

The second relationship of Proposition 2 says that the initial slope of $\mathrm{E}[\Omega_\tau|A(\Theta)]$ is negative for short spells, $\Theta < \Omega^2$, even when the drift is positive, $\pi > 0$. For these spells, $\mathrm{E}[\Omega_\tau|A(\Theta)]$ must decline immediately for $\Omega_\Theta = 0$. The third expression shows that the expected surplus declines infinitely fast just before separation. This is consistent with empirical evidence by Jacobson, LaLonde and Sullivan (1993) on the decline in relative wages in the period just before firing. The final expression shows that the second derivative is always negative. Hence, $\mathrm{E}[\Omega_\tau|A(\Theta)]$ is concave in τ; it is monotonically decreasing for short spells $\Theta < \Omega^2$ and it is hump-shaped for longer spells. Contrary to the case of completed spells, there is no explicit expression for $\mathrm{E}[\Omega_\tau|B(\Psi)]$. Hence, we use numerical integration in this case, see the Appendix.

Figure 2 plots the evolution of $\mathrm{E}[\Omega_\tau|\cdot]$ for $\Omega = 0.30$ and $\pi = 0.14$, for both completed spells (continuous lines) and incomplete spells (dashed lines), with durations Θ and respectively Ψ in $\{0.1,\ 2, 5, 10\}$. Moreover, the straight line shows the drift: $\Omega + \pi\tau$. With respect to completed spells, $\mathrm{E}[\Omega_\tau|A(\Theta)]$ is monotonically decreasing for $\Theta \leq 0.1$ year, and concave for larger Θ. The top of the profile is increasing in Θ, showing the importance of conditioning on the eventual tenure. With respect to incomplete spells, $\mathrm{E}[\Omega_\tau|B(\Psi)]$ is increasing in Ψ. The reason is that higher values of Ψ imply greater selectivity, since $\Theta > \Psi$. Trajectories are strongly concave, indicating that selection plays an important role. This can explain the observed concavity of tenure profiles in log wages: the underlying profile might be linear, with the observed concavity simply due to selection. The trajectories for both completed and incomplete spells are far above the deterministic part, except for the final year(s) before separation: selection dominates.

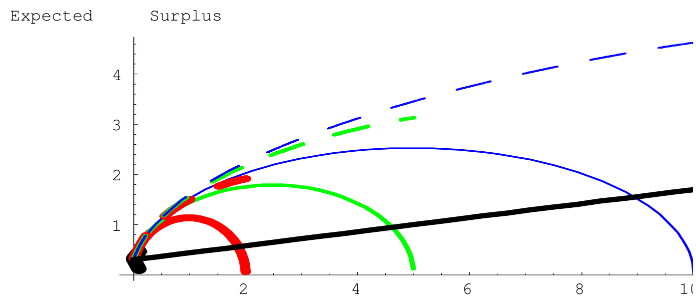

Figure 2: Expected surplus in completed and incomplete spells

The same analysis can be done for the second moment of Ω_τ. Expressions for $\mathrm{Var}[\Delta\Omega_\tau|A(\Theta)]$ and $\mathrm{Var}[\Delta\Omega_\tau|B(\Psi)]$ can be found in the Appendix. In the absence of condition $A(\Theta)$, $\mathrm{Var}[\Delta\Omega_\tau]$ would be equal to unity, see definition (6). However, conditions $A(\Theta)$ and $B(\Psi)$ introduce selectivity in the trajectories of the random walk.

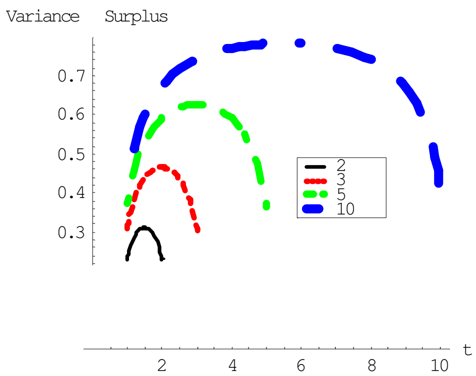

Figure 3: Variance surplus in completed spells

This selectivity reduces the variance, as shown in Figure 3. The variance is low initially, because the positive constraint Ω_τ over the course of a job spell is quite informative, and Ω_0 is small. The same argument applies towards the end of completed spells, since $\Omega_\Theta = 0$ by construction. For longer spells, the variance converges to unity in the middle part of the spell.

2.5 A Reinterpretation of the Wage Equation

The implications of our analysis on the selectivity in tenure profiles surface most clearly when we rewrite equation (10), benefitting from a decomposition of the random variables $[\Delta p_t, \Delta r_t]$ into two orthogonal components $[\Delta b_t, \Delta z_t]$. We normalize Δz_t such that its marginal effect on Δr_t and Δp_t is equal to unity:

$$\begin{aligned} \Delta r_t &= \Delta z_t - \gamma \Delta b_t \\ \Delta p_t &= \Delta z_t + \overline{\gamma} \Delta b_t \end{aligned}$$

where $\Delta z_t \tilde{} N\left(\mu_z, \sigma_z^2\right)$ and where $\overline{\gamma} \equiv 1 - \gamma$, with:

$$\mu_z = \mu_r + \gamma \mu_p, \;\; \sigma_z^2 = \sigma^{-2}\left(\sigma_p^2 \sigma_r^2 - \sigma_{pr}^2\right), \;\; \gamma = \sigma^{-2}\left(\sigma_r^2 - \sigma_{pr}\right) \tag{11}$$

This decomposition satisfies the constraint $\Delta p_t - \Delta r_t = \Delta b_t$ imposed by equation (4).

Since separation decisions are determined by the evolution of b_t, and since Δb_t and Δz_t are uncorrelated, selectivity affects Δb_t, but not Δz_t. Combining these definitions with equation (10) yields:

$$\Delta w_t = \Delta z_t + \overline{\gamma} \beta \Delta b_t + \Delta u_t = \Delta z_t + \overline{\gamma \sigma} \Delta \Omega_\tau + \Delta u_t \tag{12}$$

The parameter γ is a reflection of the correlation between the match surplus and the reservation wage. In the one extreme case, $\gamma = 0$, we can write $\Delta p_t = \Delta r_t + \Delta b_t$, where the right-hand side variables are uncorrelated. Then Δr_t reflects the evolution of the general human capital of the worker, which evolves independently of the value of the specific capital in the present job, Δb_t. Hence, the duration of the actual job is fully determined by its own (mis)fortune. Though the distinction between quits and layoffs makes little sense in this model, separations look like layoffs in this case: the firm fires the worker since she is no longer productive. In the opposite extreme case, $\gamma = 1$, we can write $\Delta r_t = \Delta p_t - \Delta b_t$, again with uncorrelated right-hand side variables. Now Δp_t reflects the evolution of the general human capital of the worker; Δb_t reflects the specific evolution of outside opportunities, e.g. new technologies emerging in other firms. Separations look like quits in this case: the worker quits because she can get a better job elsewhere. In this case, the selectivity of job relocation is not so much that of the type "only good jobs survive outside offers", but more of the type "only good outside offers kill the job".

