Tenure Profiles and Efficient Separation in a Stochastic Productivity Model*

Sebastian Buhai† and Coen Teulings‡

May 3, 2008

Abstract

This paper provides a new way of analyzing tenure profiles in wages, by modelling simultaneously the evolution of wages and the distribution of tenures. We develop a theoretical model based on efficient bargaining, where both log outside wage and log wage in the current job follow a random walk, as found empirically. This setting allows the application of real option theory. We derive the efficient separation rule. The model fits the observed distribution of job tenures well. Since we observe outside wages only at job start and job separation, our empirical analysis of within job wage growth is based on expected wage growth conditional on the outside wages at both dates. The model is estimated on the PSID.

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

JEL codes: C51, C52, J63

*We thank two anonymous referees, our editor Bernard Salanié, Rob Alessie, Gadi Barlevi, Richard Blundell, Jeff Campbell, Bas van der Klaauw, Jim Heckman, Dale Mortensen, Derek Neal, Randall Olsen, Miguel Portela, Robert Shimer, Chris Taber, Aico van Vuuren, Robert Waldmann, and participants in seminars at the Univ. of Chicago, Northwestern Univ., Federal Reserve Bank of Chicago, London School of Economics, Univ. College London, Univ. of Copenhagen, and Tinbergen Institute, for helpful comments. We also thank Nick Williams and Lennart Janssens for assistance in handling the data. The usual disclaimers apply.

†Aarhus School of Business, University of Aarhus and Tinbergen Institute, Erasmus University Rotterdam; sebu@asb.dk, buhai@tinbergen.nl; www.sebastianbuhai.com

‡CPB Netherlands Bureau for Economic Policy Analysis, University of Amsterdam and Tinbergen Institute; C.N.Teulings@cpb.nl

1 Introduction

A large empirical literature has looked at wage returns to job seniority, using a whole arsenal of econometric techniques, 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): while some authors find that large estimated returns are spurious and wage returns to tenure are actually very small, e.g. Altonji and Shakotko (1987), Abraham and Farber (1987), Altonji and Williams (1997, 2005), Abowd et al (1999), others confirm large and significant wage returns close to cross-section estimates, e.g. Topel (1991), Dustmann and Meghir (2005), Buchinsky et al (2005). Here we provide a new direction for investigating the wage-tenure relationship. From a theoretical point of view, large "true" returns to tenure are problematic. Were there really large returns, the worker-firm match would spoil large gains from trade at the moment of separation. Why would a worker separate when he loses his 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 not renegotiate the wage instead of separating? Although some models, such as efficiency wage models, can explain why this renegotiation process might not be fully efficient, the size of the wage returns to seniority reported in some papers remains puzzling. In fact, the empirical evidence offers support for at least some form of renegotiation. For instance Jacobson, LaLonde and Sullivan (1993) have shown that displaced workers face severe wage cuts of up to 25% just before separation. This paper addresses explicitly whether the existing evidence is consistent with efficient separations by modelling simultaneously the evolution of wages and the distribution of job tenures.

We take efficient bargaining as benchmark. Hence, quits and job layoffs are observationally equivalent, as in McLaughlin (1991). The model explains the correlation between wages and job tenure from the random evolution of both the job's productivity and the outside option. Separation occurs when the value of the productivity in the job falls sufficiently compared to the productivity of the outside option. This outside option is the productivity in the best alternative job that is available at that point in time. We refer to productivity at the job and in the best alternative as the *inside* and *outside productivity* respectively. By some form of bargaining, wages at the job are a linear combination of the inside and the outside productivity. Then, wages and tenure are correlated because only jobs where inside productivity evolves favorably compared to the outside productivity survive. Hence, there is no such thing as "the" return to tenure in this model. In some jobs wages go up because the job's productivity value evolves favorably. In others wages go down for *mutatis mutandis* the same reason. However, the latter group is gradually eliminated from the stock of ongoing employment relations just because there are no options for mutually gainful renegotiation left and hence separation becomes efficient.

The evolution of an individual's within-job log wage is reasonably described by a random walk with transitory shocks, as previously found by Abowd and Card (1989), Topel (1991) and Topel and Ward (1992). We verify that hypothesis in our PSID sample. Whereas this observation received little attention among labor economists, we take it as cornerstone of our modelling. Both log in- and outside productivity are assumed to follow a random walk. Our model implies that log wages are a linear combination of both, which implies that log wages in the job follow a random walk as well. Hence, the difference in the drift between the log wage in the job and the log outside productivity is what we traditionally call "the return to tenure".

Starting a job requires an irreversible specific investment, which is lost upon separation. Hence, this investment has an option value. The combination of irreversibility and productivity following a random walk implies that we can apply the theory of real options, see for example Dixit (1989), Bentolila and Bertola (1990) and Dixit and Pindyck (1994), compare Teulings and van der Ende (2000). The predicted hazard rates of this model are well in line with the empirical distribution of the job exits. Our model is similar to Mortensen's (1988) dual "on-the-job-training and matching" model. While we focus on firm tenure, our model could equally well be applied to industry or occupation tenure, as suggested by Neal (1995).

From the distribution of job tenures we are able to estimate the surplus of the job's productivity above its reservation value and a (linear) drift of this surplus, up to a normalizing constant (the variance of the random walk). We obtain a positive drift surplus, indicating that some 10% of all jobs will end only by retirement. We use these parameters to compute the evolution of the expected surplus in both complete and incomplete job spells, which will enable us to estimate its impact on wages. The typical problem in this literature is that the researcher observes the outside productivity only at job start and at job separation, assuming that the worker has a new job immediately afterwards. At job start, the worker chooses the best alternative that is available at that moment, which is by definition equal to the outside productivity. Our estimation procedure exploits both pieces of information on the outside productivity. To that end, we apply an idea first explored by Abraham and Farber (1987), conditioning the expected wage growth on both the current and the remaining tenure at that job. We can calculate a closed form expression for this expectation. As a first result, we show that this expression does not depend on the drift surplus. This implies that the evolution of wages in completed spells is uninformative on the return to tenure. This is a remarkable conclusion given the fact that so many papers have tried to identify the return to tenure from this type of data. The only sources of information on the return to tenure are the distribution of completed tenures and the evolution of wages in incomplete job spells. The fat right tail in the tenure distribution, with many jobs never ending, is an indication of large returns to tenure: the return to tenure is so high that separation is rarely efficient.

Secondly, we show that our model can explain the observed concavity in the tenure profile. Since the "true" tenure profile, the drift in the difference between inside and outside productivity, is linear by assumption, this concavity is fully due to selection. One could argue that our identification procedure relies heavily on functional form assumptions. However, there is one strong test of our assumptions: the estimated variance of the innovation in wages is sufficiently large for selectivity to generate the observed degree of concavity in the tenure profile.

Thirdly, we show that the problem in estimating the tenure profile in wages is not so much the selectivity in the inside productivity (and hence in the wage rate at the job), but in the outside productivity. Workers switch jobs only when the outside productivity is high. This source of selectivity usually receives less attention than the selectivity in the inside productivity. We show that this effect can be identified from the wage change for job movers. Surprisingly, selectivity in the outside wage turns out to be an empirically important phenomenon, accounting for 95% of the tenure profile, though this estimate is sensitive to mispecification of the model. In particular, our estimation results provide some indications of downward rigidity in wages, as discussed for example by Beaudry and DiNardo (1991), who find that within a job spell wages go up in the upturn, but do not go down in the downturn. However, in our estimation results, this gap is filled by an additional wage decline for job changers. This downward rigidity does not fit the efficient bargaining hypothesis. The estimated tenure profile is on the high end of the spectrum, almost 3% per year, though almost all of that takes the form of a declining outside productivity instead of a rising inside productivity. If we were to exclude this part of the profile, our estimates would be on the low end of the spectrum, 0.15% per year.

The paper is structured as follows: the model is discussed in Section 2, the empirical analysis in Section 3 and Section 4 concludes.

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 productivity P_t evolves according to a geometric Brownian with drift; P_t is job specific. 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 job. However, the firm retains the property right on the vacant job. Hence, the firm can hire a new worker for that job at any future time, provided that the cost of the specific investment is paid again. This cost of the specific 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. Since there are always vacant jobs available, a worker has a shadow price R_t, which is equal to the return in the best alternative vacant job, net of the cost of investment for that job. For the sake of convenience, we treat this shadow price as an exogenous variable here. Like P_t, it evolves according to a geometric Brownian with drift; since workers are homogeneous, R_t is common to all of them. Both workers and firms are perfectly informed about the current value of the P_t's for each job and of 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:

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

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.

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 surpluses $P_t - R_t$, $t > s$, 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 relative bargaining power. 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 identical, as in McLaughlin (1991). For the sake of convenience, we shall refer to separations as the firm firing the worker, though they can be both quits or layoffs. Given these assumptions, wage setting and separation decisions can be analyzed separately, since, in the spirit of the Coase theorem, hiring and firing decisions maximize the joint expected surplus, regardless of its precise distribution.

