Tenure Profiles and Efficient Separation in a Stochastic Productivity Model*

I. Sebastian Buhai† and Coen N. Teulings‡

July 18, 2011

Abstract

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. 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. 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

*For helpful comments and suggestions we thank 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 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. We are also grateful to Nick Williams and Lennart Janssens for assistance in handling the data. The usual disclaimers apply.

†Aarhus University and Northwestern University; sebu@asb.dk, sbuhai@northwestern.edu

‡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 tenure, 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 that actual returns are 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 (2010). 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 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 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 relative 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), or 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 random walks. 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 stands here for what we traditionally call the return to tenure.

Starting a job requires an irreversible specific investment, which is lost upon separation. This investment has therefore 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, this model could equally well be applied to industry or occupation tenure, as suggested, e.g. 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 tenure distribution 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, which is a remarkable conclusion given 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 50 to 80% of the tenure profile, with the estimate range due to possible mis-specification of the model. In particular, our estimation results provide some indication 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. Our estimates also provide evidence of excess variance in wages for job movers, suggesting the failure of our assumption of Walrasian market for job alternatives. The apparent tenure profile is estimated to about 0.6% 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 for the wage returns to tenure would be on the very low end of the spectrum, 0.10%- 0.25% per year.

The paper is structured as follows: the model is discussed in Section 2, the identification and the estimation strategy are set out in Section 3, the empirical analysis is presented in Section 4, and Section 5 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:

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

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 from 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 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 and 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):

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

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:

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

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:*

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

with $j(\cdot)$ and $v(\cdot)$ satisfying:

$$\begin{aligned} \left(\rho - \mu_r - \frac{1}{2}\sigma_r^2\right) j &= \exp(b_t) - 1 + \left(\mu + \sigma_{pr} - \sigma_r^2\right) j' + \frac{1}{2}\sigma^2 j'' \\ \left(\rho - \mu_r - \frac{1}{2}\sigma_r^2\right) v &= \left(\mu + \sigma_{pr} - \sigma_r^2\right) v' + \frac{1}{2}\sigma^2 v'' \end{aligned} \tag{6}$$

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:*

$$\begin{aligned} j\left(b^S\right) &= v\left(b^S\right) + I \\ v\left(b^T\right) &= j\left(b^T\right) \\ j'(b^S) &= v'(b^S) \\ v'(b^T) &= j'(b^T) \end{aligned} \tag{7}$$

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 two differential equations have an analytical solution which is presented in Appendix A. 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

[1] We use:

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

and likewise for $V\left(\cdot\right)$.

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 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 a 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 $\tau \equiv t - S$, 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. Further define:

$$\begin{aligned} \Omega_\tau &\equiv \frac{b_t - b^T}{\sigma} \\ \Omega &\equiv \frac{b^S - b^T}{\sigma} > 0 \\ \pi &\equiv \frac{\mu}{\sigma} \end{aligned} \tag{8}$$

Thus Ω_τ is a Brownian with drift π and unit variance per unit time.[2] By construction $\Omega_0 = \Omega$ and $\Omega_\Theta = 0$. In Appendix A.1 we show that Ω is an implicit function of the model's structural parameters, $I \equiv I\left(\Omega, \underline{\mu}, \Sigma\right)$, with $I_\Omega(\cdot) > 0$. Hence, we treat the parameter Ω as the parameter of interest. If desired, the underlying structural parameter I can be recovered via the function $I\left(\Omega, \underline{\mu}, \Sigma\right)$.

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_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)[3]:

$$g(\omega, \tau) = \frac{1}{\sqrt{\tau}}\left[\phi\left(\frac{\omega - \Omega - \pi\tau}{\sqrt{\tau}}\right) - e^{-2\Omega\pi}\phi\left(\frac{\omega + \Omega - \pi\tau}{\sqrt{\tau}}\right)\right] \tag{9}$$

[2] Using the moment of the job start instead of some other point in time as our point of reference does not rule out a return to experience. It is just a convenient renormalization of the time variable.

[3] 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.

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, $\overline{F}(\tau) = \Pr(\Theta > \tau)$:

$$\overline{F}(\tau) \equiv \Phi\left(\frac{\Omega + \pi\tau}{\sqrt{\tau}}\right) - e^{-2\Omega\pi}\Phi\left(\frac{-\Omega + \pi\tau}{\sqrt{\tau}}\right) \tag{10}$$

where $\Phi(.)$ is the standard normal CDF. This expression above is identical to Whitmore (1979: eq. 2)[4].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:

$$f(\tau) = \frac{\Omega}{\tau\sqrt{\tau}}\phi\left(\frac{\Omega + \pi\tau}{\sqrt{\tau}}\right) \tag{11}$$

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 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 (10) for $\tau \to \infty$, hence by $1 - e^{-2\Omega\pi}$. In

[4]Remark that $\overline{F}(\tau)$ in (10) is also an instance of the well-known Black-Scholes formula. In fact, already from above, the two expressions in (6) are familiar Black–Scholes second order PDEs for the valuation of financial assets; our theoretical framework bears a perfect analogy to the problem of pricing European call (or put) options, here the underlying "stock" being the difference between the inside and outside productivities.

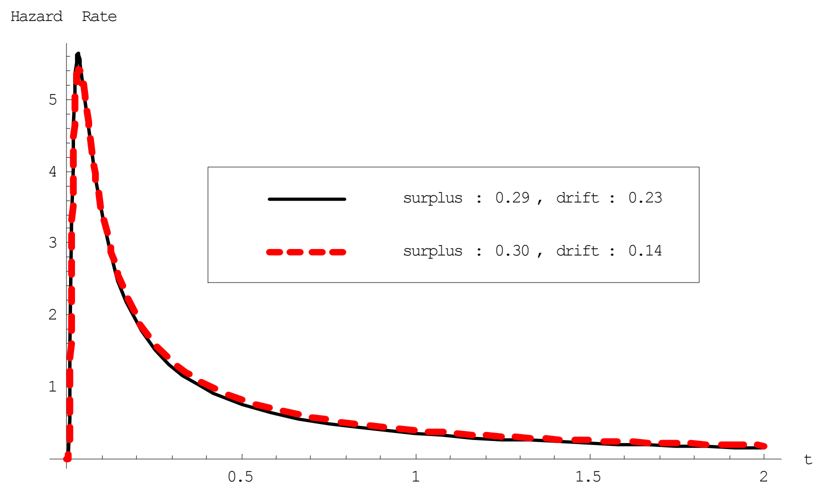

Figure 1: Predicted Job Hazards

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 Ω and π estimates for the mean values of the observed and unobserved worker characteristics, see Section 3, Table 2 below.[5] 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, imposing the log linearity of the sharing rule *a priori*, and deriving the intercept of that rule from the assumption of efficient bargaining.[6] According to this rule, the worker's log

[5] In terms of equation (15), $\Omega = \exp(\omega_0)$ and $\pi = \pi_0$

[6] 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.

wage w_t satisfies:

$$w_t = \psi + r_t + \beta \left(b_t - b^T\right) + u_t = \psi + r_t + \overline{\sigma}\Omega_\tau + u_t \tag{12}$$

where $\overline{\sigma} \equiv \beta\sigma$. We assume u_t is an i.i.d. random variable distributed $N(0, \sigma_u^2)$: 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).[7] The parameter β can be interpreted as the worker's bargaining power. If $\beta = 0$, the worker has her 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. 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.