3 Estimation Methodology

3.1 Identification

The model has in total eight structural parameters: 2 drift parameters $\underline{\mu}$, 3 (co)variances Σ, the initial surplus Ω, the worker's bargaining power β, and the variance of the transitory shock σ_u^2. As shown in Section 2.3, the distribution of completed job tenures is fully determined by two parameters, π and Ω, while wages are characterized by five parameters, $\overline{\sigma}, \gamma, \mu_z, \sigma_z$, and σ_u. Hence, one can never hope to identify more than seven parameters, two from the tenure distribution, and five from wages. The model is therefore identified up to one degree of freedom. Equation (12) reveals why this is the case. Only the product $\overline{\sigma} \equiv \beta\sigma$ shows up, not its components β and σ. Data on either the cost of necessary investment I or the productivity p_t would resolve this underidentification, offering direct information on σ. Then, β could be established as $\overline{\sigma}/\sigma$. However, neither type of data is available here.

We assume Ω and μ_p to depend on personal characteristics. We allow for both observed and unobserved characteristics. As observable we enter experience at the moment of job start S, measured in deviation from its sample mean. Since we have longitudinal data with multiple job spells per individual, we can account for random worker effects; e_Ω and e_π are normally distributed, uncorrelated, random worker effects with zero mean and standard deviations σ_Ω

and σ_π respectively. We apply an exponential specification for Ω since this parameter is positive by definition:

$$\begin{aligned} \Omega &= \exp\left(\omega_0 + \omega_1 S + e_\Omega\right) \\ \pi &= \pi_0 + \pi_1 S + e_\pi \\ \mu_z &= \mu_0 + \gamma\overline{\sigma}\left(\pi_1 S + e_\pi\right) \end{aligned} \tag{13}$$

Since experience at job start S enters the analysis in deviation from its mean across jobs, the intercepts ω_0, π_0, and μ_0 can be interpreted as the mean value for $\ln\Omega, \pi$, and μ_z respectively.

Our estimation strategy uses the recursive feature of our model. The parameters $\omega_{0,1}, \pi_{0,1}, \sigma_\Omega$ and σ_π are estimable from job spell data. Estimates of these parameters are then used to calculate conditional expectations and variances of the change in surplus $\Delta\Omega_\tau$, in both completed and incomplete job spells. The resulting expressions are subsequently employed in the analysis of wage dynamics. Below, we first show how $\omega_{0,1}, \pi_{0,1}, \sigma_\Omega$ and σ_π can be estimated by maximum likelihood from the tenure distribution; next, we derive a set of moment conditions for estimating $\mu_0, \sigma_z^2, \overline{\sigma}, \gamma$, and σ_u^2 from the evolution of wages.

3.2 Maximum Likelihood Estimation of $\omega_{0,1}$ and $\pi_{0,1}$

The contribution to the log likelihood function for an individual reads:

$$\log L = \ln \int\int \prod_{j=1}^{J} \overline{F}(\Psi_j)^{1-d_j} \cdot f(\Theta_j)^{d_j} d\Phi\left(\frac{e_\Omega}{\sigma_\Omega}\right) d\Phi\left(\frac{e_\pi}{\sigma_\pi}\right) \tag{14}$$

where j is the j^{th} job spell, and where d_j is a dummy variable, taking the value $d_j = 1$ if the job spell is completed and $d_j = 0$ otherwise. There are two reasons why we need to make amendments to this likelihood function.

First, we could restrict the estimation to job spells starting within the observation range of our dataset. However, then we would not consider any of the jobs started before they were first reported in the data; by construction, this would limit the maximum completed tenure to the maximum time span, 17 years, covered by our PSID sample, cf. Section 4. Since long tenures contain relevant information, we want to include also the spells ongoing at the start of the sample. Since we observe these spells only conditional on the fact that they have lasted till the start of the sample period, we correct the initial likelihood function for this conditioning:

$$\log L = \ln \int \int \overline{F}(\tau_1)^{-1} \prod_{j=1}^{J} \overline{F}(\Psi_j)^{1-d_j} \cdot f(\Theta_j)^{d_j} d\Phi(\frac{e_\Omega}{\sigma_\Omega}) d\Phi(\frac{e_\pi}{\sigma_\pi}) \tag{15}$$

where τ_1 is the tenure in the job at the start of its observation in the PSID (for which $j = 1$).

Second, since the PSID stock samples data at yearly intervals, job spells completed in less than a year are underreported. We know the elapsed tenure in months at the first moment a job spell is observed, by a retrospective question, but do not know whether there has been another job spell between the job observed a year ago and the current job. Since the hazard rate implied by our model is hump shaped, with the hump likely to be within the first year, cf. Farber (1994), we are likely to overestimate Ω and π, as we miss part of the short tenures in our data. One solution to this problem is to apply a similar conditioning as in equation (15), where τ_j is the initial tenure at the first moment the job is observed. However, this approach does not use the distribution of τ_j's itself. We can use this distribution if we are prepared to make the additional assumption that the starting date of job spells is distributed uniformly over the first year. Then, the density $q(\cdot)$ of initial dates of spells that started throughout the year and are still incomplete at the end of the year satisfies: $q(\tau) = \overline{F}(\tau) / \int_0^1 \overline{F}(x)\, dx$. The total contribution to the likelihood of a spell with initial tenure τ and completed tenure Θ is thus:

$$\frac{f(\Theta)}{\overline{F}(\tau)} q(\tau) = \frac{f(\Theta)}{\int_0^1 \overline{F}(x)\, dx}$$

Hence, the final log likelihood, accounting both for the jobs started before their first reporting in the PSID and for the underreporting of spells shorter than a year due to the stock sampling scheme of the data, can be written:

$$\log L = \ln \int \int \prod_{j=1}^{J} \frac{\overline{F}(\Psi_j)^{1-d_j} \cdot f(\Theta_j)^{d_j}}{\overline{F}(\tau_1)^{I(j=1)} \left(\int_0^1 \overline{F}(x)dx \right)^{I(j\neq 1)}} d\Phi \left(\frac{e_\Omega}{\sigma_\Omega} \right) d\Phi \left(\frac{e_\pi}{\sigma_\pi} \right) \tag{16}$$

where $I(y)$ is the indicator function, taking value 1 if y is true and 0 otherwise. We estimate the log-likelihood function in (16) by Simulated Maximum Likelihood (SML). We report estimates for two samples, a "small" one including only jobs starting within the observation range of our PSID extract, and a "large" one including also the jobs ongoing at the start of the dataset.