First, we analyze hiring and firing. Since hiring is in fact an irreversible investment, while firing is an irreversible disinvestment, both can be analysed using real option theory, see 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:

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

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, which state that the values before and after the hiring or firing should be equal. The first condition for hiring states that at the moment of hiring S the value of a job must be equal to the value of the vacancy plus the cost of specific investment. The second condition for firing states that at the moment of firing T the value of the job must be equal to the value of a vacancy. The last four conditions are the smooth pasting conditions. Value matching conditions impose value equality before and after hiring and firing; on top of that, smooth pasting conditions require that slight variations in the stochastic variables p_t and r_t should not affect the value equality, since hiring and firing decisions are irreversible. Hence, a decision maker should not regret her decision a minute later, due to slight variations in p_t or r_t. Smooth pasting requires thus the partial derivatives of the value matching condition with respect to p_t and r_t to be zero. These 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 of the relative surplus of P_t over R_t. By (1), we have:

$$\begin{aligned} b_t - b_s &\sim N\left[(t-s)\mu, (t-s)\sigma^2\right] \\ \sigma^2 &\equiv \sigma_p^2 + \sigma_r^2 - 2\sigma_{pr} \\ \mu &\equiv \mu_p - \mu_r \end{aligned} \tag{4}$$

Since p_t is specific for each job, so is b_t.

Proposition 1 *The value functions $J(\cdot)$ and $V(\cdot)$ can be written as:*

$$\begin{aligned} J(p_t, r_t) &= \exp(r_t)\, j(p_t - r_t) \\ V(p_t, r_t) &= \exp(r_t)\, v(p_t - r_t) \end{aligned} \tag{5}$$

where $j(\cdot)$ and $v(\cdot)$ satisfy:

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

where we leave out the argument of $j(\cdot)$ and $v(\cdot)$ for convenience. The value matching and smooth pasting conditions at the moment of job start and job separation read:

$$\begin{aligned} j\left(b^S\right) &= v\left(b^S\right) + I \\ v\left(b^T\right) &= j\left(b^T\right) \\ j'(b^S) &= v'(b^S) \\ v'(b^T) &= j'(b^T) \end{aligned}$$

where b^S, b^T are the values of b_t at the moment of hiring and firing respectively.

Proof: The proposition follows directly from substitution[1] of equation (5) in the Bellman equations (2) and the value matching and smooth pasting conditions (3).■

The smooth pasting conditions for p_t and r_t are identical, so we are left with only two independent smooth pasting conditions. 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 a rate ρ, but increase in expectation at a 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. This is exactly the "basic model" described by Dixit and Pindyck (1994; ch. 5.1-5.2), to whom we refer for the subsequent arguments. The solution to the two differential equations has four constants of integration. Two of these constants have to be zero due to a transversality condition. 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 $b^T < 0 < b^S$. Hiring occurs at the first moment that b_t rises to $b^S > 0$. Hence, $P_t > R_t$, 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 hiring later, while in the meantime b_t might fall below b^S again. Subsequent firing occurs at the first moment that b_t falls to b^T. Hence, $P_t < R_t$ because the firm accepts some losses before firing the worker, since it loses the option value of firing the worker

[1] We use:

$$\begin{aligned} J_p &= \exp(r_t)\, j', J_{pp} = \exp(r_t)\, j'' \\ J_r &= \exp(r_t)\,(j - j'), J_{rr} = \exp(r_t)\,(j - 2j' + j'') \\ J_{pr} &= \exp(r_t)\,(j' - j'') \end{aligned}$$

and likewise for $V(\cdot)$.

later.

2.3 Job Tenure Distribution

The next step is to analyze the distribution of job tenure in a job spell. 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. Analogously to the probit model, where the variance of the error term is non-identified because we observe only whether the indicator variable is positive or negative, 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 $\tau \equiv t - S, \tau \geq 0$, and $\Theta \equiv T - S, \Theta > 0$; τ is the incomplete tenure, while Θ is the completed tenure of that job spell. Define:

$$\begin{aligned} \Omega_\tau &\equiv \frac{b_t - b^T}{\sigma} \\ \Omega &\equiv \frac{b^S - b^T}{\sigma} > 0 \\ \pi &\equiv \frac{\mu}{\sigma} \end{aligned}$$

Thus Ω_τ is a Brownian with drift π and unit variance per unit time. By construction $\Omega_0 = \Omega$ and $\Omega_\Theta = 0$. 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_ζ has not been negative for any $\zeta \in [0, \tau]$. Hence, we are interested in the density of Ω_τ conditional on $\Omega_\zeta > 0, \forall \zeta \in [0, \tau]$. For this conditioning, we can apply the Reflection

Principle, first discussed by Feller (1968) for the case without drift, $\pi = 0$: 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)[2]:

$$g(\omega, \tau) = \frac{1}{\sqrt{\tau}} \left[\phi\left(\frac{\omega - \Omega - \pi\tau}{\sqrt{\tau}} \right) - e^{-2\Omega\pi} \phi\left(\frac{\omega + \Omega - \pi\tau}{\sqrt{\tau}} \right) \right] \tag{6}$$

where $\phi(.)$ is the standard normal density function. The first term in square brackets is the unconditional density; the second term 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 obtain the cumulative distribution of jobs surviving at τ, $\overline{F}(\tau) = \Pr(\Theta > \tau)$:

$$\overline{F}(\tau) \equiv \Phi\left(\frac{\Omega + \pi\tau}{\sqrt{\tau}} \right) - e^{-2\Omega\pi} \Phi\left(\frac{-\Omega + \pi\tau}{\sqrt{\tau}} \right) \tag{7}$$

where $\Phi(.)$ is the standard normal CDF. This expression above 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:

$$f(\tau) = \frac{\Omega}{\tau\sqrt{\tau}} \phi\left(\frac{\Omega + \pi\tau}{\sqrt{\tau}} \right) \tag{8}$$

[2]Kijima (2003, pp. 185-187) derives the precise expression of the transition density for our case, namely for a standard Brownian with drift $\pi > 0$, starting at $\Omega_0 = \Omega > 0$, and one absorbing barrier at $\Omega_\Theta = 0$. Many other books on stochastic processes derive the similar conditional density but for a standard driftless Brownian $\pi = 0$ and/or starting at $\Omega_0 = 0$ and/or with positive absorbing barrier $\Omega_\Theta = a > 0$. See for instance Cox and Miller (1968, pp. 219-223), Feller (1968, vol.2, p. 328), Zhang (1998, p. 208-218) etc. As shown by these authors, one can use various methods to derive the expression, the Reflection Principle being the most intuitive.

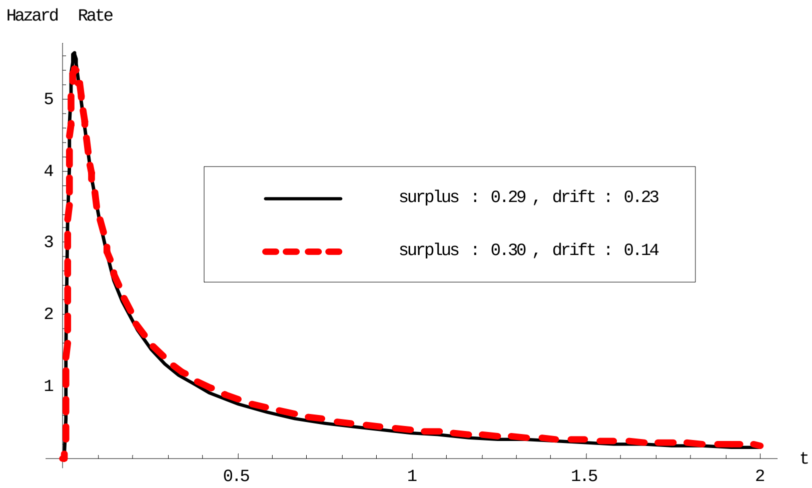

Figure 1: Predicted Job Hazards

where we have used $\phi\left(\frac{\Omega+\pi\tau}{\sqrt{\tau}}\right) = e^{-2\Omega\pi}\phi\left(\frac{-\Omega+\pi\tau}{\sqrt{\tau}}\right)$. The job exit rate is then given by $f(\tau)/\overline{F}(\tau)$. It is straightforward to check that the exit rate is hump shaped, starting from 0, reaching a peak at τ^*, $0 < \tau^* < \frac{2}{3}\Omega^2$, and afterwards either declining monotonically to 0 when the drift is positive $\pi > 0$ or to ${}^1/_2\pi^2$ when the drift is negative $\pi < 0$. Farber (1994), Teulings and Van der Ende (2000) and Horowitz and Lee (2002) have documented this hump shaped pattern using NLSY data. A positive drift implies a non exhaustive behavior, where some jobs never end. The fraction of surviving job spells for $\pi > 0$ is given by the value of the survivor function (7) for $\tau \to \infty$, hence by $1 - e^{-2\Omega\pi}$. In Figure 1, we plot the exit rates for $\Omega = e^{-1.20} \simeq 0.30, \pi = 0.14$ and respectively for $\Omega = e^{-1.24} \simeq 0.29, \pi = 0.23$, the mean estimated values for Ω and π, see Section 3, Table 2 below. 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.