To close the model, the parameter ψ and the worker's share in the cost of investment remain to be determined. Explicit expressions for these parameters in terms of the structural parameters of the model can be derived from the solution to the Bellman equation for the worker's value of holding the job, which reads analogous to the value functions in equation (2). Since these expressions are irrelevant for the subsequent empirical analysis, as they do not contribute to the identification of the model's structural parameters, their derivation is relegated to Appendix A.2.

2.4.2 Selectivity in Tenure Profiles

Equation (12) 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 and the secular growth of real wages due to technological progress; $\overline{\sigma}\pi$ is the deterministic part of the tenure profile. Were the

[7] Most researchers report some mean reversion in wages with the AR(1) coefficient being in the range of 0.95-1.00, and a MA(2) transitory process with the second lag being much smaller than the first.

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) $\Omega_\zeta > 0, \forall\zeta, 0 \leq \zeta < \Theta$. 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) $\Theta > \Psi > \tau$, and hence (iii) $\Omega_\zeta > 0, \forall\zeta, 0 \leq \zeta \leq \Psi$. We refer to this second information set as $B(\Psi)$.

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

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

Proof See Appendix B.1.■

This proposition implies that $\mathrm{E}[\Omega_\tau|A(\Theta)]$ does not depend on the tenure profile in wages, $\overline{\sigma}\pi$; see also Van der Ende (1997) for a similar result. 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. Hence, the deterministic part of the tenure profile does affect the job separation rate, in particular for long tenures (since the drift increases linearly in time, while the standard deviation of the random walk increases only proportional to the square root of time), but it does not affect the evolution of wages conditional on the moment of separation Θ.

This conclusion depends crucially on the assumption of efficient bargaining. This assumption dictates that the evolution of wages over a job spell satisfies

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

see equation (12); $w_t - r_t - u_t$ is the log of the wage on the job spell relative to market wage at time t, corrected for transitory shocks. The difference between the final and the starting value of this log relative wage is equal to 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 length of the job spell Θ, and ignoring the effect of transitory shocks, log relative wages decline by $\overline{\sigma}\Omega$ over the duration of a completed job 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. From this perspective, data on the tenure distribution are more informative on the return to tenure than data on wages.

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

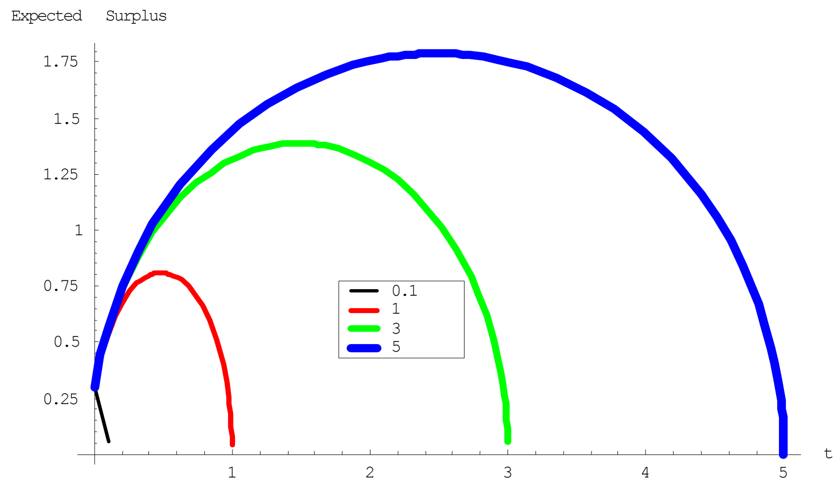

Figure 2: Expected Surplus in Completed Job Spells

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 B.2. 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, 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 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. The intuition is that jobs that last longer pay higher wages half way the job spell, see Figure 2. A higher "true" return to tenure π implies that jobs last longer on average, and hence pay higher wages measured by $\mathrm{E}[\Omega_\tau|B(\Psi)]$, so that the "true" return to tenure is correlated to the "observed" tenure profile. The impact of the drift, the "true" return to tenure π, is negligible compared to that of selectivity, as documented by Figure 4. This figure compares the trajectories for completed spells, $\mathrm{E}[\Omega_\tau|A(\Theta)]$, for incomplete spells,

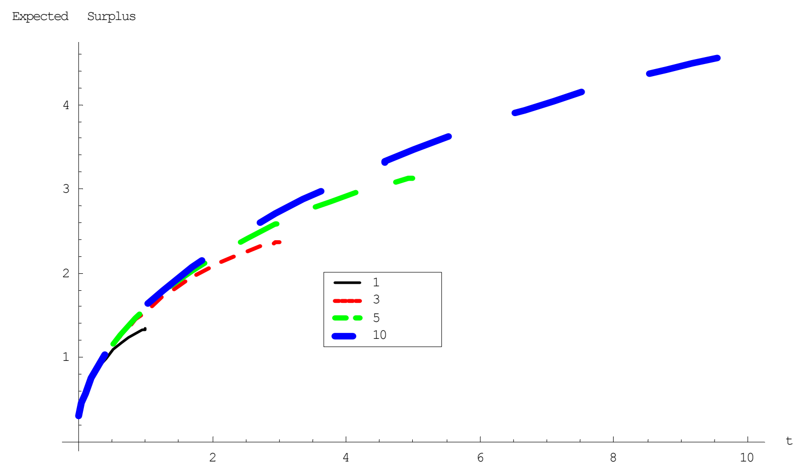

Figure 3: Expected Surplus in Incomplete Job Spells

$E[\Omega_\tau|B(\Psi)]$, and the deterministic part, $E(\Omega_\tau)$, which is equivalent to the tenure profile in the absence of selectivity. The trajectories for complete and incomplete spells are far above the deterministic part, except for the final year(s) before separation. This shows that the effect of selectivity on the observed tenure profile is far more important than the true return to tenure. In Figure 5 we plot $E[\Omega_\tau|A(\Theta)]$ and $E[\Omega_\tau|B(\Psi)]$ for long job durations, $\Theta = 10$, 20 and respectively $\Psi = 10$, 20.

The same analysis can be done for the second moment of Ω_τ. In the absence of the condition $A(\Theta)$, $\mathrm{Var}[\Delta\Omega_\tau]$ would be equal to unity, see equation (8). However, the conditions $A(\Theta)$ for complete and $B(\Psi)$ for incomplete spells introduce selectivity in the trajectories of the random walk that are considered. This selectivity reduces the variance. Expressions for $\mathrm{Var}[\Delta\Omega_\tau|A(\Theta)]$ and $\mathrm{Var}[\Delta\Omega_\tau|B(\Psi)]$ can be found in Appendix B.