3.3 Moment Conditions for Wage Dynamics

Using the maximum likelihood estimates of the parameters $\omega_{0,1}, \pi_{0,1}, \sigma_\Omega$ and σ_π, we can calculate the conditional expectations and variances of $\Delta\Omega_\tau$. These expressions are then used for the estimation of the parameters $\overline{\sigma}, \gamma, \mu_0, \sigma_z$, and σ_u by a method of moments, using data on wage changes. The conditions for the first two moments can be derived by substitution of equation (13) in equation (12), and taking the expectation and the variance. This yields the following system of equations:

$$\begin{aligned}
\Delta w_t &= \mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\tau + \varepsilon_t \qquad (17)\\
w_t - w_{t-1}^* &= \mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\Theta^* + \overline{\sigma}\Omega_0 + \zeta_t \\
\Delta w_t^2 &= \sigma_z^2 + 2\sigma_u^2 + \left(\overline{\gamma\sigma}\right)^2 \mathrm{Var}\Delta\Omega_\tau + \left(\mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\tau\right)^2 + \eta_t \\
\left(w_t - w_{t-1}^*\right)^2 &= \sigma_z^2 + 2\sigma_u^2 + \left(\overline{\gamma\sigma}\right)^2 \mathrm{Var}\Delta\Omega_\Theta^* + \left(\mu_0 + \gamma\overline{\sigma}\pi S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\Theta^* + \overline{\sigma}\Omega_0\right)^2 + \nu_t \\
\Delta w_t \Delta w_{t-1} &= -\sigma_u^2 + \left(\mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\tau\right)\left(\mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_{\tau-1}\right) + \upsilon_t
\end{aligned}$$

where ε_t, ζ_t, η_t, ν_t, and υ_t are i.i.d. errors, where superscript $*$ indicates the log wage or respectively change in surplus just before separation from the previous job, and where $\mathrm{E}\Delta\Omega_\tau \equiv \mathrm{E}[\Delta\Omega_\tau|A(\Theta), B(\Psi)]$, and $\mathrm{Var}\Delta\Omega_\tau \equiv \mathrm{Var}[\Delta\Omega_\tau|A(\Theta), B(\Psi)]$. We impose no constraints upon the covariance matrix of these five error terms. In the third and fourth lines we use $\mathrm{Var}[\Delta w_t] = \mathrm{E}\left[\Delta w_t^2\right] - \mathrm{E}[\Delta w_t]^2$, where we substitute $\mathrm{E}[\Delta w_t]$ with the deterministic part of the right hand side in the first, and respectively the second equation. The first two equations are the conditions for the first moment, for within and between job spell wage changes, respectively. The second pair of equations are the conditions for the second moment, again for within and between job spell wage changes. The final moment condition is the covariance between subsequent wage changes due to the transitory shocks u_t.

The system of equations (17) is characterized by additive disturbances and nonlinear cross-equation restrictions in the parameters. It can be estimated by Feasible Generalized Non-linear Least Squares (FGNLS). Since we use $\omega_{0,1}, \pi_{0,1}, \sigma_\Omega$ and σ_π as estimated by Simulated Maximum Likelihood (SML) in the first step analysis, c.f. section 3.2, in this second step we need to correct the standard errors of the FGNLS estimates for the estimation error introduced by the SML. We follow the methodology outlined by, e.g., Murphy and Topel (1985). Details on the FGNLS estimation and on the adjustment of the standard errors of the FGNLS estimates are relegated to our web appendix.

4 Empirical Analysis

4.1 Data

We use as data a PSID extract of 18 waves, covering the years 1975 through 1992. Our model does not work well when employed people consider other alternatives than switching to another job, such as retirement, leaving the labor force, or taking up full time education. The availability of these other alternatives yields two problems. First, we do not observe the reservation wage at the point of separation when people do not accept another job. Second, with only one alternative to the present job, the decision problem is simply whether a particular indicator switches signs. With more alternatives, that choice process becomes far more complicated. Therefore we restrict the sample to people who do not switch in and out the labor force regularly, and for whom retirement is not a relevant option: white male heads of household, with more than 12 years of education, and less than 60 years of age. Furthermore, we restrict the sample to those individuals that were employed, temporarily laid off, or unemployed at the time of the survey. We exclude people from Alaska or Hawaii. Finally, we discard all observations on unionized jobs. Through these initial data selection procedures we discard in total 10351 observations from the original dataset used by Altonji and Williams (1999).

For the analysis of wage dynamics, observations for which wages are missing (2404 obs.) or topcoded (254 obs.) are deleted, as well as observations for which $|\Delta w_t| > 0.50$ (276 obs.). This leaves a total of 8082 observations on within-job wage changes, and 462 observations on between-job wage changes. Wages are deflated by the implicit price deflator, using 1982 as base year, as in Altonji and Williams (1999).

Table 1: Summary statistics

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**Observations**
logwage$^{(1)}$	2.42	0.52	0.17	4.82	13660
tenure (years)	6.67	7.42	0.08	43.69	15504
experience (years)	14.58	9.21	0.12	43.69	16179

Dataset for Estimating the Tenure Distribution Parameters$^{(2)}$	
Number of individuals	2421
Total number job spells	4681
- started before the observation range	1512
- started within the observation range	3169
Completed job spells	1712
- started before the observation range	372
- started within the observation range	1340
Incomplete job spells	2969
- started before the observation range	1140
- started within the observation range	1829

$^{(1)}$reported average hourly wage, deflated using the implicit price deflator with 1982 base year
$^{(2)}$subset of the data summarized in the top panel, keeping one observation for each job spell

Table 1 presents summary statistics of the data. Tenure and experience measures are constructed by Altonji and Williams (1999). Observations with missing wage information are included in the tenure distribution analysis. One can distinguish four types of job spells. Apart from the distinction between completed and incomplete spells (right censoring), one can also make a distinction between spells that start before the time span covered by our data and spells that start afterwards. The lower half of the summary statistics table informs on the number of spells for each of these four types.

4.2 The Parameters of the Tenure Distribution

Estimation results for the tenure distribution analysis, see equation (16) and parameter specification (13), are presented in Table 2. Theoretically, the results for the "large sample", including the jobs ongoing at the start date of the PSID sample, and for the "small sample", excluding those jobs, should be identical. Both job exit hazards look indeed almost identical for the estimated mean values of $\ln\Omega$ and π, see again Figure 1, with the peak in the hazard somewhat lower for the large sample. Inspecting the estimates in Table 2 yields the same conclusion.

Table 2: SML estimates tenure distribution parameters equation (16)

	Small Sample$^{(1)}$		**Large Sample**$^{(2)}$	
Variable	π	**ln**Ω	π	lnΩ
Intercept	0.226**	-1.243 **	0.141**	-1.197**
(t-val)	(9.60)	(-14.22)	(83.96)	(-73.50)
Initial experience	0.0088**	-0.0057	0.012**	0.0025
(t-val)	(2.95)	(-0.57)	(53.57)	(1.38)
Random worker effects σ	0.309**	0.0022	5.76E-07	3.66E-05
(t-val)	(5.77)	(1.76E-03)	(3.34E-04)	(2.89E-03)
Observations (job spells)	3169		4681	

Significance levels: $\dagger$: 10% $*$: 5% $**$: 1% . Statistical t-values in parentheses under estimated coefficients.
$^{(1)}$Small sample= sample of job spells starting within the observation range of the PSID sample
$^{(2)}$Large sample= sample of all job spells, including those ongoing at the PSID sample start
All covariates are taken in deviations from their means over jobs

The intercepts and the coefficients for experience at job start are very similar for both samples. The positive effect of experience on the drift is consistent with the idea that workers start their career with some initial job hopping, before settling down in a job that fits their comparative advantages best. The intercept for π is positive and large for both samples. In both cases, there are hardly observations for which π is negative. This implies that a fraction of the job spells will last until the retirement of the worker. The fraction of jobs that never end is about 10%, computed for the mean values of $\ln \Omega$ and π.