2.4 Tenure Profile in Wages

2.4.1 Sharing Rule of Surpluses and Wages

We extend the 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 however a simpler approach, by imposing the log linearity of the sharing rule *a priori*, and deriving the intercept of that rule from the assumption of efficient bargaining.[3] According to this rule, the worker's log wage w_t satisfies:

$$w_t = r_t + \beta\left(b_t - b^T\right) + \psi = r_t + \overline{\sigma}\Omega_\tau + \psi \tag{9}$$

where $\overline{\sigma} \equiv \beta\sigma$. The parameter β can be interpreted as the worker's bargaining power. If $\beta = 0$, the worker receives a wage proportional to her shadow price R_t, while if $\beta = 1$, she receives a wage proportional to her productivity at the job, P_t. To close the model, the parameter ψ and the worker's share in the cost of investment remain to be determined. Since we do not need their expressions for the subsequent empirical analysis, we only provide an heuristic argument here. Let $Q\left(p_t, r_t\right)$ be the worker's asset value of holding a job, net of the discounted expected value of her shadow price R_t. Analogous to Proposition 1, $Q\left(\cdot\right)$ can be written as:

$$\begin{aligned} Q\left(p_t, r_t\right) &= \exp\left(r_t\right) q\left(p_t - r_t\right) \\ \left(\rho - \mu_r - \frac{1}{2}\sigma_r^2\right) q &= \exp\left[\beta\left(b_t - b^T\right) + \psi\right] - 1 + \left(\mu + \sigma_{pr} - \sigma_r^2\right) q' + \frac{1}{2}\sigma^2 q'' \end{aligned}$$

leaving out the argument of $q\left(\cdot\right)$ in the second line. The first term in the second line captures the wage at the job, compare equation (9), while the second term captures the outside wage. Note that we have divided both sides of the equation by $\exp\left(r_t\right)$. Efficient bargaining implies that it is optimal for the worker to quit when $b_t = b^T$. Hence,

[3] Nash bargaining would lead to a linear, instead of a log linear, sharing rule. Apart from this, the two approaches are identical. Note that Nash bargaining satisfies the assumption of efficient bargaining.

value matching and smooth pasting conditions must apply: $q\left(b^T\right)=0, q'\left(b^T\right)=0$. The value matching condition states that at the moment of separation the net asset value of continuation is zero. Again, the solution of the differential equation has two constants of integration, one of which has to be zero due to a transversality condition. Hence, the solution to the differential equation and the two boundary conditions determine the constant of integration and ψ. Finally, the assumption of verifiability of the cost of specific investment at the moment of job start implies that the worker's share in this cost must be equal to the net value of holding a job at the moment of job start, $\exp(r_S)\, q\left(b^S\right)$. Hence, this share is equal to $q\left(b^S\right)/I$.

2.4.2 Selectivity in Tenure Profiles

Equation (9) implies that log wages within a job follow a Brownian with drift $\mu_r + \overline{\sigma}\pi$; μ_r is the sum of the return to experience for that cohort of workers and the secular growth of real wages due to technological progress common to all cohorts; $\overline{\sigma}\pi$ 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. Empirically, wages are observed in discrete time. Hence, for estimation, the easiest way would be to first difference equation (9):

$$\begin{aligned} \text{job stayers} &: \quad \Delta \mathrm{E}\left(w_t | 1 < t - S < \Theta\right) = \mu_r + \overline{\sigma}\pi \\ \text{job changers} &: \quad \Delta \mathrm{E}\left(w_t^* | t = T\right) = \mu_r - \overline{\sigma}\pi\left(\Theta - 1\right) \end{aligned}$$

where Δ is the first difference operator and where the superscript $*$ indicates that we compare log wages in the new and the old job; hence, Δw_T^* compares the starting wage in the new job to the wage one year before separation in the old job. Hence, $\overline{\sigma}\pi$ can be estimated from data on job changers. However, in completed job spells, Ω_τ is correlated to Θ for three reasons: (i) $\Omega_0 = \Omega$, (ii) $\Omega_\Theta = 0$, and (iii) $\Omega_\zeta > 0, \forall \zeta, 0 \leq \zeta < \Theta$. For the sake of brevity, we refer to this information set as $A(\Theta)$. Our strategy for estimation is to calculate $\mathrm{E}(\Omega_\tau | A(\Theta))$ and to enter its first difference as a regressor in a regression on Δw_t's within completed job spells. *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) $\Theta > \Psi > \tau$, and hence (iii) $\Omega_\zeta > 0, \forall\zeta, 0 \leq \zeta \leq \Psi$. We refer to this second information set as $B(\Psi)$. Again, we can calculate $\mathrm{E}(\Omega_\tau | B(\Psi))$ and use its first difference as a regressor, see also Van der Ende (1997).

Proposition 2 *$E(\Omega_\tau | A(\Theta))$ and its derivatives satisfy:*

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

Proof See Appendix A1.■

This proposition implies that $\mathrm{E}(\Omega_\tau | A(\Theta))$ does not depend on the tenure profile in wages, $\overline{\sigma}\pi$. 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 staggering conclusion. The intuition for this result is that an increase in $\overline{\sigma}\pi$ has two offsetting effects on $\Delta\mathrm{E}(\Omega_\tau | A(\Theta))$. On the one hand, it raises the deterministic part of the tenure profile, so that the change in the unconditional expectation $\Delta\mathrm{E}(\Omega_\tau)$ goes up. On the other hand, it makes separation a less likely event, so that the condition $A(\Theta)$ becomes more selective: the non-deterministic part of $\Delta\Omega_\tau$ must have evolved unfavorably for a job spell to end after Θ, even though the deterministic part of the tenure profile pushes Ω_τ up. This conclusion depends crucially on the assumption of efficient bargaining. This assumption dictates

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

see equation (9). Hence, irrespective of the steepness of the tenure profile $\overline{\sigma}\pi$ or the length of the job spell Θ, log relative wages decline by $\overline{\sigma}\Omega$ over the duration of the spell. However, as noted in Section 2.3, 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 implies that jobs will survive longer. We return to the issue of identification in Section 3.

The third line 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 fourth line shows that the expected surplus declines infinitely fast just before separation. This result 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 line 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. The tenure profile is plotted for $\Omega = 0.30$ and for various values of Θ in Figure 2. For $\Theta \leq 0.1$ years the tenure profile is monotonically decreasing, while for larger Θ it is concave. The top of the profile is increasing in Θ, showing the importance of conditioning on the eventual tenure.

Contrary to the case of completed spells, there is no explicit expression for $\mathrm{E}(\Omega_\tau|B(\Psi))$. Hence, we use numerical integration, see Appendix A2. Figure 3 presents the trajectory of $\mathrm{E}(\Omega_\tau|B(\Psi))$ for $\Omega = 0.30, \pi = 0.14$ and various values of Ψ. $\mathrm{E}(\Omega_\tau|B(\Psi))$ is increasing in Ψ. The reason is that a higher value of Ψ provides more information on Θ since $\Theta > \Psi$. Hence, higher values of Ψ imply a greater selectivity. Were there no selectivity, then the trajectory would be linear, $\mathrm{E}(\Omega_\tau|B(\Psi)) = \mathrm{E}(\Omega_\tau) = \pi t$. The trajectories are strongly concave, implying that selection plays an important role. This offers an explanation of the

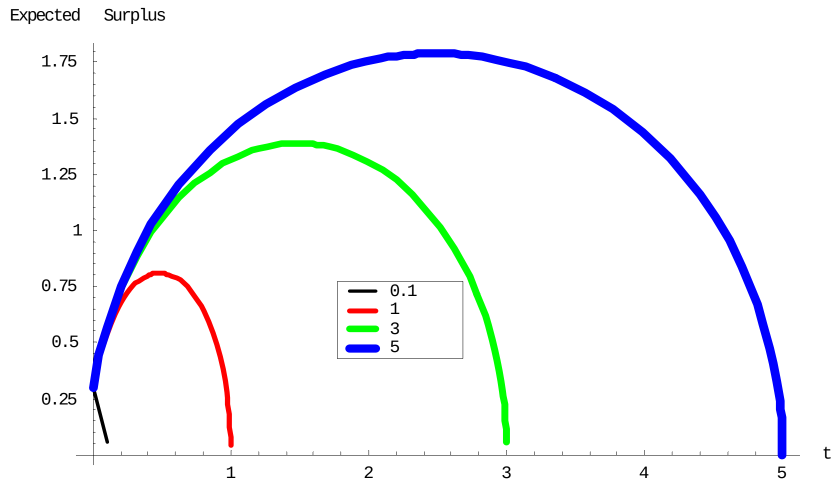

Figure 2: Expected Surplus in Completed Job Spells

observed concavity of tenure profiles in log wages: the underlying profile might be linear and the observed concavity might simply be due to selection. Contrary to the completed spells case, incomplete spells do provide information on the drift π. Nevertheless, the impact of the drift is negligible compared to that of selectivity, as documented by Figure 4, which compares the trajectories of $\mathrm{E}(\Omega_\tau|A(\Theta))$, $\mathrm{E}(\Omega_\tau|B(\Psi))$, and $\mathrm{E}(\Omega_\tau)$. The concavity outweighs the linear trajectory by far, at least for the first five years. In Figure 5 we plot $\mathrm{E}(\Omega_\tau|A(\Theta))$ and $\mathrm{E}(\Omega_\tau|B(\Psi))$ for long job durations, $\Theta = 10, 20$ and respectively $\Psi = 10, 20$.