Figures 6 and 7 depict the evolution of the variance in completed and respectively incomplete job spells, for 2, 3, 5 and 10 years of tenure. Without the conditioning, the variance is equal to unity. The conditioning limits the variance: initially, the variance is low, because the constraint Ω_τ being positive over the course of a job spell is quite informative, since Ω_0 is quite small. The same argument applies towards the end of completed job spells, since $\Omega_\Theta = 0$ by construction. Selectivity does not play a large role

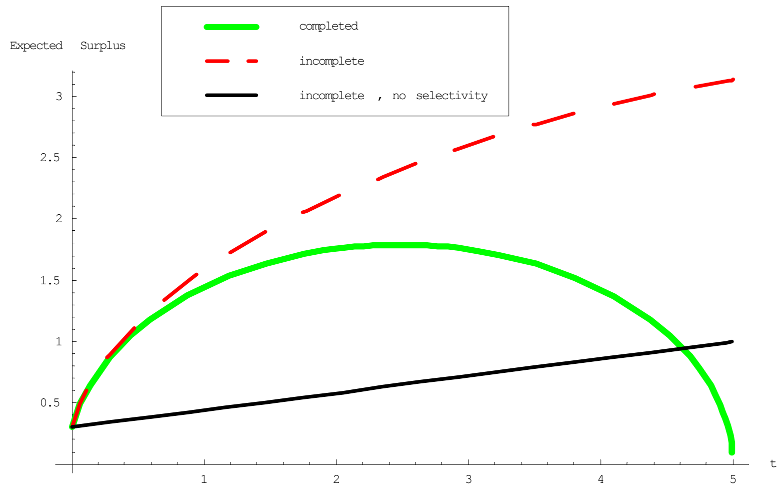

Figure 4: Selectivity versus Drift in the Expected Surplus

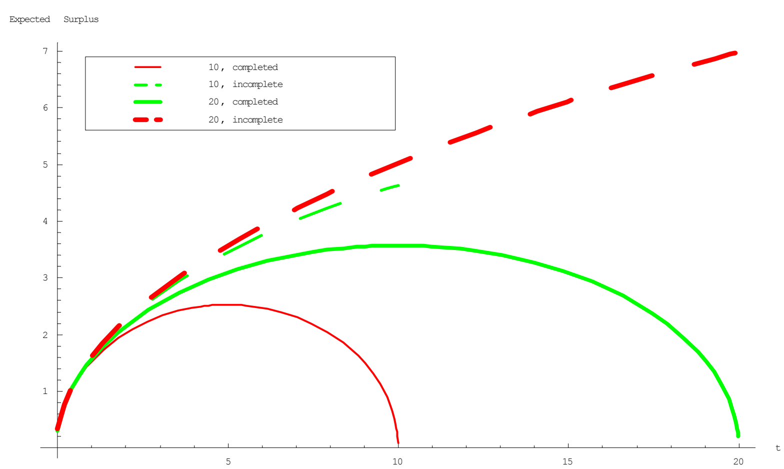

Figure 5: Expected Surplus in Long Spells

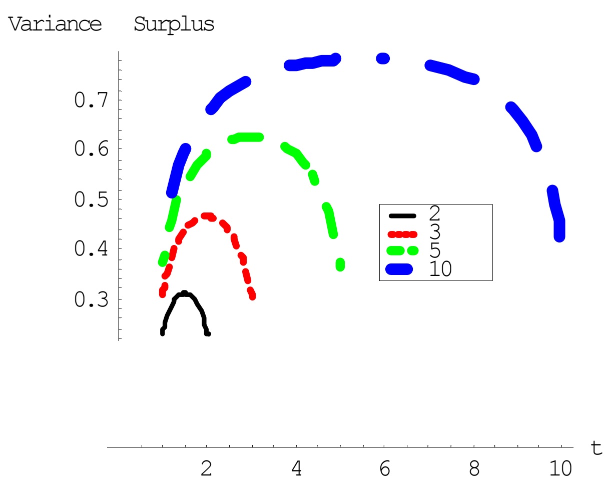

Figure 6: Variance surplus completed spells

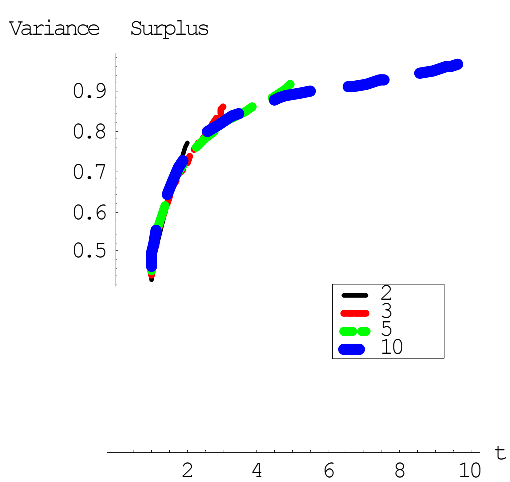

Figure 7: Variance surplus incomplete spells

for incomplete spells. For longer spells, the variance converges to unity, as expected, given that the unconstrained variance of Ω_τ was normalized to unity.

2.5 A Reinterpretation of the Wage Equation

The implications of our analysis of the selectivity in tenure profiles surfaces most clearly when we rewrite equation (12) 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 both components such their marginal effect on Δr_t and Δp_t is equal to unity:

$$\begin{aligned} \Delta r_t &= \Delta z_t - \gamma \Delta b_t \\ \Delta p_t &= \Delta z_t + (1-\gamma)\Delta b_t \end{aligned}$$

This decomposition satisfies the constraint $\Delta p_t - \Delta r_t = \Delta b_t$ imposed by equation (4). Given the joint normality of Δp_t and Δr_t, this decomposition is unique. The parameters μ_z, σ_z, and γ are fully determined by the structural parameters of the model:[8]

$$\begin{aligned} & \Delta z_t \tilde{} N\left(\mu_z, \sigma_z^2\right) \\ \gamma &= \frac{\sigma_r^2 - \sigma_{pr}}{\sigma^2} \\ \mu_z &= \mu_r + \gamma \mu_p \\ \sigma_z^2 &= \frac{\sigma_p^2 \sigma_r^2 - \sigma_{pr}^2}{\sigma^2} \end{aligned} \tag{13}$$

where 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

[8] γ and σ_z^2 can be found by solving the system:

$$\begin{aligned} \sigma_{pr} &= \sigma_z^2 - \gamma(1-\gamma)\sigma^2 \\ \sigma_r^2 &= \sigma_z^2 + \gamma^2 \sigma^2 \end{aligned}$$

with equation (12) yields:

$$\Delta w_t = \Delta z_t + (1-\gamma)\beta\Delta b_t + \Delta u_t = \Delta z_t + (1-\gamma)\overline{\sigma}\Delta\Omega_\tau + \Delta u_t \tag{14}$$

where $\overline{\sigma} \equiv \sigma\beta$. The wage formation process is therefore characterized completely by five parameters: $\overline{\sigma}, \gamma, \mu_z, \sigma_z$, and σ_u. The parameter γ is a reflection of the correlation between the match surplus and the reservation wage. In one extreme case, $\gamma = 0$, we can write $\Delta p_t = \Delta r_t + \Delta b_t$, where both 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 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".