An interesting observation is that while we find unobserved random worker effects for the small sample, we do not find any for the large sample. The same result is found for a slightly different specification of the model, see Teulings and Van der Ende (2000), or for slightly different samples or specifications that we estimated, but do not report in the paper (i.e., including unionized jobs, or including education as an explanatory variable). The main difference between the samples is that the large one contains more long job spells, since it includes ongoing spells at the start of the PSID extract. As pointed out earlier, estimates for the two samples should be identical; since this is not the case, we suspect some form of misspecification of the empirical model. Given that some prior studies have estimated considerable unobserved heterogeneity in the PSID, see Browning et al (2010), this misspecification must be related to the sample containing the long job spells. What form of misspecification of the hazard rate for long job durations might explain our result? Part of the jobs last till retirement. Since our model does not describe the outside option of leaving the market due to retirement, and we therefore exclude workers above 60, a share of the jobs will never end according to this model. This explains why we find the drift to be positive. As discussed in section 2.3, a positive drift implies the hazard rate to be steadily declining to 0. Heterogeneity in the drift would strenghten this decline. Due to selection, the sample of surviving job spells will be increasingly made up of workers with a high drift. Hence, the hazard rate will decline more rapidly when there is heterogeneity in the drift than when there is not. The data do not support this rapid decline in the hazard, hence the estimated zero variance of the random worker effect.

From this perspective, it is pertinent to compare our model with Jovanovic's (1979) Bayesian learning model, where the firm learns about the productivity of a match by subsequent realizations of its output. That model has a stochastic structure similar to the model considered here. New information about the quality of a match is orthogonal to the previously collected information. However, as times goes by, the quality of the firm's information set improves and new information will have a smaller impact. Hence, the hazard rate in learning models declines much faster than in our model. The data strongly reject this rapid decline. Our model comes a long way in explaining the slow decline in the hazard rate for long job spells. However, the absence of unobserved heterogeneity in the drift suggests that the actual decline is even slower than our model predicts.

Since the long spells started before the first wave of our PSID sample contain crucial information, we focus on the results for Ω and π obtained from the full sample of job spells for the subsequent wage dynamics analysis.

4.3 Wage Dynamics

The FGNLS estimation results of system (17) are reported in Table 3. The equations in (17) impose a linear experience profile. However, the model can be easily extended with a concave experience profile, since this affects r_t and p_t equally. We do so throughout the subsequent analysis. As stated in Section 3.3, since we make use of the earlier estimated tenure distribution parameters, we need to correct the variances of the FGNLS estimates for the error introduced by the SML estimation. This two-step correction in the spirit of, e.g., Murphy and Topel (1985), has absolutely no effects on the standard errors for any of our reported estimations (adjusted variance-covariances are identical to the unadjusted ones up to 10 decimal digits).

Estimates for a number of subsamples are reported in different horizontal panels. The first panel uses all available data. All coefficients are significant and have the expected sign. The coefficients on experience (t and t^2) point to a standard concave experience profile. The coefficient γ is estimated to be 0.792, relatively close to unity, implying that the correlation between Δp_t and Δb_t is low. Separations look more like quits: they are driven more by random positive shocks to the outside productivity than by negative shocks to the inside productivity. Hence, the correlation of $\Delta\Omega_t$ with Δw_t is low, leading to a high estimated value of γ. Part of the reason for the low correlation might be downward rigidity in wages; if so, the declining part of the wage profile for a complete spell, cf. Figure 2, will not be realized.

Table 3: FGNLS estimates system (17)

	μ_0	γ	$\bar{\sigma}^2$	${\sigma_u}^2$	${\sigma_z}^2$	t	t^2	Avg Nobs
				1: All Stayers+ Movers				
coef	0.069**	0.729**	0.0012†	0.0046**	0.011**	-0.0056**	9.80E-05**	4575
(t-val)	(12.11)	(4.25)	(1.75)	(14.90)	(14.74)	(-9.02)	(6.53)	
				2: Incomplete and Positive Completed Surplus Change Spells for Stayers + Movers				
coef	0.066**	0.512†	0.0014†	0.0046**	0.010**	-0.0057**	9.90E-05**	3957
(t-val)	(9.19)	(1.64)	(1.79)	(14.12)	(12.65)	(-8.83)	(6.17)	
				3: As panel 2 above, but using -Max($\Omega_\tau{}^*$) as regressor for job movers				
coef	0.067**	0.547*	0.0015†	0.0046**	0.010**	-0.0057**	9.90E-05**	3957
(t-val)	(9.66)	(2.03)	(1.87)	(14.12)	(12.81)	(-8.40)	(6.20)	

Significance levels: † : 10% $*$: 5% $**$: 1%. Statistical t-values in parentheses under estimated coefficients.
We allow for a concave experience (t, t^2) profile in all equations of system (17).

We investigate this issue by leaving out all observations for which $\Delta\Omega_\tau$ is negative, i.e. roughly the second half of all completed spells. This second set of estimates are reported in panel 2 of Table 3. They are virtually the same, except for γ, which is now estimated to be 0.512, though not statistically significant. The downward rigidity in wages implies a large fall in wages at the moment of separation. Hence, we further enter, with a negative sign, the maximum of Ω_τ in the previous job, $-\mathrm{Max}(\Omega_\tau^*)$, as regressor in the equation for job movers, instead of the decline in the surplus in the last year before separation, $\mathrm{E}\Delta\Omega_\Theta^*$. We expect its coefficient to be

$\overline{\gamma\sigma}$. The estimation results for this model are reported in panel 3. Once again, the estimation results are virtually the same as in panel 2 of the table, except that the standard errors of all coefficients become somewhat smaller; γ is also significant in this specification. The difference in the estimated value of γ in panels 1 and 3 suggests that downward rigidity plays indeed a role. Later on, we present a Wald test of this hypothesis.

The ranges of estimates obtained for the main parameters, $\gamma = \{0.51, 0.73\}$, and $\overline{\sigma} = 0.04$, enable us to compute the 'true' return to tenure, $\overline{\sigma}\pi = 0.04 \times 0.14 = 0.56\%$ (taking the estimated mean value of $\pi = 0.14$ from Table 2 above). However, the high values of γ imply that most of the return to tenure, between 50% and 75%, takes the form of the log reservation wage r_t having a negative drift, instead of the inside wage w_t having a positive drift, see equation (12). The return to tenure measured as the rise in log productivity in the current job p_t, is thus even smaller, between $\overline{\gamma\sigma}\pi = (1 - 0.73) \times 0.04 \times 0.14 = 0.15\%$ and $(1 - 0.51) \times 0.04 \times 0.14 = 0.27\%$. Apart from this true return to tenure (linear by assumption), there is also a return to tenure due to the selectivity in the evolution of b_t in surviving jobs. Complete spells yield a hump shaped pattern for Ω_τ, while incomplete spells yield an increasing concave pattern for Ω_τ, see Figure 2. When there is downward rigidity, the hump shape for complete spells is reduced to an increasing concave pattern, too. Hence, the concavity in the tenure profiles can be fully explained by selectivity.