2.4.3 Expected Within-Job and Between-Job Wage Growth

We apply the conditional expectations of Ω_τ for the empirical analysis of the growth in w_t and r_t. We observe r_t only at the moment of job change. Hence, whereas we can use information of within job wage growth for the analysis of w_t, we have to rely on between job wage growth for the analysis of r_t. For this purpose, we linearly decompose the random variables $[\Delta p_t, \Delta r_t]$ in two orthogonal components Δb_t and Δz_t, such that

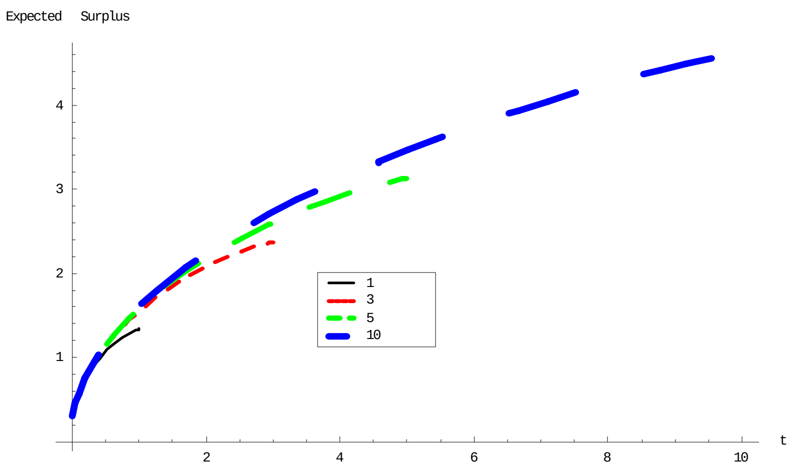

Figure 3: Expected Surplus in Incomplete Job Spells

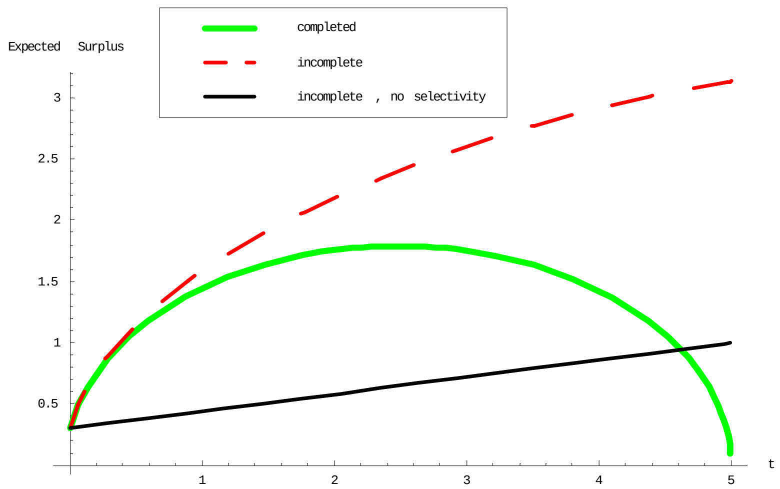

Figure 4: Selectivity versus Drift in the Expected Surplus

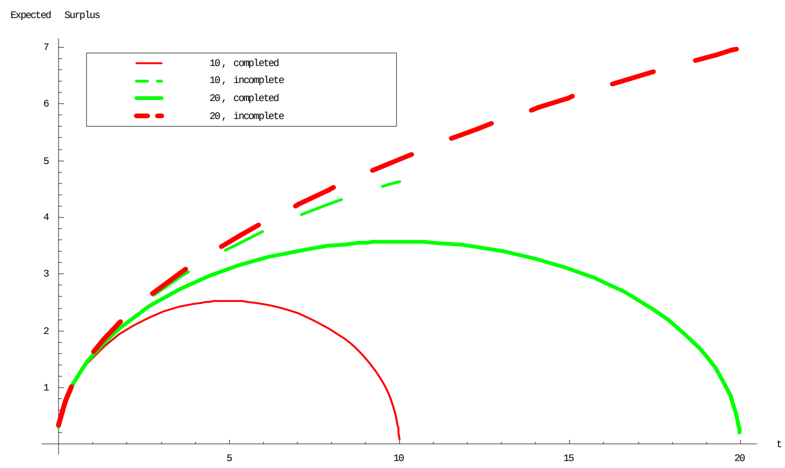

Figure 5: Expected Surplus in Long Spells

$\mathrm{Cor}(\Delta b_t, \Delta z_t) = 0$, and that the effect of Δz_t on Δr_t and Δw_t is equal to unity. Then:

$$\begin{aligned} \Delta r_t &= \Delta z_t - \gamma\beta\Delta b_t = \Delta z_t - \gamma\overline{\sigma}\Delta\Omega_\tau \\ \Delta w_t &= \Delta z_t + (1-\gamma)\beta\Delta b_t = \Delta z_t + (1-\gamma)\overline{\sigma}\Delta\Omega_\tau \end{aligned} \tag{10}$$

Given the joint normality of Δp_t and Δr_t, such a decomposition always exists. The advantage of this decomposition is that, 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. The parameter γ can be expressed in terms of the covariance matrix Σ and the bargaining power β, but that is of little help here. It is more useful to interpret it as a reflection of the correlation between the match surplus and the reservation wage. In the one extreme case where $\gamma = 0$, we can write $\Delta p_t = \Delta r_t + \Delta b_t$, with both right-hand side variables being 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 where $\gamma = 1$, we can write $\Delta r_t = \Delta p_t - \Delta b_t$, again with both right-hand side variables being uncorrelated. 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".

It is immediately clear from equation (10) that there is no hope of identifying β and σ separately from data on wages, since only their product $\overline{\sigma}$ shows up in the final expression for Δw_t. There is a clear intuition for this result. We do not observe the productivity p_t or the surplus b_t, but only the share that goes to workers as a wage payment w_t.[4]

Taking expectations in the second equation of (10) yields:

$$\mathrm{E}(\Delta w_t|A(\Theta)) = \mu_z + (1-\gamma)\overline{\sigma}\Delta \mathrm{E}(\Omega_\tau|A(\Theta)) \tag{11}$$

where $\mu_z \equiv \mathrm{E}(\Delta z_t)$. This equation applies for completed spells; replacing the condition $A(\Theta)$ by $B(\Psi)$ yields the equation for incomplete spells. As discussed in Section 2.3, Ω and π can be estimated from the distribution of completed tenures. These parameters are sufficient statistics for the calculation of $\mathrm{E}(\Omega_\tau|A(\Theta))$. Equation (11) allows the estimation of $(1-\gamma)\overline{\sigma}$ by OLS. However, γ and $\overline{\sigma}$ are not separately identified. The intuition is that we observe the current wage of a worker, but not her shadow price for an alternative job.

For job changers, we can write a similar equation:

$$\mathrm{E}(\Delta w_T^*) = \mu_z + (1-\gamma)\overline{\sigma}\Delta \mathrm{E}(\Omega_\Theta|A(\Theta)) + \overline{\sigma}\Omega \tag{12}$$

where, as before, the superscript $*$ indicates that we compare log wages in the new and the

[4]Teulings and Van der Ende (2000) work out a method that allows the estimation of σ. They interpret hours spend on training at the start of a job spell as a source of variation in I. The estimated covariation between I and Ω and the assumption of the absence of hold up problems allow the identification of σ.

old job. $\Delta E(\Omega_\Theta|A(\Theta))$ is the wage decline in the old job the year before separation.[5] It is always negative, see Figure 2. The term $\overline{\sigma}\Omega = \overline{\sigma}(\Omega_0 - \Omega_\Theta)$ reflects the wage increase due to entering the new job. The final term of equation (12) allows the separate identification of γ and $\overline{\sigma}$. The intuition is that at the moment a worker changes jobs, the wage jumps up by $\beta\left(b^S - b^T\right) = \overline{\sigma}\Omega$. This conclusion is very much similar to the standard result in this literature that information on job movers is crucial for the estimation of the tenure profile.

Till sofar we focused on the model's implications for the first moment of Δw_t. However, the model has also implications for the second moment of Δw_t. Taking the variance in the second equation of (10) yields:

$$\sigma_w^2 = \sigma_z^2 + (1-\gamma)^2\overline{\sigma}^2$$

where $\sigma_w^2 \equiv$Var(Δw_t) and $\sigma_z^2 \equiv$Var(Δz_t). Equation (9) implies:

$$w_T^* - w_S = r_T - r_S = z_T - z_S - \gamma\overline{\sigma}\Omega \tag{13}$$

where, as before, the superscript $*$ indicates that we refer to the starting wage in the new job starting at T, and where we use the first equation of (10), and $\Omega_0 = \Omega$ and $\Omega_\Theta = 0$ for the second equality. Since Var$(\Omega) = 0$ and Var$(z_T - z_S) = \Theta$Var(Δz_t) we have:

$$\text{Var}\left(w_T^* - w_S\right) = \Theta\sigma_z^2$$

Hence:

$$(1-\gamma)^2\overline{\sigma}^2 = \sigma_w^2 - \Theta^{-1}\text{Var}\left(w_T^* - w_S\right) \tag{14}$$

This test relates the observed variance of wage changes σ_w^2, net of the variance of the overall

[5] Implicitly, we assume here that separation takes place exactly at the end of the year of observation. This is an important assumption, since wages decline steeply in the last year before separation, see Figure 2. If separation occurs earlier on during the year of observation, part of the fall in wages during the last year before separation is captured by the previous observation.

shock z, to the degree of concavity in wages, $(1-\gamma)\overline{\sigma}\Delta\mathrm{E}(\Omega_\tau|A(\Theta))$, in other words, it tests whether there is sufficient yearly variation in wages to generate the observed concavity from selection.[6]