3 Estimation Methodology

3.1 Identification

The model has eight structural parameters: 2 drift parameters $\underline{\mu}$, 3 (co)variances Σ, the investment level I, the worker's bargaining power β, and the variance of the transitory shock, σ_u^2. In Section 2.3, we have shown that the distribution of completed job tenures is fully determined by two parameters, π and Ω, while in Section 2.4, wage formation process is characterized completely by five parameters $\overline{\sigma}, \gamma, \mu_z, \sigma_z$, and σ_u. Hence, one can never hope to identify more than these seven parameters, two from data on the tenure distribution, and five from panel data on wages. The relation of these seven parameters to the eight structural parameters is given by equations (8), (26, in Appendix A.1), and (13). Given that our model has eight structural parameters, the model is therefore identified up to one degree of freedom. Equation (14) reveals why this is the case. Only the product $\overline{\sigma} \equiv \beta\sigma$ shows up in the both expressions, and neither β or σ separately. Both these seven and the eight parameters are structural, since the derived parameters are simple functions of the underlying parameters. Given an assumption on one of the eight structural parameters, there is a one-to-one correspondence between the other seven underlying parameters and the derived parameters, see equations (8), (26), and (13). Data on either the cost of investment or productivity would resolve this underidentification. Data on investment would offer direct information on the necessary specific investment I.[9] Data on productivity would offer direct information on σ. Then, β could be established as $\overline{\sigma}/\sigma$. However, neither type of information is available here.

Since the data generating process is fully determined by the seven derived parameters $\pi, \Omega, \mu_z, \sigma_z^2, \overline{\sigma}, \gamma$ and σ_u^2, the subsequent discussion of the estimation results focuses on these parameters. As discussed in Section 2.1, we assume that personal characteristics affect the parameters I and μ_p. We allow for both observed and unobserved characteristics. As observed characteristic we enter experience at the moment of job start S, measured in deviation from its sample mean. By letting I and μ_p depend on the initial job experience we preserve a crucial feature of our model, namely that the level of investment I and the drifts $\underline{\mu}$ are constant over the course of a job spell[10]. We can allow μ_p to take a particular value at the beginning of a job spell and to remain constant afterwards. However, we cannot allow μ_r to do so, since contrary to the productivity p_t at this particular job, the

[9] 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.

[10] The results are identical when allowing the parameters to vary with higher order effects of the experience at job start.

log reservation wage r_t is independent of what job the worker actually occupies. Since we deal with longitudinal data, we account for random worker effects.[11]

$$\begin{aligned} \Omega &= \exp\left(\omega_0 + \omega_1 S + e_\Omega\right) \\ \pi &= \pi_0 + \pi_1 S + e_\pi \\ \mu_z &= \mu_0 + \gamma\overline{\sigma}\left(\pi_1 S + e_\pi\right) \end{aligned} \tag{15}$$

where e_Ω and e_π are normally distributed random worker effects with zero mean and standard deviations σ_Ω and σ_π respectively[12]. We apply an exponential specification for Ω since this parameter must be positive by definition.[13]

Our estimation strategy uses the recursive feature of our model. The parameters Ω and π can be estimated from the tenure distribution. These parameters can then be used to calculate conditional expectations and variances of the change in surplus $\Delta\Omega_\tau$, in both completed and incomplete job spells. These expressions are then used for the analysis of wage dynamics, for which we use the method of moments. In the next subsection, we show how the parameters π and Ω can be estimated by maximum likelihood from the tenure distribution. Next, we derive a set of moment conditions that (over)identify the parameters $\mu_0, \sigma_z^2, \overline{\sigma}, \gamma$and σ_u^2.

3.2 Maximum likelihood estimation of Ω and π

Ω and π are estimated by maximum likelihood, using the density function (11). Since experience at job start enters the analysis in deviation from its mean across jobs, the

[11] We do not consider also random job effects for both theoretical and empirical reasons. From a theoretical point of view, given our assumption of a frictionless market for alternative job opportunities, 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.

[12] For the third line in (15): since $\mu_z = \mu_r + \gamma\mu_p = \mu_r + \gamma\overline{\sigma}\pi$ and since μ_r does not depend on $\widetilde{E}$, we have $\partial\mu_z/\partial\widetilde{E} = \gamma\overline{\sigma}\partial\pi/\partial\widetilde{E}$. The same argument applies for e_π.

[13] Job tenures are determined by the first passage time distribution for Ω_τ becoming negative. Hence, $\Omega_0 = \Omega > 0$.

intercepts can be interpreted as the mean value for $\ln \Omega, \pi$ and μ_z respectively. We assume the two random effects e_π and e_Ω to be uncorrelated. Then, the contribution to the log likelihood function for an individual reads:

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

where j is the j^{th} job held by the worker, and 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:

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

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[14], but we do not know whether there

[14]Initial tenures are either reported or inferred by making them consistent with the latest reported

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 (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.[15] 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:

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

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

tenures –see Altonji and Williams (1999 and previous working versions).

[15] 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 than can be justified from the distribution of τ_j for jobs started after the first wave.

of the PSID, can thus be written as:

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

where $I(y)$ is the indicator function, taking value 1 if y is true and value 0 otherwise. We report results both 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 they are first observed in the PSID[16].

3.3 Moment Conditions for Wage Dynamics

Conditional on the maximum likelihood estimates of the parameters Ω and π, the conditional expectations and variances of $\Delta\Omega_\tau$ can be computed, accounting for the effect of S and e_Ω, see equation (15). These expressions can be used for the estimation of the parameters $\overline{\sigma}, \gamma, \mu_z, \sigma_z$, and σ_u by the method of moments, using data on (within-job and respectively, between-job) wage changes. We use the first two moments. We start with the conditions for the first moment of Δw_t. First, consider within-job wage changes. Taking expectations in equation (14) yields:

$$\mathrm{E}\left[\Delta w_t | A(\Theta)\right] = \mu_0 + \gamma\overline{\sigma}\pi_1 S + (1-\gamma)\overline{\sigma}\mathrm{E}\left[\Delta\Omega_\tau | A(\Theta)\right], 1 \leq \tau \leq \Theta \tag{20}$$

where we use $\mathrm{E}[\mu_z] = \mu_0 + \gamma\overline{\sigma}\pi_1 S$, see equation (15). This equation applies for completed spells; replacing the condition $A(\Theta)$ by $B(\Psi)$ yields the equation for incomplete spells. Next, consider between job wage changes. The change in log wages at the moment of job