In the course of the life cycle, workers with the same level of experience t end up with different elapsed tenure lengths, depending on their history of job mobility. The existence of a tenure profile in wages implies that these differences translate into wage inequality. Since the tenure profile can be decomposed in a deterministic part, $\gamma\overline{\sigma}\pi\Psi$, and a random part, $\overline{\gamma\sigma}\Omega_\tau$, one can ask what is the contribution of these two factors to expected wage growth and wage inequality. We address this issue using the estimated parameter values $\pi = 0.14$ and $\Omega = 0.30$, respectively $\gamma = 0.60$ (about the middle of the interval 0.51-0.73) and $\overline{\sigma} = 0.04$. We do this decomposition for $t =$

$10, 20, 30$ years of experience, using a recursive computation for identical persons, starting with 0 years of experience, and characterized by the tenure CDF given by (8). With respect to expected log wage growth, the contribution of the deterministic component $\gamma\overline{\sigma}\pi\mathrm{E}[\Psi|t]$ is 1.68%, 3.31%, and 4.94% for $t = 10, 20, 30$ respectively. The contribution of the expected value of the random component $\overline{\gamma\sigma}\mathrm{E}_{\Psi}\left[\mathrm{E}\left[\Omega_{\tau}|B\left(\Psi\right)\right]|t\right]$ is 3.55%, 5.39%, and 6.94% respectively. Hence, the random component has a larger effect on log wage growth, in particular for low levels of experience. At higher levels of experience, job change becomes an unlikely event anyway and hence the contribution of selectivity converges to a fixed number. With respect to wage inequality, the contribution of the deterministic component is $(\gamma\overline{\sigma}\pi)^2\mathrm{Var}[\Psi|t]$ and the stochastic component is $(\overline{\gamma\sigma})^2\mathrm{E}_{\Psi}\left[\mathrm{Var}\left[\Delta\Omega_{\tau}|B\left(\Psi\right)\right]|t\right]$. In this case, given the long experience lengths considered, both the deterministic and the random component have almost identical overall contributions to wage variance for these experience levels; however, the random component has a much larger contribution to the wage inequality, about 20 times larger, compared to the deterministic one. The deterministic components for 10, 20 and 30 years are 5.08e-06, 5.35e-06 and 5.37e-06, with the corresponding stochastic components at 1.11e-04, 1.20e-04 and respectively 1.21e-04.

The estimation results in Table 3 use all available information on first and second moments of wage changes, simultaneously. A specification test is to estimate the model separately for relevant subsets of the data, e.g. job stayers versus job movers, complete versus incomplete spells, or first versus second moments, and verify whether the coefficient estimates are the same. Before reporting formal nested hypotheses tests, linear regression estimates for such data subsets can already be informative. Table 4 displays OLS estimates for the first moments, i.e. the first two equations of system (17). The regressions in columns D and G, complete spells with an increasing surplus ($\Delta\Omega_{\tau} \geq 0$) and respectively job movers, are badly identified due to a low number of observations. The other columns reveal some common patterns. First, the intercept and the experience profile are virtually the same in all regressions. Second, the coefficient of $\gamma\overline{\sigma}$

is negative, though never statistically significant, while it is expected to be positive. This term captures the earlier result from the tenure distribution analysis that the drift in the surplus Ω_t depends positively on experience at job start S, equivalent to stating that jobs starting at later age last longer. Given that $\gamma > 0$, the model predicts that workers are able to capture part of the surplus increase, and hence, that the tenure profile in jobs starting at higher ages should be steeper. The data reject this implication.

Table 4: OLS estimates 1st moments (first 2 equations) of system (17)

	A	B	C	D	E	F	G
μ_0	0.059**	0.061**	0.043**	-0.021	0.050**	0.046**	-8.593
(t-val)	(9.00)	(4.12)	(3.34)	(-0.36)	(2.65)	(4.19)	(-0.97)
$\gamma\bar{\sigma}$	-0.013	-0.071	-0.054	-0.529	-0.011	-0.048	0.497
(t-val)	(-0.57)	(-0.10)	(-1.60)	(-1.51)	(-0.14)	(-1.56)	(1.07)
$(1\text{-}\gamma)^2\bar{\sigma}^2$	6.40E-05*	6.40E-05	9.61E-04*	0.0045†	4E-06	8.41E-04†	0.0031
(t-val)	(2.37)	(1.53)	(2.14)	(1.96)	(0.27)	(2.34)	(0.58)
$\bar{\sigma}^2$							857.43
(t-val)							(1.00)
t	-5.01E-03**	-5.25E-03**	-4.39E-03**	-2.92E-03	-4.64E-02*	-4.57E-03**	0.039
(t-val)	(-7.17)	(-3.00)	(-4.99)	(-0.06)	(-2.31)	(-5.58)	(0.76)
t^2	9.19E-05**	1.08E-04*	8.33E-05**	1.44E-04	9.44E-05†	8.62E-05**	-8.23E-05
(t-val)	(5.6)	(2.23)	(4.42)	(1.29)	(1.73)	(4.78)	(-0.32)
Nobs	8082	1572	6510	435	1137	6945	462

Significance levels: † : 10% $*$: 5% $**$: 1%. Statistical t-values in parentheses under estimated coefficients.
Columns A to F correspond to the 1st moment eq. for job stayers (1st eq. of system 17), where A- All Stayers, B- Completed Spells, C- Incomplete Spells, D- Completed Positive Surplus Change Spells, E- Completed Negative Surplus Change, F- Incomplete plus Completed Positive Surplus Change; Column G corresponds to the 1st moment eq for job movers (2nd eq of system 17).
We allow for a concave experience (t, t^2) profile in each estimated equation.

Table 5 presents linear regression results for the 2nd moments, i.e. the last three equations of system (17). The estimates invoke three observations. First, contrary to the predictions of our model, the estimates for $(\overline{\gamma\sigma})^2$ are negative. The model predicts a hump shape in the variance of Δw_t over the course of a job spell, with low variances in the beginning and the end of a job. The data tell the opposite. Thus, while the model accurately captures the concavity in the tenure profile in the first moment of Δw_t, in particular when accounting for downward rigidity in wages, it does not capture the pattern in its second moment. Second, the variance $\sigma_z^2 + \sigma_u^2$

is a factor four times higher for job movers than for job stayers. This suggest that the labour market is not a Walrasian market with a continuum of outside offers available at any time, where workers who want to change jobs can just the pick the best option out of this continuum. Outside offers come along randomly, so that there are large jumps in the wage profile at the moment of job change. Third, within the group of job stayers, the variance does not seem to be constant across subgroups either; it is the largest for the incomplete spells and the smallest for the complete spells with a declining surplus ($\Delta\Omega_\tau < 0$), whereas the complete spells with an increasing surplus fall somewhere in between. The low variance for complete spells with a declining surplus fits the notion of downward wage rigidity. When wages are rigid, one would not expect a whole lot of variance.