3 Empirical Analysis

3.1 The Data

We use a dataset based on a PSID extract of 18 waves, covering the years 1975 through 1992, same as the one used by Altonji and Williams (1997, 1999). Our model does not work well when employed people consider other alternatives than switching to another job, like 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 (we also drop the few observations that have a missing value for education) and less than 60 years of age. Our reasoning is similar to the one used in Mincer and Jovanovic (1981), who also use job separation synonymous to job change, thereby also defining labor mobility as change of employer and excluding other alternatives, which are minor phenomena in the case of the full-time male working force. Furthermore, we restrict the attention to those individuals that were employed, temporarily laid off, or unemployed at the time of the survey, and were not from Alaska

[6] This test does not account for effect of the condition $A(\Theta)$ on the observed variance in Δw_t. Due to this conditioning, the observed variance is a lower bound for σ_w^2. The expressions for the conditional variance are extremely cumbersome. However, for longer spells, $\Theta > 5$, the effect of conditioning on the variance is small.

or Hawaii. Finally, we discard all observations on unionized jobs.[7]. We use the tenure and experience measures constructed by Altonji and Williams (1999). Table 1 presents summary statistics of the data. Since we do not need wages in the tenure distribution analysis, observations with missing wage information are included in that 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 the data, and spells that start afterwards (left censoring). The lower half of the summary statistics table informs on the number of spells for each of these four types.

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
No. of obs. discarded from the data in Altonji and Williams(1997)					10351

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 data summarized in the top panel, keeping one observation for each job spell

3.2 The Parameters of the Tenure Distribution

Our estimation strategy uses the recursive feature of our model, that Ω and π can be estimated from the tenure distribution, and that these parameters can then be used to calculate $\Delta \mathrm{E}(\Omega_t | A(\Theta))$ and $\Delta \mathrm{E}(\Omega_t | B(\Psi))$, that are entered in the analysis of wage

[7]The previous working paper version of this paper includes unionized spells and controls for other covariates, see Buhai and Teulings (2006).

dynamics. Ω and π are estimated by maximum likelihood, using the density function (8). For the theoretical analysis, we have treated both parameters as constants that do not depend on worker characteristics. Empirically, one can expect that workers choose their optimal job type according to their characteristics. Hence, Ω and π are likely to differ according to both observed and unobserved worker characteristics. As observed worker characteristics we enter only experience at the start of the job, S. Since we deal with longitudinal data, we can take into account random worker effects. We do not consider random job effects for both theoretical and empirical reasons. From a theoretical point of view, our assumption of a frictionless market for alternative job opportunities, where the only constraint on instantaneous mobility is the specific investment in the present job and not the cost of getting another job offer, each worker type will choose that job type that fits best her comparative advantages, like in Sherwin Rosen's hedonic world of kissing curves. Hence, job characteristics are implied by worker characteristics. From an empirical point of view, we observe each job only once, thus we have no basis for identifying random job effects other than from functional form assumptions. Taking into account that Ω has to be positive, the following specification for Ω and π is adequate:

$$\begin{aligned} \Omega &= \exp\left(\beta_{\Omega 0} + \beta_{\Omega}\widehat{S} + u_{\Omega}\right) \\ \pi &= \beta_{\pi 0} + \beta_{\pi}\widehat{S} + u_{\pi} \end{aligned} \tag{15}$$

where u_{Ω} and u_{π} are normally distributed random worker effects with zero mean and standard deviations σ_{Ω} and σ_{π}, and where hat on $\widehat{S}$ denotes deviations from its mean over jobs. Hence, the intercept can be interpreted as the mean value for Ω and π respectively. We assume both random effects to be uncorrelated. Then, the contribution to the log likelihood function for an individual reads:

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

where j is the j^{th} job held by the worker, where d_j is a dummy variable, taking the value $d_j = 1$ if the job spell is completed and the value $d_j = 0$ otherwise. There are two reasons why we have to make amendments to the simple likelihood function in equation (16).

First, we could restrict the estimation to job spells starting within the observation range in the PSID extract. However, this means that 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 in the data to the maximum time span covered by the PSID sample, that is 17 years. Since long tenures contain relevant information, we want to include also spells started before their first wave in the PSID. We know either Θ_j or Ψ_j for these spells and we can compute experience at the beginning of a job by subtracting current tenure from current experience. However, we observe these spells only conditional on the fact that they have lasted till the start of our observation period. We should correct the log likelihood function for this condition:

$$\log L = \ln \int \int \overline{F}(\tau_1)^{-1} \prod_{j=1}^{J} \overline{F}(\Theta_j)^{1-d_j} \cdot f(\Theta_j)^{d_j} d\Phi(\frac{u_\Omega}{\sigma_\Omega}) d\Phi(\frac{u_\pi}{\sigma_\pi}) \tag{17}$$

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 collects data at a yearly interval, 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[8], but we do not know whether there has been another job spell between the job observed a year ago and the job observed now. Since the hazard rate implied by our model is hump shaped, with the hump likely to be within the first year, cf. Farber (1994), this phenomenon is expected to have a large impact on the estimation results. We are likely to overestimate Ω and π, since we miss part of the short tenures in our data. Hence, we have to correct for this form of left censoring. One solution to this problem is to use a similar conditioning as in equation

[8]Initial tenures are either reported or inferred by making them consistent with the latest reported tenures- see Altonji and Williams (1999 and previous working versions).

(17), where τ_j is the initial tenure at the first moment the job is observed (measured in months in PSID). However, this approach does not use the distribution of these τ_j's itself.[9] 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:

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

The total contribution to the likelihood of a spell with initial tenure τ and completed tenure Θ is therefore:

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

Hence, the log likelihood reads:

$$\log L = \ln \int\int \prod_{j=1}^{J} \frac{\overline{F}(\Theta_j)^{1-d_j} \cdot f(\Theta_j)^{d_j}}{\int_0^1 \overline{F}(x)dx} d\Phi(\frac{u_\Omega}{\sigma_\Omega}) d\Phi(\frac{u_\pi}{\sigma_\pi}) \tag{18}$$

The log likelihood that accounts both for jobs starting before their first reporting in the PSID and for the left censoring of spells shorter than a year started after the first wave of the PSID, can thus be written as:

$$\log L = \ln \int\int \prod_{j=1}^{J} \frac{\overline{F}(\Theta_j)^{1-d_j} \cdot f(\Theta_j)^{d_j}}{\overline{F}(\tau_1)^{I(j=1)} \left(\int_0^1 \overline{F}(x)dx\right)^{I(j\neq 1)}} d\Phi(\frac{u_\Omega}{\sigma_\Omega}) d\Phi(\frac{u_\pi}{\sigma_\pi}) \tag{19}$$

where $I(y)$ is the indicator function, taking value 1 if y is true and value 0 otherwise. We report results for (18), where we use only the sample of jobs that start within their observation period, and for (19), where we use all job spells, including those started before

[9] Maximum likelihood estimation using this approach yields a huge hump in the hazard rate, which implies a much higher share of spells shorter than a year that can be justified from the distribution of τ_j for jobs started after the first wave.

they are first observed in the PSID[10]. The estimation results are presented in Table 2.

Theoretically, the results for the two likelihood functions should be identical. The theoretical hazards for both models look indeed almost identical (cf. Figure 1 above), the only difference being the height of the peak, lower for the case where we use all job spells. The same can be concluded also by inspecting Table 2, where the estimated intercepts are very similar, while the coefficients for experience at job start are virtually the same, for both Ω and respectively π. The positive effect of experience on the drift π would be consistent with the idea that workers start their career with some initial job hopping, before settling down in a job that fits one's comparative advantages best. Furthermore, the intercept for π is positive and large in both estimations. In both cases, there are hardly observations for which π is negative. This implies that some job spells will last until the retirement of the worker. The fraction of jobs that never end, for mean values of the parameters, is about 10%.

Table 2: MLE Tenure Distribution Parameters

	Small Sample$^{(1)}$		**Large Sample$^{(2)}$**	
Variable	**Drift π**	**Dist Ω**	**Drift π**	**Dist Ω**
Intercept	0.226**	-1.243 **	0.141**	-1.197**
(st. errors)	(0.023)	(0.087)	(0.002)	(0.016)
Initial experience	0.009**	-0.006	0.012**	0.002
(st. errors)	(0.003)	(0.010)	(0.0002)	(0.002)
Random worker effects σ	0.309**	0.002	5.75e-007	3.66e-005
(st. errors)	(0.053)	(1.219)	(0.002)	(0.013)
Observations (job spells)	3169		4681	

$^{(1)}$Small sample= sample of job spells starting within the range covered in the PSID
$^{(2)}$Large sample= sample of all job spells
All covariates are taken in deviations from their means over jobs
Significance levels: † : 10% $*$: 5% $**$: 1%

[10] In order to estimate the log-likelihood functions above, we used simulated maximum likelihood, cf. Stern (1997). Sampling from a joint normal distribution with mean 0 and variances σ_π^2 and σ_Ω^2 and using a sampling size of 500 sampling points (the results are robust to altering the sampling dimension to any size between 100 and 500 sampling points) we achieved strong convergence in a reasonable number of iterations. We used the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method for convergence of derivatives, allowing for a tolerance of 1E-4 times the absolute value of the log likelihood.