[16] 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.

switch satisfies:[17]

$$\mathrm{E}[w_S - w^*_{T-1}|t=T] = \mu_0 + \gamma\overline{\sigma}\pi_1 S + (1-\gamma)\overline{\sigma}\mathrm{E}\left[\Delta\Omega^*_\Theta|A(\Theta^*)\right] + \overline{\sigma}\mathrm{E}[\Omega|S] \qquad (21)$$

where the superscript $*$ indicates that we refer to the previous job, and where we use $\mathrm{E}[\Omega|S] = e^{\omega_0+\omega_1 S+\frac{1}{2}\sigma^2_\Omega}$ in the last term, see equation (15). $\mathrm{E}[\Delta\Omega^*_\Theta|A(\Theta^*)]$ is the surplus change in the old job during the last year before separation. It is always negative, see Figure 2. The term $\overline{\sigma}\mathrm{E}[\Omega|S]$ is due to the starting surplus in the new job, since $\Omega_0 = \Omega$.

We turn now to the conditions for the second moments of Δw_t. Taking the variance in the second equation of (14) for within job wage changes yields:

$$\mathrm{Var}\left[\Delta w_t|A(\Theta)\right] = \sigma^2_z + \gamma^2\overline{\sigma}^2\sigma^2_\pi + (1-\gamma)^2\overline{\sigma}^2\mathrm{Var}\left[\Delta\Omega_\tau|A(\Theta)\right] + 2\sigma^2_u, 1 \leq \tau \leq \Theta \qquad (22)$$

since $\Delta z_t, \Delta b_t, u_t$ and u_{t-1} are all mutually uncorrelated. Again, if we replace $A(\Theta)$ by $B(\Psi)$ above, we obtain the variance condition for incomplete spells. The variance for between job wage changes reads similarly:

$$\mathrm{Var}[w_S - w^*_{T-1}|t=T] = \sigma^2_z + \gamma^2\overline{\sigma}^2\sigma^2_\pi + (1-\gamma)^2\overline{\sigma}^2\mathrm{Var}\left[\Delta\Omega^*_\Theta|A(\Theta^*)\right] + \overline{\sigma}^2\mathrm{Var}[\Omega|S] + 2\sigma^2_u \qquad (23)$$

where we use $\mathrm{Var}(\Omega|S) = (e^{\sigma^2_\Omega} - 1)e^{2(\omega_0+\omega_1 S)+\sigma^2_\Omega}$, see equation (15), and where the final term accounts for the variability in Ω due to unobserved worker characteristics. Finally, the presence of transitory shocks makes the following moment condition informative :

$$\mathrm{Cov}[\Delta w_t, \Delta w_{t-1}|A(\Theta)] = \gamma^2\overline{\sigma}^2\sigma^2_\pi - \sigma^2_u \qquad (24)$$

since $\Delta z_t, \Delta z_{t-1}, \Delta b_t, \Delta b_{t-1}, u_t, u_{t-1}$ and u_{t-2} are all mutually uncorrelated; again, by replacing $A(\Theta)$ with $B(\Psi)$ above we get the covariance condition for incomplete spells. The

[17] w^*_{T-1}represents the last wage that we observe in the previous job. The job separation happens at T, but we do not observe w^*_T; instead, the next wage we observe for the worker is her starting wage in the new job, w_S. We have: $w_S - w^*_{T-1} = w_S - w^*_T + \Delta w^*_T = w_S - r_S + \Delta w^*_T$

equations (20), (21), (22), (23), and (24) offer five moment conditions, two first moments and three second moments. The identification of μ_0 fully depends on the conditions for the first moments, while the identification of σ_z^2 and σ_u^2 fully depends on the conditions for the second moments. The identification of $\overline{\sigma}$ and γ depends on the conditions for both the first and the second moments. The degree of concavity of the wage profile, which affects the first moment, depends on the quantitative importance of selectivity, that is, on the second moment. The equivalence of the estimates from data on complete and on incomplete job spells, and from data on job stayers and job movers can be used to test the model. For the parameters $\overline{\sigma}$ and γ that appear in the conditions for both the first and the second moment, we shall also use the equivalence of these estimates as a test. Details of the estimation procedure for the system of equations (20)-(24) are discussed below, in the empirical analysis section for the wage dynamics and in the corresponding Appendix C.

4 Empirical Analysis

4.1 The Data

We use as data a PSID extract of 18 waves, covering the years 1975 through 1992, the 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, 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 (we 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 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 or Hawaii. Finally, we discard all observations on unionized jobs.[18]. Through these initial data selection procedures we discard in total 10351 observations from the original dataset used by 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. As tenure and experience we use the measures constructed by Altonji and Williams (1999). 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

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

afterwards (left censoring). 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

The estimation results of our empirical analysis on the tenure distribution are presented in Table 2.

Theoretically, the results for the two likelihood functions derived above, (18) and respectively (19), should coincide. 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 by inspecting Table 2, where the estimated intercepts and the coefficients for experience at job start are very similar in the small and large sample, for both Ω and respectively π. The positive effect of experience on the drift π is 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**
(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	

$^{(1)}$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% . Statistical t-values in parentheses under estimated coefficients.

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 fully observed shorter spells. Since the long spells started before the first wave of our PSID sample contain crucial information, we focus on the estimation results obtained from the full sample of job spells in the subsequent wage dynamics analysis. For the analysis of wage dynamics we have to delete observations for which wages are missing (2404 obs.) or topcoded (254 obs.), as well as observations for which $|\Delta w_t| > 0.50$ (276 obs.), since wage changes of this magnitude are likely due to be due to reporting error, see e.g. Fehr and Goette (2005, p. 785). This data cleaning procedure leaves a total of 8082 observations on within-job wage changes, and 462 observations on between-job wage changes. We deal with inflation by deflating wages with the implicit price deflator, using 1982 as base year as in Altonji and Williams (1999).