Table 5: OLS estimates 2nd moments (last 3 equations) of system (17)

	A	B	C	D	E	F	G	H	I
$(1-\gamma)^2\bar{\sigma}^2$	-0.0081**	-0.010†	-0.025**	-0.016	-0.010*	-0.019**	-0.089		
(t-val)	(-3.28)	(-1.88)	(-4.55)	(-1.16)	(-1.82)	(-4.42)	(-0.53)		
${\sigma_z}^2+2{\sigma_u}^2$	0.027**	0.025**	0.042**	0.032**	0.025**	0.036**	0.156**		
(t-val)	(12.28)	(8.33)	(8.49)	(3.51)	(7.73)	(9.60)	(2.98)		
${\sigma_u}^2$								0.0042**	0.0043**
(t-val)								(14.01)	(13.28)
Nobs	8082	1572	6510	435	1137	6945	462	5789	4972

Significance levels: † : 10% * : 5% ** : 1%. Statistical t-values in parentheses under estimated coefficients.
Columns A to F correspond to the 2nd moment for job stayers (3rd eq. of system 17), where A- All Stayers, B- Completed Spells, C- Incomplete Spells, D- Completed Positive Surplus Change Spells, E- Completed Negative Surplus Change, F- Incomplete plus Completed Positive Surplus Change; Column G corresponds to the 2nd moment for job movers (4th eq. of system 17); Columns H to I correspond to the covariance moment (last eq. of system 17), with H using All Stayers and I using the Incomplete plus Positive Completed Surplus Change Spells.
We use the residuals computed from corresponding 1st moments of system (17), upgraded with concave experience profiles, as dependent variables in all 2nd moment equations.

Table 6 presents formal Wald tests for the above hypotheses, reporting χ^2 statistics and associated p-values for tests of equality of estimates across nested model specifications for the system (17). All tests start from the full model (using the whole number of observations, on both stayers and movers), except for horizontal panel 6 which presents nested hypotheses tests for the subsample of job stayers. Panel 1 presents three Wald tests for the null hypotheses

$\{\gamma_s = \gamma_m\}$, $\{\overline{\sigma}_s^2 = \overline{\sigma}_m^2\}$, and the joint null $\{\gamma_s = \gamma_m$ and $\overline{\sigma}_s^2 = \overline{\sigma}_m^2\}$, where s indexes stayers and m movers. The null $\{\gamma_s = \gamma_m\}$ cannot be rejected, but the null $\{\overline{\sigma}_s^2 = \overline{\sigma}_m^2\}$ can. Hence, the joint null $\{\gamma_s = \gamma_m$ and $\overline{\sigma}_s^2 = \overline{\sigma}_m^2\}$ is also rejected. This suggests that there is excess wage variance for job movers. Hence, our assumption of a Walrasian market for outside job offers is not respected in the data. Panel 2 presents a Wald test for $\{\gamma_{neg} = \gamma_{rest}\}$, where γ_{neg} is estimated only for the completed job spells with a negative surplus change, $\Delta\Omega_\tau < 0$, while γ_{rest} is estimated on the rest of the sample. We fail to reject the null of γ being the same for the negative job spells and for the rest of the observations.

Table 6: Nested hypotheses tests for subset estimates system (17)

1: γ_s=γ_m, $\bar{\sigma}_s{}^2$=$\bar{\sigma}_m{}^2$	
Wald γ_s=γ_m	χ^2=0.37 (Prob>χ^2=0.54)
Wald $\bar{\sigma}_s{}^2$=$\bar{\sigma}_m{}^2$	χ^2=5.60*(Prob>χ^2=0.018)
Joint Wald γ_s=γ_m and $\bar{\sigma}_s{}^2$=$\bar{\sigma}_m{}^2$	χ^2=5.80†(Prob>χ^2=0.055)
2: γ_{neg}=γ_{rest}	
Wald γ_{neg}=γ_{rest}	χ^2=1.64 (Prob>χ^2=0.20)
3: γ_{drift}=$\gamma_{nondrift}$	
Wald γ_{drift}=$\gamma_{nondrift}$	χ^2=31.43**(Prob>χ^2<.0001)
4: γ_{first}=γ_{second}	
Wald γ_{first}=γ_{second}	χ^2=0.45 (Prob>χ^2=0.50)
5: $\sigma_{z,s}{}^2$=$\sigma_{z,m}{}^2$	
Wald $\sigma_{z,s}{}^2$=$\sigma_{z,m}{}^2$	χ^2=37.45**(Prob>χ^2<.0001)
6: $\sigma_{z,s,inc}{}^2$=$\sigma_{z,s,neg}{}^2$, $\sigma_{z,s,inc}{}^2$=$\sigma_{z,s,pos}{}^2$, $\sigma_{z,s,neg}{}^2$=$\sigma_{z,s,pos}{}^2$	
Wald $\sigma_{z,s,inc}{}^2$=$\sigma_{z,s,neg}{}^2$	χ^2=2.08 (Prob>χ^2=0.149)
Wald $\sigma_{z,s,inc}{}^2$=$\sigma_{z,s,pos}{}^2$	χ^2=0.04 (Prob>χ^2=0.833)
Wald $\sigma_{z,s,neg}{}^2$=$\sigma_{z,s,pos}{}^2$	χ^2=1.82 (Prob>χ^2=0.177)
Joint Wald $\sigma_{z,s,inc}{}^2$=$\sigma_{z,s,neg}{}^2$ and $\sigma_{z,s,inc}{}^2$=$\sigma_{z,s,pos}{}^2$	χ^2=2.17 (Prob>χ^2=0.338)

Significance levels: $\dagger$: 10% $*$: 5% $**$: 1%. Statistical p-values in parentheses.
Detailed description for each hypothesis test can be found in the text; "m" indexes movers, "s" stayers, "inc" incomplete job spells, "pos" ("neg") completed positive (negative) surplus change job spells. The test in panel 4 is de facto implemented as H_0:$k_1{}^2$=k_2, where k_1=1-γ_{first} and k_2=(1-γ_{second})^2, since, as k_2 is estimated negative in the corresponding linear regression, γ_{second} cannot take a real value in that particular specification.
All specifications allow for concave experience (t, t^2) profiles.

Panel 3 presents the Wald test of $\{\gamma_{drift} = \gamma_{nondrift}\}$, where γ_{drift} is estimated from the drift term in the first moment equations, while $\gamma_{nondrift}$ is estimated from all other instances where it appears in the system (17). As expected from earlier remarks, we strongly reject this null. Panel 4 presents the Wald test of $\{\gamma_{first} = \gamma_{second}\}$. This restriction cannot be rejected. Panel

5 displays the result of the Wald for $\{\sigma^2_{z,s} = \sigma^2_{z,m}\}$, i.e. testing for equality of the variance of the permanent shocks, σ^2_z, across movers and stayers. This null is clearly rejected. Given our result from panel 5, we now start from a model where we allow for different $\sigma^2_{z,s}$ and $\sigma^2_{z,m}$: in panel 6 we test in subsamples of job stayers only the following null hypotheses: $\{\sigma^2_{z,s,inc} = \sigma^2_{z,s,neg}\}$, $\{\sigma^2_{z,s,pos} = \sigma^2_{z,s,neg}\}$, $\{\sigma^2_{z,s,inc} = \sigma^2_{z,s,pos}\}$, and respectively the joint $\{\sigma^2_{z,s,inc} = \sigma^2_{z,s,neg}$ and $\sigma^2_{z,s,inc} = \sigma^2_{z,s,pos}\}$, where *inc* indexes incomplete job spells, *neg* completed job spells with $\Delta\Omega_\tau < 0$, as above, and *pos* completed job spells with $\Delta\Omega_\tau \geq 0$ These tests show that once we account for the differences in σ^2_z between movers and stayers, there are no further statistical differences between the estimates for subsamples of stayers: indeed, we cannot reject any of the null hypotheses from panel 6 of table 6.