One remarkable conclusion is that there are no unobserved random worker effects when we use the sample of all jobs spells, while there is unobserved heterogeneity in the drift for the sample including only the shorter spells within the observation range. Since the long spells started before the first wave contain crucial information, we focus on the estimation results obtained from the full sample of job spells in the subsequent wage dynamics analysis.

As a test of the goodness of fit of the model, we compute the predicted distribution of incomplete tenures after 32 years experience and compare that to the observed distribution.[11] Figure 6 depicts both the predicted and the empirical density of incomplete job spells. There is a reasonable correspondence between both densities. The peak in the first year is overestimated, but otherwise the shapes of the two densities are identical. Note the small peak in the density for short incomplete spells, which is due to the hump shape pattern in the hazard: if your job ends for instance in the last five years before the end of the observation period, there is a substantial probability that you experience further separations afterwards due to the peak in the hazard rate, leading to a peak of short incomplete tenures. Close alignment of the predicted and the empirical densities suggests that our model works well.

3.3 Test of the Random Walk Hypothesis in Wages

To prepare the ground for our formal analysis of wage dynamics, we document some stylized facts on the stochastic dynamics of wages. In particular, we verify that log wages follow a random walk using the methodology applied by MaCurdy (1982), Abowd and

[11] This density is calculated by a recursive scheme. We divide the 32 years time period in 32×256 subperiods. We calculate the distribution of completed tenures for jobs starting at the beginning of the career, in the first subperiod. For some of these jobs, $t > 32$, which is the density of incomplete tenures of 32 years. Then we calculate the distribution of completed tenures for jobs starting in the second subperiod, which is the number of jobs started in the first subperiod that separate in the second. We add this number to the corresponding completed tenures of the jobs started in the first subperiod. Then we calculate the completed tenure for jobs started in the third subperiod, etc. In these calculations we account for the effect of experience at the job start on the parameters Ω and π, which we estimated above.

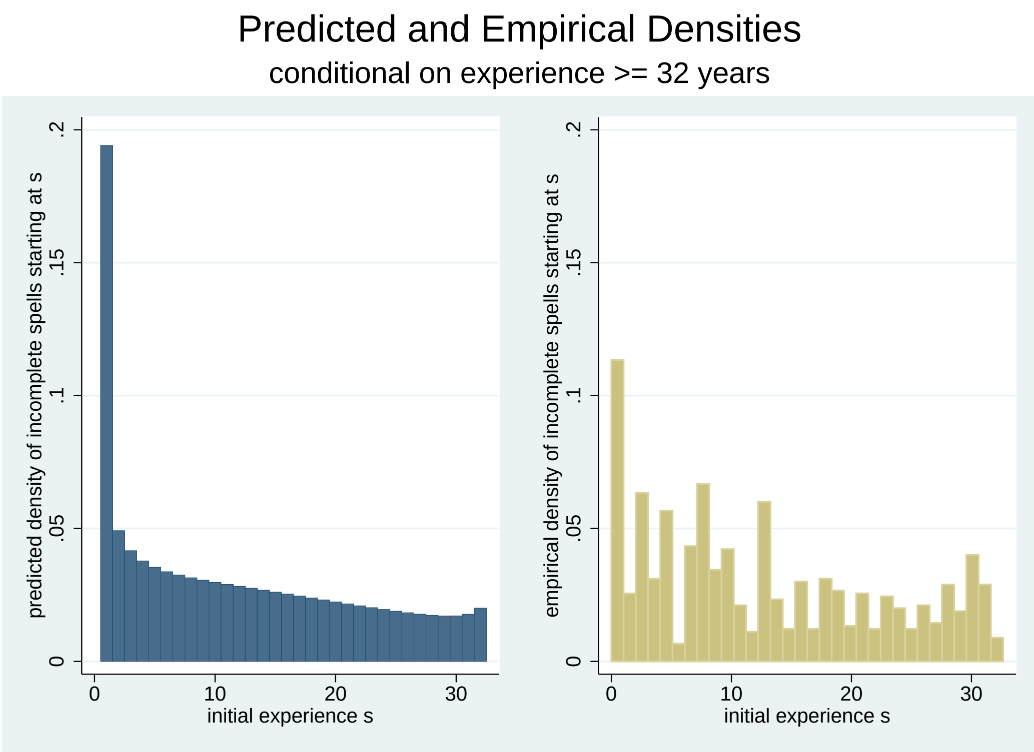

Figure 6: Density of Incomplete Job Spells with Exit Option

Card (1989) and Topel and Ward (1992). First, we run a regression of within-job log wage differentials on a number of controls:

$$\Delta w_t = \delta_0 + \delta_t t + \delta_\tau \tau + \Delta u_t \tag{20}$$

Though our theoretical model allows only for linear tenure and experience profiles, we enter both variables (t=tenure, τ=experience) in the first differenced equation to capture the eventual concavity of both profiles. Tenure is entered as a proxy for the concavity due to the selection process as discussed in Section 2.4. We run separate regressions for job stayers and switchers[12]. The regression results are displayed in Table 3.

Table 3: Within and Between-Jobs Wage Change Regressions

	Within-Jobs	**Between-Jobs**
Intercept	0.0458**	0.1109**
(std. err.)	(0.0042)	(0.0183)
Tenure	0.0001	-0.0034
(std. err.)	(0.0003)	(0.0028)
Experience	-0.0016**	-0.0039**
(std. err)	(0.0003)	(0.0014)
Observations	8802	1243
SER$^{(1)}$	0.1951	0.3584
R^2	0.0049	0.0092

Significance levels : † : 10% $*$: 5% $**$: 1%

$^{(1)}$ SER= standard error of the regression (root mean square error)

In the regression for stayers, we find some evidence for concavity in the experience, but not in the tenure profile. For job switchers we find that experience at separation negatively affects the wage differential to the new job, but that tenure at separation has no statistically significant effect. The results concerning experience are consistent with the findings in Buchinsky et al (2005)[13].

[12] In the case of job switchers we regress the change in the log wages from one job to the other, on the last tenure and respectively experience, at the moment of separation.

[13] Buchinsky et al (2005) also find a positive effect of the seniority at job separation on the starting wage in the new job and evidence of concavity in the seniority profile for job stayers, which are both statistically

Next, we construct a covariogram of residuals Δu_t of the within-job wage change regression, see Table 4.

Table 4: Residual Autocovariances for Within-Job Wage Innovations

Lag	Autocovariance	Standard Deviation
0	.0380**	.0020
1	-.0130**	.0017
2	-.0012*	.0006
3	.0009	.0007
4	.0004	.0007
5	-.0006	.0007
6	.00004	.0008
7	-.0005	.0008
8	-.0004	.0009
9	.0015	.0010
10	.00007	.0012
11	-.0011	.0015
12	.0004	.0013
13	.0016	.0012
14	-.0017	.0013
15	-.0022	.0017
16	-.0013	.0018

Significance levels : † : 10% $*$: 5% $**$: 1%

Residuals are strongly negatively correlated to their first lag, while autocovariances for longer lags are statistically insignificant beyond lag 2. This outcome is similar to results obtained by MaCurdy(1982), Abowd and Card (1989) and Topel and Ward (1992). Our covariogram is thus typical of an MA(2) process or even an MA(1) once we note that the second order lag autocovariance is close to 0. For simplicity, we focus on the MA(1) case. We decompose the stochastic time-variant component of the wage equation from (20) in a martingale persistent component e_t and a transitory component η_t:

$$\begin{aligned} u_t &= e_t + \eta_t \\ \Delta e_t &= \varepsilon_t \end{aligned} \qquad (21)$$

insignificant in our case. This difference can be due to slight differences either in the specification or in the PSID sample they use, as can be seen from comparing summary statistics.

where η_t and ε_t are i.i.d. with $\text{Var}(\eta_t) = \sigma_\eta^2$ and $\text{Var}(\varepsilon_t) = \sigma_w^2$. Then:

$$\begin{aligned}
\text{Var}(\Delta u_t) &= \sigma_w^2 + 2\sigma_\eta^2 \\
\text{Cov}(\Delta u_t, \Delta u_{t-1}) &= -\sigma_\eta^2 \\
\text{Cov}(\Delta u_t, \Delta u_{t-k}) &= 0, k > 1
\end{aligned}$$

This is a good description of the pattern of autocovariances in Table 4. Hence, a random walk with transitory shocks provides a fairly accurate description of the dynamics of log wages. Using the values in Table 4, we have $\sigma_\eta^2 = 0.0130$ and $\sigma_w^2 = 0.0380 - 2 \times 0.0130 = 0.0120$. The standard deviation of permanent innovations is substantial, $\sigma_w = 11\%$ per year. The variance of the transitory shocks σ_η^2 is consistent with the variance of the measurement error for earnings of US males reported by Bound and Krueger (1991). Hence, transitory shocks can be ignored in the analysis of job relocation.

Finally, we inspect whether the variance of the innovations in wages depends on tenure or experience. Table 5 present results for the Koenker (1981) Studentized LM version of the Breusch-Pagan (1979) test for homoskedasticity of u_t for both stayers and movers: the squared residuals from (20) are regressed on a constant term and the control variables.