4.3 Wage Dynamics

Equations (20)- (24) can be used[19] for the formulation of an empirical regression model for the moments of Δw_t, as the following system of five equations (25):

$$\begin{aligned}
\Delta w_t &= \mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\tau + \varepsilon_t \qquad (25)\\
w_t - w_{t-1}^* &= \mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\Theta^* + \overline{\sigma}\Omega_0 + \zeta_t\\
\Delta w_t^2 &= \sigma_z^2 + 2\sigma_u^2 + \overline{\gamma}^2\overline{\sigma}^2\mathrm{Var}\Delta\Omega_\tau + \left(\mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\tau\right)^2 + \eta_t\\
\left(w_t - w_{t-1}^*\right)^2 &= \sigma_z^2 + 2\sigma_u^2 + \overline{\gamma}^2\overline{\sigma}^2\mathrm{Var}\Delta\Omega_\Theta^* + \left(\mu_0 + \gamma\overline{\sigma}\pi S + \overline{\gamma\sigma}\mathrm{E}\Delta\Omega_\Theta^* + \overline{\sigma}\Omega_0\right)^2 + \nu_t\\
\Delta w_t \Delta w_{t-1} &= -\sigma_u^2 + \left(\mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\Delta\Omega_\tau\right)\left(\mu_0 + \gamma\overline{\sigma}\pi_1 S + \overline{\gamma\sigma}\Delta\Omega_{\tau-1}\right) + \upsilon_t
\end{aligned}$$

[19] Note that some terms vanish from the second moment equations (22), (23), and (24) when implemented in the empirical system (25), as we have estimated $\sigma_\pi^2 = \sigma_\Omega^2 = 0$ earlier in the tenure distribution analysis.

where ε_t, ζ_t, η_t, ν_t, and υ_t are i.i.d. errors and where $\mathrm{E}\Delta\Omega_\tau \equiv \mathrm{E}[\Delta\Omega_\tau | A(\Theta), B(\Psi)]$, $\mathrm{Var}\Delta\Omega_\tau \equiv \mathrm{Var}[\Delta\Omega_\tau | A(\Theta), B(\Psi)]$, and $\overline{\gamma} \equiv 1 - \gamma$. We impose no constraints upon the covariance matrix of these five error terms. In the third and the fourth line we use $\mathrm{Var}[\Delta w_t] = \mathrm{E}[\Delta w_t^2] - \mathrm{E}[\Delta w_t]^2$, where we substitute $\mathrm{E}[\Delta w_t]$ for the deterministic part of the right hand side of the first and the second equation respectively, while in the fifth line we do the analogue substitution for the covariance moment. This system of equations is characterized by additive disturbances and nonlinear cross-equation restrictions in the parameters. It is estimated by Feasible Generalized Nonlinear Least Squares (FGNLS). Details of the estimation procedure are relegated to Appendix C.

The results for the FGNLS estimation of the system of equations (25) are reported in Table 3 below. The equations in (25) 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. We report estimation results for a number of subsamples, each in a different horizontal panel. 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 option than by negative shocks to the current job. 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 this low correlation might be downward rigidity in wages. If there is downward rigidity, the declining part of the wage profile for a complete spell, see e.g. Figure 2, might not be realized. 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 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, $E\Delta\Omega^*_\Theta$. We expect its coefficient to be $(1-\gamma)\,\overline{\sigma}$. 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. Hence, the significance of all estimates goes up and also γ is significant in this specification. The difference in the estimated value of γ in panels 1. and 3. suggests that downward rigidity plays a role indeed. Later on, we present a formal Wald test of this hypothesis.

Table 3: FGNLS system equations 25, with experience profile t

	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.

The range 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 value for γ implies 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 (14). Realized job changes are a highly selective sample from the total set of outside options that is available to the worker during the worker's tenure in her previous jobs. Only the highly favourable options are realized. The return to tenure measured as the rise in log productivity in the current job p_t, is thus even smaller, between $(1-\gamma)\overline{\sigma}\pi = (1-0.73)\times 0.04\times 0.14 = 0.15\%$ and $(1-0.51)\times 0.04\times 0.14 = 0.27\%$. To our knowledge, this paper is the first to account for selectivity in the realised outside options r_t. Apart from this true return to tenure, which is assumed to be linear, 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, c.f. Figure 2, while incomplete spells yield an increasing concave pattern for Ω_τ, c.f. Figure 3. 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.

Table 4: OLS first moments (first two equations) of system 25

	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 first moment eq. for job stayers (1st equation of system 25), 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 first moment equation for job movers (2nd equation of system 25)

The estimation results in Table 3 use all available information on first and second moments of wage changes simultaneously. An eyeball approach to specification testing is to estimate the model separately for relevant subsets of the data, e.g. job stayers versus job movers, complete versus incomplete spells, and first versus second moments. Later on we present formal Wald tests of these specification tests.

Table 4 presents the linear regressions for the first moments, i.e. the first two equations in the system (25), for various subgroups. The regression in column D and G, complete spells with an increasing surplus ($\Delta\Omega_\tau \geq 0$) and job movers respectively, are badly identified due to a low number of observations, so we do not consider these results. 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 significant, while it is expected to be positive. This term captures the fact that the estimation results on the tenure distribution show that the drift in the surplus Ω_t depends positively on experience at job start S, which is equivalent to the statement that jobs starting at later age last longer. Given the fact that $\gamma > 0$, the model predicts that workers are able to capture part of the increase in the surplus, and hence, the tenure profile in jobs starting at a higher age should be steeper. This eyeball test suggest this not to be the case. We present later on a Wald test of the hypothesis that the value of γ estimated from the drift effect is the same as its value estimated from the term $\mathrm{E}\Delta\Omega_\tau$.

Table 5 presents the linear regression results for the second moment equations[20] in system (25). The estimation results invoke three observations. First, in contrast to the predictions of our model, the estimation results for the term $(1-\gamma)^2\,\overline{\sigma}^2$ 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

[20] We use residuals from the first moment equations as dependent variables in the individual estimations of the second moment equations, so that we do not have to include the expectation of the first moments as regressors in the second moments, as was necessary in the FGNLS system estimation.

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 where a continuum of outside offers is available at any time and 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.

Table 5: OLS second moments (last three equations) of system 25

	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 second moment for job stayers (3rd equation of system 25), where A- All Stayers, B- Completed Spells, C- Incomplete Spells, D- Completed Positive Surplus Change Spells, E- Completed Negative Surplus Change Spells, F- Incomplete plus Completed Positive Surplus Change Spells; Column G corresponds to the second moment for job movers (4th equation of system 25); Columns H to I correspond to the covariance moment (last equation of system 25), with H using All Stayers, while I using the Incomplete plus Positive Completed Surplus Change Spells.

Third, within the group of job stayers, the variance does not seem to be constant across subgroups either. The variance is the largest for the incomplete spells and the smallest for the complete spells with a declining surplus ($\Delta\Omega_\tau < 0$), and where 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. We present Wald tests for these three hypotheses later on.

Table 6: Hypotheses tests equality of coefficients nested models

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 of the paper; "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 model variant estimation.