Table 7: FGNLS estimates system (17), with different wage variances for stayers and movers

	μ_0	γ	$\bar{\sigma}_s{}^2$	$\bar{\sigma}_m{}^2$	$\sigma_u{}^2$	$\sigma_z{}^2$	t	t^2	Avg Nobs
1: All Stayers+ Movers									
coef	0.071^{**}	0.812^{**}	0.0018^{**}	0.301^{**}	0.0046^{**}	0.011^{**}	-0.0057^{**}	9.9E-05**	4575
(t-val)	(13.94)	(36.61)	(3.17)	(2.78)	(14.90)	(14.82)	(-9.48)	(6.66)	
2: Incomplete and Positive Completed Surplus Change Spells for Stayers + Movers									
coef	0.073^{**}	0.811^{**}	0.0024^{**}	0.291^{**}	0.0046^{**}	0.010^{**}	-5.96E-03**	1.03E-04**	3957
(t-val)	(13.39)	(35.83)	(2.54)	(2.73)	(14.13)	(13.59)	(-9.18)	(6.52)	
3: As panel 2 above, but using -Max($\Omega_\tau{}^*$) as regressor for job movers									
coef	0.074^{**}	0.852^{**}	0.0021^{*}	0.150^{**}	0.0046^{**}	0.010^{**}	-0.0059^{**}	1.03E-04**	3957
(t-val)	(13.35)	(35.94)	(2.28)	(3.15)	(14.13)	(13.69)	(-9.11)	(6.50)	

Significance levels: † : 10% $*$: 5% $**$: 1%. Statistical t-values in parentheses under estimated coefficients.
We allow for a concave experience (t, t^2) profile in all equations of system (17).

Once we allow movers and stayers to possibly have different wage variances, the new FGNLS estimates reported in Table 7 show that movers have indeed a much higher wage variance at job separation; our favorite specification in panel 3 suggests $\overline{\sigma}_m \simeq 8.4\overline{\sigma}_s$. In this case, γ is estimated in a much narrower range, between 0.81 and 0.85, with all the other parameter estimates close to the values from Table 3. We repeat the previous calculation of the return to tenure, this time with $\overline{\sigma}_s = 0.04$, and $\gamma = 0.8$. We obtain a return which is exactly the same as computed before, $\overline{\sigma}_s\pi = 0.04 \times 0.14 = 0.56\%$, again mostly due to the fall in the outside option. The drift in p_t accounts now for only $\overline{\gamma\sigma}_s\pi = (1 - 0.8) \times 0.56\% = 0.11\%$. Hence, once we adjust our empirical model to allow for downward rigidity of wages and for the non-Walrasian market for job switches, about 80% of the wage returns to tenure is due to selectivity on the outside wages.

A Conditional Expectation and Variance of Ω_τ

A.1 Completed Spells

For the subsequent derivations, it is useful to add the parameter for initial surplus, Ω, as an argument to the survival function of job tenures in equations (8) and (9), thus $\overline{F}(\tau, \Omega)$ and $f(\tau, \Omega)$. Let $h(\omega, \tau, \Theta, \Omega)$ be the density of $\Omega_\tau = \omega$ for $0 < \tau < \Theta$ conditional on $A(\Theta)$. Comparing this density to $g(\omega, \tau)$, there is one additional condition: $\Omega_\Theta = 0$. Hence, $h(\omega, \tau, \Theta, \Omega)$ can be calculated by applying Bayes's rule. Since Ω_τ is a martingale, the distribution of Θ conditional on $\Omega_\tau = \omega$ is equal to the distribution of $\Theta - \tau$ conditional on $\Omega_0 = \omega$. Hence, its density is $f(\Theta - \tau, \omega)$. Then $h(\omega, \tau, \Theta, \Omega)$ can be calculated from $f(\cdot)$ and $g(\cdot)$, by Bayes's rule:

$$h(\omega, \tau, \Theta, \Omega) = \frac{f(\Theta - \tau, \omega) g(\omega, \tau)}{\int_0^\infty f(\Theta - \tau, x) g(x, \tau)\, dx}$$

Substitution of equation (7) in the above yields:

$$h(\omega, \tau, \Theta, \Omega) = \frac{\omega}{\Omega m \sqrt{m\tau}} \left[\phi\left(\frac{\omega - m\Omega}{\sqrt{m\tau}}\right) - \phi\left(\frac{\omega + m\Omega}{\sqrt{m\tau}}\right) \right]$$

where $m \equiv (\Theta - \tau)/\Theta$. Hence, $\mathrm{E}(\Omega_\tau | A(\Theta))$ satisfies:

$$
\begin{aligned}
\mathrm{E}(\Omega_\tau|A(\Theta)) &= \int_0^\infty \omega h(\omega,\tau,\Theta,\Omega)d\omega \\
&= \int_0^\infty \frac{\omega^2}{\Omega m\sqrt{m\tau}}\left[\phi\left(\frac{\omega-m\Omega}{\sqrt{m\tau}}\right)-\phi\left(\frac{\omega+m\Omega}{\sqrt{m\tau}}\right)\right]d\omega \\
&= 2\sqrt{m\tau}\phi\left(\sqrt{\frac{m}{\tau}}\Omega\right)-\left(\frac{\tau}{\Omega}+m\Omega\right)\left[1-2\Phi\left(\sqrt{\frac{m}{\tau}}\Omega\right)\right]
\end{aligned}
$$

For the calculation of the second moment of a first differential of Ω_τ, $\mathrm{E}\left[\Delta\Omega_\tau^2|A(\Theta)\right]$, we apply the joint density of $\Omega_{\tau-1}=\omega$ and $\Omega_\tau=\omega+\chi$, for $1\leq\tau\leq\Theta$, conditional on $A(\Theta)$:

$$
h(\omega+\chi,1,\Theta-\tau+1,\omega)\cdot h(\omega,\tau-1,\Theta,\Omega)
$$

The second moment of $\Delta\Omega_\tau$ is thus given by:

$$
\mathrm{E}\left[\Delta\Omega_\tau^2|A(\Theta)\right]=\int_0^\infty\int_{-\omega}^\infty \chi^2 h(\omega+\chi,1,\Theta-\tau+1,\omega)d\chi\cdot h(\omega,\tau-1,\Theta,\Omega)d\omega
$$

We use numerical integration for the evaluation of the integral above. The variance is subsequently derived by the standard expression $\mathrm{Var}[\Delta\Omega_\tau|A(\Theta)]=\mathrm{E}\left[\Delta\Omega_\tau^2|A(\Theta)\right]-\mathrm{E}[\Delta\Omega_\tau|A(\Theta)]^2$.