Table 5: Heteroskedasticity Test for Wage Changes Within and Between-Jobs

	Within-Jobs	**Between-Jobs**
Intercept	0.0458**	0.1120**
(std. err.)	(0.0042)	(0.0191)
Tenure	-0.0005	-0.0013
(std. err.)	(0.0004)	(0.0030)
Experience	0.0006*	0.0017
(std. err)	(0.0003)	(0.0015)
Observations	8802	1243
SER(1)	0.1861	0.3743
R^2	0.0006	0.0011
Breusch-Pagan χ^2 test	$\chi^2_{(2)}$: N*R^2=5.28	$\chi^2_{(2)}$: N*R^2=1.37

Significance levels : † : 10% * : 5% ** : 1%

(1) SER= standard error of the regression (root mean square error)

First, in a joint test for tenure and experience, the null hypothesis of homoskedasticity cannot be rejected for both stayers or movers. At the level of single variables, only the effect of experience for job stayers is marginally significant. A learning model would imply a higher variance early on in the job, when the firm still has to learn the capability of the worker, see Jovanovic (1979) and Topel and Ward (1992). The results reported in Table 5 do not support this idea. The variance of wage changes is substantially higher for job movers than for stayers. This result might be explained by the strong fall in wages in the last year before separation, see Figure 2 and Proposition 2 above. Alternatively, it might be due to search frictions, so that workers cannot collect the best alternative job option, unlike we assume in this paper.

3.4 Wage Dynamics

Table 6 presents estimation results for equations (11) and (12), both for the whole sample and separately for the complete and incomplete job spells, and respectively for the job transitions, with heteroskedasticity-robust standard errors[14]. We add as an additional control experience, to account for concavity in the experience profile, which has been documented in Section 3.3. As long as the concavity in the experience profile affects p_t and r_t in the same way, it hardly affects the theoretical structure of our model.[15] The theoretically relevant regressors, $\Delta \mathrm{E}(\Omega|\cdot)$ and Ω, have the right sign in all specifications. The coefficients in the regression for job switchers in column 4 are badly determined. This is not surprising since Ω, the intercept, and the coefficient on experience are highly collinear (since Ω varies by initial experience, see Table 2). Hence, the subsequent discussion on Ω focuses on the first of the three columns. The coefficient on $\Delta \mathrm{E}(\Omega|\cdot)$, $(1-\gamma)\,\overline{\sigma}$, varies substantially between the three specifications, and is statistically insignificant in

[14] We tested for the absence of individual specific effects in all regressions.

[15] Since the concavity of the experience profile affect p_t and r_t in the same way, it drops out in $b_t = p_t - r_t$. Hence, as long as separations are driven by b_t falling below a constant separation threshold b^T, this concavity does not matter for our analysis. Strictly speaking, the Bellman equations (2) do not apply, since the concavity in the experience profile adds another state variable, and hence the separation threshold b^T will depend on t. We ignore this effect.

the regression for completed job spells in column 2. Although the differences in the estimated value across columns are insignificant, these results suggest that there is downward rigidity in wages. Where the model predicts a fall in wages at the job compared to the outside wage at the end of a job spell, see Figure 2, the data do not seem to support this idea.

Table 6: Wage Change Regressions: Overall, Completed Spells, Incomplete Spells, Job Transitions

	1: Overall	2: Completed	3: Incomplete	4: Job Switch
Intercept	0.047**	0.049**	0.031**	-8.253
(st. errors)	(0.004)	(0.009)	(0.008)	(6.738)
ΔE(Ω|.)	0.011*	0.007	0.035*	0.138*
(st. errors)	(0.005)	(0.007)	(0.016)	(0.068)
Ω	0.203**			28.293
(st. errors)	(0.064)			(22.252)
Experience	-0.0016**	-0.0015**	-0.0013**	0.0403
(st. errors)	(0.0002)	(0.0006)	(0.0003)	(0.035)
Observations	8994	1648	6833	513
SER	0.209	0.185	0.197	0.369
R^2	0.008	0.006	0.006	0.024

Significance levels : † : 10% $*$: 5% $**$: 1%

As further check of the notion of downward rigidity, we re-estimated the equation for all completed job spells, eliminating the last within spell wage change before separation.[16] Then, the coefficient for $\Delta E(\Omega|\cdot)$ goes up to 0.037 (std.err. 0.011) in column 1 and 0.043 (std.err. 0.023) in column 2. So indeed, the last observation before a job change, where $\Delta E(\Omega|\cdot)$ is negative to accommodate the downward part of the hump shape profile, does not fit the model well, consistent with the idea of downward rigidity. The variable Ω only shows up in the equation for job switchers, compare equations (11) and (12). Hence, its coefficient, $\overline{\sigma}$, is only identified by comparing wage changes within job spells and between

[16] The wage change Δw_T^* is included in the regression for job switchers, so the last change included in the regression for completed spells is Δw_{T-1}. In the alternative specification, these observations are excluded.

job spells, that is, in column 1, by identifying the intercept and the coefficients on $\Delta\mathrm{E}(\Omega|\cdot)$ and experience on wage changes within job spells, and then identifying the coefficient on Ω as the differential impact on wage changes for job movers.

The estimation results from column 1 imply $\gamma = 1 - 0.011/0.203 = 0.946$. Apparently, separation is driven by selectivity in the shocks to the worker's reservation wage r_t, not to the current job's productivity p_t, which seems somewhat counter-intuitive. Just as a way to illustrate the sensitivity of this estimate to the occurrence of downward rigidity, we redo this calculation for the alternative specification, where we omit the last observation on Δw_t before job change; then $\gamma = 1 - 0.037/0.203 = 0.818$.[17] We can calculate the return to tenure, $\overline{\sigma}\pi = 0.203 \times 0.14 = 2.8\%$ (taking the estimated mean value of $\pi = 0.14$ from Table 2 above). However, the high value of γ implies that most of the return to tenure, almost 95%, takes the form of the log reservation wage r_t falling, instead of the inside wage w_t rising, cf. equation (10). The tenure profile due to the rise in log productivity in the current job p_t is really small, $(1-\gamma)\overline{\sigma}\pi = 0.011 \times 0.14 = 0.15\%$ (or $0.037 \times 0.14 = 0.52\%$ in the alternative specification). Although the discussion on downward rigidity has shown that this is a rather thin line of identification of γ, this is the first research to actually account for selectivity in the realised outside wages.

Table 7 below presents the estimation results for equation (14), the variance of σ_z^2. Column 1 reports the coefficients for equation (13). Column 2 regresses the squared residuals from the first column on an intercept and the completed tenure. We restrict the sample to jobs lasting more than 1 year, since wage changes for jobs that last less than one year are noisy. The intercept captures the excess variance for job movers, see Table 5, and the variance of the transitory shocks in wages σ_η^2, see equation (21). The coefficient for tenure is an estimator of $\sigma_z^2 = \Theta^{-1}\mathrm{Var}(w_T^* - w_S)$. However, due to the low sample size, this coefficient is not statistically significant. Using the estimator for σ_w^2 of 0.0120, see Section 3.2, equation (14) implies $(1-\gamma)\overline{\sigma} = \sqrt{0.0120 - 0.0104} = 0.040$. Given the noisiness of the estimate of σ_z^2, this test does not have much power. However, the point

[17] This calculation is hazardous, since downward wage rigidity is also affecting the value of $\Delta\mathrm{E}(\Omega|\cdot)$ in the regression for job switchers, and hence our estimate of the coefficient for Ω.

estimate is consistent with the estimates of $(1-\gamma)\overline{\sigma}$ derived from Table 6, in particular in the alternative specification where we allow for downward rigidity in wages.

Table 7: Regression on Changes in Initial Wages between Jobs

	1: Initial wages	**2: Residual variance**
Intercept		0.130^{**}
(st. errors)		(0.030)
Tenure	0.033^{**}	0.010
(st. errors)	(0.011)	(0.008)
$\Delta\Omega$	0.168	
(st. errors)	(0.142)	
Observations	328	328
SER	0.400	0.355
R^2	0.121	0.003

Significance levels : † : 10% $*$: 5% $**$: 1%

We use completed jobs that last more than 1 year (Tenure >1)

The dependent variable in column 2 is the residual variance from column 1

A final question we ask is to what extent the option to switch jobs limits the growth of the variance in log wages over time. Without the option to switch jobs, the variance of log wages would increase linearly over time, due to the fact that z_t and b_t follow a random walk. However, the option to switch jobs allows the worker to eliminate bad trajectories of b_t, thereby compressing its variance. This can be seen from the distribution of incomplete tenures, see Figure 6 above, showing that a substantial fraction of the jobs has an incomplete tenure of less than 32 years. There are two mechanisms that lead to compression. First, many jobs have an incomplete tenure of less than 32 years and hence a smaller variance, since the variance increases proportional to incomplete tenure. Second, those jobs that are still going on after some period are a selective sample of all the trajectories that have started initially, namely those which never crossed the separation threshold. This selection process compresses the variance. We use the computed density of incomplete tenures from Figure 6, and the density of $\Omega_\tau = b_t/\sigma$ conditional on the incomplete tenure τ, $g(\Omega_\tau, \tau)/\overline{F}(\tau)$, see equations (6) and (7). In Figure 7, we plot together 2 graphs: first, the evolution of the variance of b_t without the option to switch jobs, the line $\sigma^2\Theta$, and second, the evolution of the variance with that option. The plots reveal that the option to switch jobs compresses the variance of b_t considerably: by