Table 6 presents the Wald tests for all the hypotheses discussed above, displaying the χ^2 statistic and associated p-values. All the tests start from the full model (and using the whole number of observations on stayers and movers), except for the test of σ_z^2 differences in subsamples of job stayers (in panel 6 of the table, see below), which starts from a model where we allow σ_z^2 to be larger for job movers. In the first panel we present three Wald tests for the null hypotheses $\{\gamma_s = \gamma_m\}$, $\{\overline{\sigma}_s^2 = \overline{\sigma}_m^2\}$, and respectively the joint null

$\{\gamma_s = \gamma_m$ and $\overline{\sigma}_s^2 = \overline{\sigma}_m^2\}$, where s indexes job stayers and m indexes job movers. The tests reveal that we cannot reject the null of $\{\gamma_s = \gamma_m\}$, but we reject the nulls of $\{\overline{\sigma}_s^2 = \overline{\sigma}_m^2\}$, and the joint null $\{\gamma_s = \gamma_m$ and $\overline{\sigma}_s^2 = \overline{\sigma}_m^2\}$ at 5%, and respectively 10% confidence levels. This suggests that there is excess wage variance for job movers, and hence that our assumption of a Walrasian market for job alternatives is not respected in the data. We therefore re-estimate the system of equations (25), allowing for different variances for job stayers and job movers, $\overline{\sigma}_s^2 \neq \overline{\sigma}_m^2$, with the corresponding estimates displayed and discussed later on in this section, after the presentation of the next hypotheses tests.

In panel 2 of Table 6 we have a Wald test for the null of $\{\gamma_{neg} = \gamma_{rest}\}$, 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 all other observations. Our test shows that we fail to reject the null of γ being the same for the negative job spells and for the rest of the observations.

Panel 3 presents the test of a hypotheses test for $\{\gamma_{drift} = \gamma_{nondrift}\}$, 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 system (25). As expected from our discussion above, we strongly reject the null here.

Panel 4 presents the Wald test of the hypothesis $\{\gamma_{first} = \gamma_{\sec ond}\}$. This restriction cannot be rejected.

Panel 5 displays the result of the Wald for $\{\sigma_{z,s}^2 = \sigma_{z,m}^2\}$, testing for equality of the variance of the permanent shocks, σ_z^2, across movers and stayers. This null is rejected with a lot of statistical power, showing, as stated above, the non-Walrasian market for outside offers.

Given our result from panel 5, we now start from a model where we allow for different $\sigma_{z,s}^2$ and $\sigma_{z,m}^2$, and in panel 6 we test in subsamples of job stayers for the following null hypotheses: $\{\sigma_{z,s,inc}^2 = \sigma_{z,s,neg}^2\}$, $\{\sigma_{z,s,pos}^2 = \sigma_{z,s,neg}^2\}$, $\{\sigma_{z,s,inc}^2 = \sigma_{z,s,pos}^2\}$, and respectively the joint $\{\sigma_{z,s,inc}^2 = \sigma_{z,s,neg}^2$ and $\sigma_{z,s,inc}^2 = \sigma_{z,s,pos}^2\}$, 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_z^2 between movers and stayers (see panel 5), there are no further statistical differences between these estimates for subsamples of stayers: indeed we cannot reject any of the null hypotheses from panel 6.

Once we allow for movers to have a different wage variance, as revealed by our test in panel 5 of Table 6. above, the new FGNLS estimates for system (25), displayed in the following Table 7 show that movers have a much higher wage variance at job separation, with our favorite specification in panel 3 suggesting $\overline{\sigma}_m \simeq 8.4\overline{\sigma}_s$. This time γ is estimated in a much narrower range, between 0.81 and 0.85, while all the other parameters are estimated close to the values from Table 3.

Table 7: FGNLS system equations 25, with experience profile t, and different wage variance 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.

If we repeat thus the earlier computation for the apparent and true tenure profile, but we do this only for stayers, fixing $\overline{\sigma}_s = 0.04$, and $\gamma = 0.8$, we get that the "apparent" return to 1 year spent on the job is exactly the same as computed above, $\overline{\sigma}_s\pi = 0.04 \times 0.14 = 0.56\%$, but that most of it is due to selection on the outside option, with the "true" tenure profile only $(1-\gamma)\overline{\sigma}_s\pi = (1-0.8) \times 0.04 \times 0.14 = 0.11\%$. Hence, once we adjust our empirical model for downward rigidity of wages and for the non-Walrasian market for job switches, we show that about 80% of the wage returns to tenure is due to selectivity on the outside wages.

5 Conclusions

We have 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 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. 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. There are two shortcomings of our basic model, that we are able to account for in the empirical estimation: our model does not incorporate the fact that wages are downward rigid, an indication of the failure of the efficient bargaining hypothesis; and, our assumption of frictionless market for job alternatives at job switching is proven incorrect, given excess wage variance estimated for job movers. Once we correct empirically for these misfits, we find that about 80% of our estimated tenure profile is accounted for by the selectivity in the outside productivity,

6 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

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

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

Fehr, E. and L. Goette (2005), "Robustness and real consequences of nominal wage rigidity", *Journal of Monetary Economics* 52, 779-804

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

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

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

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

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

A Derivations for I and ψ

A.1 Solution to Differential Equations and Investment Function

We follow the derivation in Dixit and Pindyck (1994). The solution to the differential equations (6) reads

$$\begin{aligned}
j(b_t) &= \left(\rho-\mu_p-\frac{1}{2}\sigma_p^2\right)^{-1}\exp(b_t)-\left(\rho-\mu_r-\frac{1}{2}\sigma_r^2\right)^{-1}+A^-\exp\left(\lambda^-b_t\right)\\
v(b_t) &= A^+\exp\left(\lambda^+b_t\right)\\
\lambda^\pm &\equiv \frac{-\left(\mu+\sigma_{pr}-\sigma_r^2\right)\pm\sqrt{\left(\mu+\sigma_{pr}-\sigma_r^2\right)^2+4\sigma^2\left(\rho-\mu_r-\frac{1}{2}\sigma_r^2\right)}}{\sigma^2}
\end{aligned}$$

where A^- and A^+ are constants of integration. Hence, the value matching and smooth pasting conditions (7) can be written as:

$$\begin{aligned}
A^+\exp\left(\lambda^+b^S\right)+I &= \left(\rho-\mu_p-\frac{1}{2}\sigma_p^2\right)^{-1}\exp\left(b^S\right)-\left(\rho-\mu_r-\frac{1}{2}\sigma_r^2\right)^{-1}+A^-\exp\left(\lambda^-b^S\right)\\
A^+\exp\left(\lambda^+b^T\right) &= \left(\rho-\mu_p-\frac{1}{2}\sigma_p^2\right)^{-1}\exp\left(b^T\right)-\left(\rho-\mu_r-\frac{1}{2}\sigma_r^2\right)^{-1}+A^-\exp\left(\lambda^-b^T\right)\\
A^+\lambda^+\exp\left(\lambda^+b^S\right) &= \left(\rho-\mu_p-\frac{1}{2}\sigma_p^2\right)^{-1}\exp\left(b^S\right)+A^-\lambda^-\exp\left(\lambda^-b^S\right)\\
A^+\lambda^+\exp\left(\lambda^+b^T\right) &= \left(\rho-\mu_p-\frac{1}{2}\sigma_p^2\right)^{-1}\exp\left(b^T\right)+A^-\lambda^-\exp\left(\lambda^-b^T\right)
\end{aligned}$$

or in matrix notation as:

$$\begin{bmatrix} \rho_r^{-1} \\ \rho_r^{-1} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} D & D^- & -D^+ & -1 \\ 1 & 1 & -1 & 0 \\ D & \lambda^- D^- & -\lambda^+ D^+ & 0 \\ 1 & \lambda^- & -\lambda^+ & 0 \end{bmatrix} \begin{bmatrix} B \\ C^- \\ C^+ \\ I \end{bmatrix}$$

where:

$$\begin{aligned} \rho_r &\equiv \rho - \mu_r - \frac{1}{2}\sigma_r^2 \\ B &\equiv \left(\rho - \mu_p - \frac{1}{2}\sigma_p^2\right)^{-1} \exp\left(b^T\right) \\ C^- &\equiv A^- \exp\left(\lambda^- b^T\right), C^+ \equiv A^+ \exp\left(\lambda^+ b^T\right) \\ D &\equiv \exp(\Omega), D^- \equiv \exp\left(\lambda^- \Omega\right), D^- \equiv \exp\left(\lambda^+ \Omega\right) \end{aligned}$$