A.2 Incomplete Spells

Let $h^*(\omega,\tau,\Psi,\Omega)$ be the density of $\Omega_\tau=\omega$ conditional on $B(\Psi)$. Application of the Bayes rule yields:

$$
h^*(\omega,\tau,\Psi,\Omega)=\frac{\overline{F}(\Psi-\tau,\omega)g(\omega,\tau)}{\int_0^\infty\overline{F}(\Psi-\tau,x)g(x,\tau)\,dx}
$$

Hence, $\mathrm{E}(\Omega_\tau|B(\Psi))$ satisfies:

$$
\mathrm{E}[\Omega_\tau|B(\Psi)]=\int_0^\infty\omega h^*(\omega,\tau,\Psi,\Omega)d\omega=\frac{\int_0^\infty\omega\overline{F}(\Psi-\tau,\omega)g(\omega,\tau)\,d\omega}{\int_0^\infty\overline{F}(\Psi-\tau,\omega)g(\omega,\tau)\,d\omega}
$$

where $\overline{F}(\Psi - \tau, \omega)$ is given by equation (8). This expression is evaluated numerically, since it does not have an analytical solution.

The variance of $\Delta\Omega_\tau = \Omega_\tau - \Omega_{\tau-1}$, for $1 \leq \tau \leq \Psi$, conditional on $B(\Psi)$ is then derived from the first and second moments of $\Delta\Omega_\tau$, analogous to the completed spells case discussed above.

Acknowledgements

We thank our editors Keisuke Hirano and Shakeeb Khan, an anonymous associate editor, two anonymous referees, Rob Alessie, Gadi Barlevi, Richard Blundell, Jeff Campbell, Bas van der Klaauw, Jim Heckman, Dale Mortensen, Derek Neal, Randall Olsen, Aureo de Paula, Miguel Portela, Jean Marc Robin, Rob Shimer, Chris Taber, Aico van Vuuren, Robert Waldmann, and audiences in a large number of seminars and conferences. We also acknowledge Nick Williams and Lennart Janssens for assistance in handling the data. The usual disclaimers apply. An online folder containing a web appendix with supplementary details on the estimations, the data, and our estimation programs is accessible at www.sebastianbuhai.com/tenureprofiles. Buhai gratefully acknowledges partial funding for this project through Marie Curie IOF grant PIOF-GA-2009-255597, awarded by the European Commission.

References

Abowd, J.M. and D. Card (1989), "On the covariance of earnings and hours changes", *Econometrica* 57(2), 411-45

Abowd, J.M., F. Kramarz and D. Margolis (1999), "High wage workers and high wage firms", *Econometrica* 67(2), 251-333

Abraham, K.G. and H.S. Farber (1987), "Job duration, seniority, and earnings", *American Economic Review* 7(3), 278-97

Altonji, J.G. and N. Williams (1997), "Do wages rise with job seniority? A reassessment.", NBER Working Paper No. 6010

Altonji, J.G. and N. Williams (1999), "The effects of labor market experience, job seniority and job mobility on wage growth". in *Research in Labor Economics*, vol. 17, edited by S.

Polacheck, JAI Press, Elsevier

Altonji, J. G. and N. Williams (2005), "Do wages rise with job seniority? A reassessment", *Industrial and Labor Relations Review*, 58 (3), 370-97.

Beaudry, P. and J. DiNardo (1991), "The effect of implicit contracts on the movement of wages over the business cycle: Evidence from micro data", *Journal of Political Economy* 99(4), 665-88

Browning, M., M. Ejrnaes and J. Alvarez (2010), "Modeling income processes with lots of heterogeneity", *Review of Economic Studies*, 77(4), 1353-81

Buchinsky, M., D. Fougere, F. Kramarz and R. Tchernis (2010), "Interfirm mobility, wages, and the returns to seniority and experience in the US", *Review of Economic Studies*, 77, 972-1001

Burdett, K. and D.T. Mortensen (1998), "Wage differentials, employer size, and unemployment", *International Economic Review* 39, 257-273

Dixit, A.K. (1989), "Entry and exit decisions under uncertainty", *Journal of Political Economy*, 97, 620-38

Dixit, A.K. and R.S. Pindyck (1994), *Investment under uncertainty*, Princeton University Press, Princeton, USA

Dustmann, C. and C. Meghir (2005), "Wages, experience and seniority", *Review of Economic Studies* 72, 77-108

Farber, H.S. (1994), "The analysis of interfirm mobility", *Journal of Labor Economics* 12(4), 554-93

Farber, H.S. (1999), "Mobility and strategy: The dynamics of job change in labor markets", in O. Ashenfelter and D. Card, eds., *Handbook of Labour Economics*, Vol 3-B, Amsterdam, North-Holland

Feller, W. (1968), *An introduction to probability theory and its applications*, vol 2, Wiley Series in Probability and Mathematical Statistics, J. Wiley and Sons, Inc.

Horowitz, J.L. and S. Lee (2002), "Semiparametric estimation of a panel data proportional hazards model with fixed effects", Chapter 3, PhD Dissertation Simon Lee, University of Iowa, 2002

Jacobson, L.S., R.J. LaLonde and D.G. Sullivan (1993), "Earnings losses of displaced workers", *American Economic Review* 84(4), 685-709

Jovanovic, B., "Job matching and the theory of turnover", Journal of Political Economy, 87(5), 972-90

Kijima, M. (2003), *Stochastic processes with applications to finance*, Boca Raton, FL, US: Chapman and Hall

Lancaster, T. (1972), "A stochastic model for the duration of a strike", *Journal of the Royal Statistical Society A*, 135(2), 257-71

Low, H., C. Meghir and L. Pistaferri (2010): "Wage risk and employment risk over the life cycle", *American Economic Review*, 100(4), 1432-67

MaCurdy, T.E. (1982), "The use of time-series processes to model the error structure of earnings in a longitudinal data analysis", *Journal of Econometrics* 18, 83-114

McLaughlin, K. J.(1991), "A theory of quits and layoffs with efficient turnover," *Journal of Political Economy*, 99(1), 1-29

Meghir, C. and L. Pistaferri (2004), "Income variance dynamics and heterogeneity", *Econometrica*, 72(1), 1-32

Murphy, K.M. and R.H. Topel (1985), "Estimation and inference in two-step econometric models",*Journal of Business and Economic Statistics*, 3(4), 370-79

Postel-Vinay, F. and H. Turon (2010), "On-the-job-search, productivity shocks, and the individual earnings process", *International Economic Review*, 51(3), 599-629

Teulings, C.N. and M.A van der Ende (2000), "A structural model of tenure and specific investments", Tinbergen Institute Discussion Paper 2000-9/3

Topel, R.H. (1991), "Specific capital, mobility and wages: wages rise with job seniority", *Journal of Political Economy*, 99(1), 145-76.

Topel, R.H. and M.P. Ward (1992), "Job mobility and the careers of young men", *Quarterly Journal of Economics*, 107(2), 439-79

Van der Ende, M. (1997), *On the values of jobs and specific training*, PhD Dissertation at the University of Amsterdam and Tinbergen Institute

Whitmore, G.A. (1979), "An inverse Gaussian model for labour turnover", *Journal of the Royal Statistical Society: A*, 142(4), 468-78