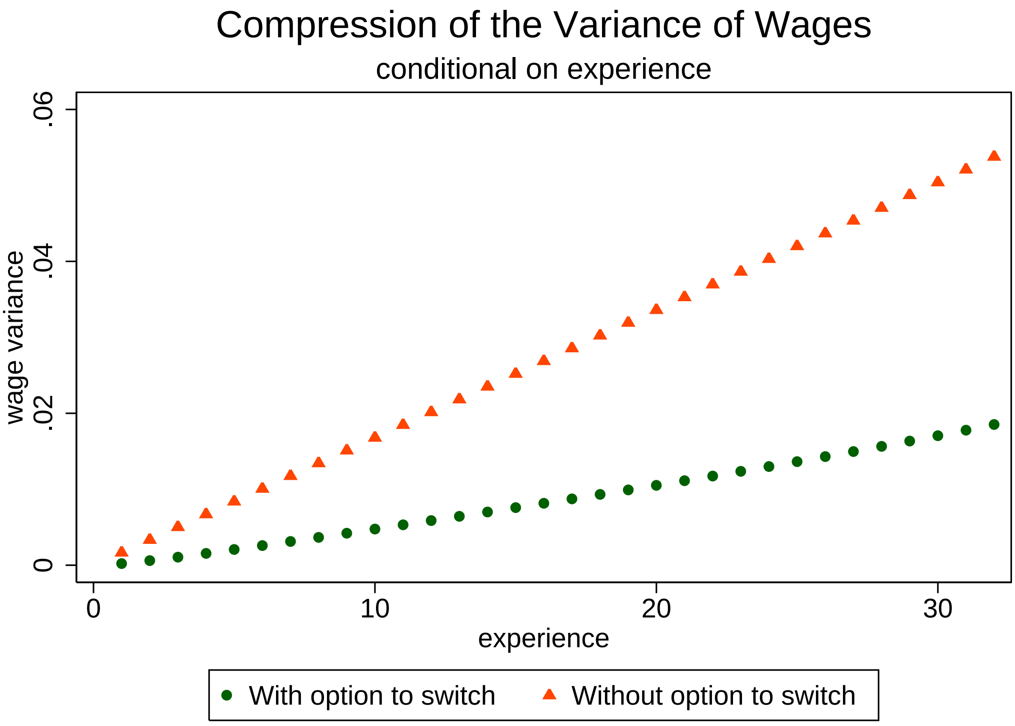

Figure 7: Compression Effect of the Exit Option on the Variance of b_t

almost 65%, after 32 years of experience[18]. Note however that the variance of z_t remains unaffected by this process, while this accounts for the main share of the total variance of w_t in our calculations.

4 Discussion and Conclusions

We analyzed a model of the evolution of wages and the duration of job spells featuring frictionless labor market at the moment of job start –enabling workers to pick the best job alternative at that moment– specific investment and hence subsequent lock-in on the

[18]For the computation we use $\sigma = 0.040$, as computed above in the variance comparison test. Hence, after 32 years of experience, without the exit option the variance of b_t would be $32 * 0.040^2 = 0.0512$. With the option to exit the wage variance is compressed after the same period to about 0.018.

current job, and efficient bargaining over the match surplus. This model explains the data on the job tenure distribution and wages for the USA surprisingly well, the main deviation being that the data suggest there is downward wage rigidity for which we do not allow in our model. We have proven the remarkable result that in this model the evolution of log wages in completed job spells does not provide any information whatsoever on wage-tenure profiles, since this evolution is independent of the drift in log wages. Hence, the tenure profile can only be estimated either from the distribution of tenures or from log wages in incomplete job spells. We have verified that the wage dynamics within jobs closely resembles a random walk; that the predicted job hazard rate is humped shaped with the peak very early in time, closely tracing the empirical evidence on job exits; and that the variance of the within-job wages does not diminish with tenure or experience, a fact that is less easily squared with the learning model. We have further shown that the concavity in the observed tenure profile is easily explained by the selection of the surviving employment matches, even when the underlying tenure profile is linear. In general, the selection effect tends to be much more important than the deterministic trend. This is in fact the first research that looks at selectivity in the realised outside productivities. Remarkably, job separation is driven more by the selectivity in the outside productivity r_t, than by shocks to the inside productivity in the job p_t. Almost 95% of our estimated tenure profile is accounted for by the selectivity in the outside productivity. However, identification of the part of the variance due to variation in r_t is fragile. A rough calculation suggest that allowing for downward rigidity in wages reduces this estimate to 81%. We find excess variance of wage changes at job transition. This might indicate that our assumption of frictionless market for job alternatives at the moment of job change is incorrect.

5 References

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

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

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

Bound, J., and A. Krueger (1991),"The Extent of Measurement Error in Longitudinal Earnings Data: Do two wrongs make a right?", *Journal of Labor Economics*, 9, 1-24.

Breusch, T.S. and A.R. Pagan (1979), "A simple test for heteroskedasticity and random coefficient variation", *Econometrica* 47, 1287-94

Buchinsky, M., D. Fougere, F. Kramarz and R. Tchernis (2005), "Interfirm mobility, wages, and the returns to seniority and experience in the US", IZA Discussion Paper No. 1521

Buhai, I.S. and C.N. Teulings (2006), "Tenure Profiles and Efficient Separation in a Stochastic Productivity Model", Tinbergen Institute Discussion Paper 2005-099/3, revised October 2006

Cox, D. R. and H.D. Miller (1965), *The Theory of Stochastic Processes*, London: Methuen

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 of 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. (1979), "Firm-specific capital and turnover", *Journal of Political Economy* 87(6), 1246-60

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

Koenker, R. (1981), "A note on studentizing a test for heteroskedasticity", *Journal of Econometrics* 17, 107-12

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

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

Mincer, J. and B. Jovanovic (1981), "Labor mobility and wages", in S. Rosen (ed), *Studies in Labor Markets*, ch. 1, 21-63

Mortensen, D.T. (1988), "Wages, separations, and job tenure: On-the-job specific training or matching?", *Journal of Labor Economics* 6(4), 445-71

Neal, D. (1995), "Industry-specific human capital: Evidence from displaced workers", *Journal of Labor Economics*, 14(4), 653-77

Stern, S. (1997), "Simulation-based estimation", *Journal of Economic Literature* 35(4), 2006-39

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", J*ournal of the Royal Statistical Society: A*, 142(4), 468-78

Zhang, P.G. (1998), *Exotic Options* (2nd edition), World Scientific Publishing Co.

A Appendix Conditional Expectation of Ω_τ

A.1 Completed spells: proof of Proposition 2

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 (7) and (8), thus $\overline{F}(\tau,\Omega)$ and $f(\tau,\Omega)$. Let $h(\omega,\tau,\Theta)$ be the density of $\Omega_\tau=\omega$ conditional on $A(\Theta)$. Comparing this density to $g(\omega,t)$, there is one additional condition: $\Omega_\Theta=0$. Hence, $h(\omega,t,\Theta)$ 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 $\Omega=\omega$. Hence, its density is $f(\Theta-\tau,\omega)$. Then $h(\omega,\tau,\Theta)$ can be calculated from $f(\cdot)$ and $g(\cdot)$ by Bayes's rule:

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

Substitution of equation (6) in the above yields:

$$\begin{aligned}
h(\omega,\tau,\Theta) &= \frac{\omega}{\Omega m(\tau)}\sqrt{\frac{1}{m(\tau)\tau}} \\
&\quad\times\left[\phi\left(\sqrt{\frac{1}{m(\tau)\tau}}\left[\omega-m(\tau)\Omega\right]\right)-\phi\left(\sqrt{\frac{1}{m(\tau)\tau}}\left[\omega+m(\tau)\Omega\right]\right)\right] \\
m(\tau) &\equiv \frac{\Theta-\tau}{\Theta}
\end{aligned}$$

Hence, $\mathrm{E}(\Omega_\tau|A(\Theta))$ satisfies:

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

The first and second derivatives of $\mathrm{E}(\Omega_\tau|A(\Theta))$ read:

$$\begin{aligned}
\frac{d\mathrm{E}(\Omega_\tau|A(\Theta))}{d\tau} &= -2\sqrt{\frac{1}{m(\tau)\tau}}\phi\left(\sqrt{\frac{m(\tau)}{\tau}}\Omega\right) + \left(\frac{1}{\Omega} - \frac{\Omega}{\Theta}\right)\left[2\Phi\left(\sqrt{\frac{m(\tau)}{\tau}}\Omega\right) - 1\right] \\
\frac{d^2\mathrm{E}(\Omega_\tau|A(\Theta))}{d\tau^2} &= -\sqrt{\frac{1}{m(\tau)\tau}}^3\phi\left(\sqrt{\frac{m(\tau)}{\tau}}\Omega\right)
\end{aligned}$$

■

A.2 Incomplete job spells

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

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

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

$$\begin{aligned}
\mathrm{E}(\Omega_\tau|B(\Psi)) &= \int_0^\infty \omega h^*(\omega,\tau,\Psi)d\omega \\
&= \frac{\int_0^\infty \omega\overline{F}(\Psi-\tau,\omega)g(\omega,\tau)\,d\omega}{\int_0^\infty \overline{F}(\Psi-\tau,\omega)g(\omega,\tau)\,d\omega}
\end{aligned}$$

where $\overline{F}(\Psi-\tau,\omega)$ is given by equation (7). This expression is evaluated numerically.