Solving this system for I yields:

$$I = \rho_r^{-1} \frac{D\left(D^- \lambda^+ - D^+ \lambda^-\right) - D^- D^+ \left(\lambda^+ - \lambda^-\right) + D\left(D^+ - D^-\right)\lambda_0 \lambda_1}{D\left(\lambda^+ - \lambda^-\right) + D^- \lambda^- - D^+ \lambda^+ + \left(D^+ - D^-\right)\lambda^- \lambda^+} - \rho_r^{-1} \equiv I(\Omega; \mu, \Sigma) \tag{26}$$

One can show that $I(\Omega; \mu, \Sigma)$ is increasing in Ω, and hence its inverse $\Omega = \Omega(I : \mu, \Sigma)$ exists.

A.2 Derivation of the worker's asset value and the parameter ψ

Let $Q(p_t, r_t)$ 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, $Q(\cdot)$ can be written as:

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

leaving out the argument of $q(\cdot)$ in the second line. The first term in the second line captures the wage at the job, compare equation (12), while the second term captures the outside wage. Note that we have divided both sides of the equation by $\exp(r_t)$, so that the term for the outside wages is equal to unity. Efficient bargaining implies that it is optimal for the worker to quit when $b_t = b^T$. Hence, value matching and smooth pasting conditions must apply: $q\left(b^T\right) = 0$, $q'\left(b^T\right) = 0$. The solution to the differential equation has two constants of integration, one of which is zero due to the transversality condition. Hence, the solution to the differential equation and the two boundary conditions determine the constant of integration and ψ. Finally, the cost of specific investment are equal to $\exp\left(r^S\right) I$. 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. Hence:

$$q\left(b^S\right) = \text{ workers' share } \times I$$

B Conditional Expectation and Variance of Ω_τ

B.1 Completed job 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 (10) and (11), 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 $A(\Theta)$ and $\Omega_0 = \Omega$. 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 (9) in the above yields:

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

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

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

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

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

For the calculation of the second moment of a first differential of Ω_τ, $\mathrm{E}[\Delta\Omega_\tau^2|A(\Theta)]$, we apply the joint density that $\Omega_{\tau-1}=\omega$ and $\Omega_\tau=\omega+\chi$ for $1\leq\tau\leq\Theta$ conditional on $A(\Theta)$. This density is equal to the density that $\Omega_{\tau-1}=\omega$ conditional on $A(\Theta)$ times the density that $\Omega_\tau=\omega+\chi$ conditional on $\Omega_{\tau-1}=\omega$ and $\Omega_\Theta=0$:

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

The second moment of $\Delta\Omega_\tau$ will thus be 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 evaluating this integral, since its analytical solution is highly complicated. The variance is then derived by the standard expression[21]:

$$\mathrm{Var}\left[\Delta\Omega_\tau|A\left(\Theta\right)\right] = \mathrm{E}\left[\Delta\Omega_\tau^2|A\left(\Theta\right)\right] - \mathrm{E}\left[\Delta\Omega_\tau|A\left(\Theta\right)\right]^2$$

B.2 Incomplete job spells

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

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

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

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

where $\overline{F}(\Psi-\tau,\omega)$ is given by equation (10), substituting appropriately the parameters of the function. 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 $1 \le \tau \le \Psi$, conditional on $B\left(\Psi\right)$ is then derived from the first and second moments of $\Delta\Omega_\tau$, analogous to the completed spells case discussed above. The second moment reads:

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

[21] The first moment is straigtforward using the analytical solution derived for $\mathrm{E}(\Omega_\tau|A\left(\Theta\right))$:

$$\mathrm{E}\left[\Delta\Omega_\tau|A\left(\Theta\right)\right] = \Delta(\mathrm{E}(\Omega_\tau|A\left(\Theta\right))$$

C Estimation procedure for the system (25)

The system of equations (25) can be written in matrix notation as:

$$y = X(\beta) + u \tag{27}$$

where y is a vector of the five dependent variables; u is a corresponding vector of error terms, with a different variance for each of the five components and cross-equation correlations; β is the vector of parameters; $X(\beta)$ is the matrix of functions of exogenous regressors and parameters. Equation (27) is estimated by feasible generalized nonlinear least squares (FGNLS). β is estimated as an iterative process: start with a value of $\widehat{\beta}$, then calculate $\widehat{\beta}^*$ as

$$\widehat{\beta}^* = \left[X_\beta\left(\widehat{\beta}\right)' X_\beta\left(\widehat{\beta}\right)\right]^{-1} X_\beta\left(\widehat{\beta}\right)' \left[y - X\left(\widehat{\beta}\right) + X_\beta\left(\widehat{\beta}\right)\widehat{\beta}\right]$$

where $X_\beta(\beta)$ is the matrix of partial derivatives of $X(\beta)$ with respect to β; repeat the procedure using $\widehat{\beta}^*$ as the new starting point, until $\widehat{\beta}^*$converges to $\widehat{\beta}$. The first stage estimation allows the calculation of $\widehat{u} = y - X\left(\widehat{\beta}\right)$, which enables us to compute an estimate of the variance of each of the five components of u. Let V be the diagonal matrix containing for each of the five components the estimated variance of the error term on the diagonal. Then, the second stage estimator for β requires a similar iterative process as above until $\widehat{\beta}^*$ converges to $\widehat{\beta}$, but using V instead of the identity matrix:

$$\begin{aligned}
\widehat{\beta}^* &= \left[X_\beta\left(\widehat{\beta}\right)' V^{-1} X_\beta\left(\widehat{\beta}\right)\right]^{-1} X_\beta\left(\widehat{\beta}\right)' V^{-1} \left[y - X\left(\widehat{\beta}\right) + X_\beta\left(\widehat{\beta}\right)\widehat{\beta}\right] \\
\text{Var}\left(\widehat{\beta}^*\right) &= \left[X_\beta\left(\widehat{\beta}\right)' V^{-1} X_\beta\left(\widehat{\beta}\right)\right]^{-1}
\end{aligned}$$