

IZA DP No. 1997

Tenure Profiles and Efficient Separation in a Stochastic Productivity Model

Sebastian Buhai
Coen Teulings

March 2006

Tenure Profiles and Efficient Separation in a Stochastic Productivity Model

Sebastian Buhai

*Tinbergen Institute, Erasmus University Rotterdam
and Aarhus School of Business*

Coen Teulings

*SEO Economic Research, University of Amsterdam,
Tinbergen Institute and IZA Bonn*

Discussion Paper No. 1997
March 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Tenure Profiles and Efficient Separation in a Stochastic Productivity Model*

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

JEL Classification: C51, C52, J63

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

Corresponding author:

Sebastian Buhai
Aarhus School of Business
Department of Economics and CCP
Prismet, Silkeborgvej 2
DK-8000, Aarhus C
Denmark
Email: sebu@asb.dk

* For helpful comments and suggestions we are grateful to Rob Alessie, Gadi Barlevi, Richard Blundell, Jeff Campbell, Bas van der Klaauw, Jim Heckman, Dale Mortensen, Derek Neal, Randall Olsen, Miguel Portela, Robert Shimer, Chris Taber, Aico van Vuuren, Robert Waldmann, 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, Tinbergen Institute, and in several conferences and workshops in Europe and the USA. We also wish to thank Nick Williams and Lennart Janssens for clarifying questions about the data. The usual disclaimers apply.

1 Introduction

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

We take efficient bargaining as benchmark. Hence, quits and job layoffs are observationally equivalent, as in McLaughlin (1991). The model explains the observed correlation between wages and job tenure from the random evolution of wages after job start. This random evolution of wages is due to 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 below the value of the outside option, which depends on the initial productivity at the job start in the best alternative job. We refer to this initial productivity in the best alternative as the "outside productivity" and to the productivity in the job as the "inside productivity". The observed correlation between wages and tenure is caused by the fact that only jobs that evolve 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) and Topel and Ward (1992), hypothesis that we verify on our PSID estimation sample. Whereas this observation received little attention among labor economists, we take it as cornerstone of our modelling. Both log in- and outside productivity are assumed to follow a random walk. Our model implies that log wages are a linear combination of both, which implies that log wages in the job follow a random walk as well. Hence, the difference in the drift between the log wage in the job and the log outside productivity is what we traditionally call "the return to tenure".

Each job requires some form of specific investment. That specificity can be anything. Training is just one aspect. For most jobs other aspects are more important, such as getting to know your new colleagues, knowing where to make photocopies or where to get a cup of coffee, organizing your home-to-work travel efficiently etc. Upon separation the worker and the firm lose the value of these specific investments. Since separation is irreversible, the investments have an option value. The combina-

tion of irreversible specific investment and inside and outside productivities following random walks implies that we can apply the theory of real options, see for example Dixit (1989), Bentolila and Bertola (1990) and Dixit and Pindyck (1994). Option theory allows us to calculate the reservation value of productivity for which separation becomes the efficient alternative. Teulings and van der Ende (2000) use this model for the analysis of distribution of job tenures. The predicted hazard rates of this model are well in line with the empirical distribution of the job exits. Our model shares some features with Mortensen's (1988) dual "on-the-job-training and matching" model, starting from the same basic idea of jointly modelling optimal separation behavior and individual wage evolution over time and providing qualitatively similar empirical implications for the job tenure distribution. One of the essential differences is the fact that our "match quality" (surplus of the job's productivity over the worker's outside option) evolves randomly and the intensity of the shocks to the match quality does not diminish with time as in a Bayesian learning framework.

From the distribution of job tenures we are able to estimate the surplus of the job's productivity above its reservation value and a (linear) drift of this surplus, up to a normalizing constant (the variance of the random walk). We obtain a positive drift surplus, indicating that some 10% of all jobs will end only by retirement. We use these parameters to compute the expected surplus in both completed and incomplete job spells, which will enable us to estimate the evolution of wages. The typical problem in this literature is that the researcher observes the outside productivity only at job start and at job separation, if the worker starts 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. The initial wage in this job is therefore a linear function of the value of the outside productivity at job start. Applying the same reasoning, the first wage in the next job is a function of the outside productivity at the moment of separation from the previous job. Our estimation procedure exploits both pieces of information on the outside productivity

to the maximum. To that end, we elaborate an idea first explored by Abraham and Farber (1987): we condition the expected wage growth within a job not only on the elapsed duration since job start -that is: current job tenure- but also on the remaining time span left till the next separation, so we take into account the complete duration of the job. We can calculate a closed form expression for this expectation. As a first result, we show that this expression does not depend on the drift surplus. This implies that the evolution of wages in completed spells is uninformative on the return to tenure. This is a remarkable conclusion given the fact that so many papers have tried to identify the return to tenure from this type of data. The only sources of information on the return to tenure are the distribution of completed tenures and the evolution of wages in incomplete job spells. The fat right tail in the tenure distribution, with many jobs never ending, is an indication of large returns to tenure: the return to tenure is so high that separation is rarely efficient, except for cases where the random walk evolves really unfavorably.

As a second result, we demonstrate the fragility of the tenure profile identification. The problem is not so much the selectivity in the observed wage profile, as the selectivity in the outside productivity. Observed outside productivities are positively selected, since we observe them only at the moment when workers switch jobs and workers switch jobs only when the outside productivity is high. This source of selectivity usually receives less attention than the selectivity in the inside wage. We show that this effect can be identified from the wage change for job movers, but that this is a thin line of identification. Surprisingly, selectivity in the outside wage turns out to be an empirically important phenomenon; this selectivity provides a new source of a "tenure profile", accounting for about 85% of the tenure profile we estimate.

The empirical results show that our model does very well in explaining the concavity in the "observed" 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. Nevertheless, there is one strong test of our assumptions: the estimated variance of the innovation in wages is consistent with the concavity in the "observed" tenure profile. There is nothing in our estimation procedure that drives this result. Furthermore, we do observe a smaller wage increase during the first half of the job spell than at similar tenures for jobs that separate later or not at all, as predicted by the model. However, we do not observe wages falling during the second half of the job spell, as also predicted by the model. This fits the idea of downward rigidity, 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. Nonetheless, at the moment of separation, the gap is filled by an additional wage decline for job changers. Hence, our empirical results provide support for an amended version of the model, where we allow for downward rigidity in wages. This rigidity does not fit the efficient bargaining hypothesis. In unionized jobs, the fall in wages at the date of separation is much larger. It is hard to see how these results can be squared with Nash bargaining. The estimated tenure profile is on the high end of the spectrum, 5% per year, though more than five sixths of the return take the form of a declining outside productivity instead of a rising inside productivity. If we were to exclude this part of the profile, our estimates would be on the low end of the spectrum, 0.6% per year.

While we focus on firm tenure, our model could equally well be applied to industry or occupation tenure, as suggested by Neal (1995). Many specific skills are likely to be industry or occupation-specific and are thus irreversibly lost if the worker quits the industry or the occupation, but not if she switches from the one to the other firm. One would expect greater losses when switching between industries or occupations than when just changing jobs, as is reported by Neal (1995). We leave this extension for future research.

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

2 The Random Productivity Growth Model

2.1 Model Assumptions

Our modelling extends the analysis on the tenure distribution by Teulings and van der Ende (2002) to wages. Consider a labor market in continuous time, where a job is a unique match between a risk neutral worker and a risk neutral firm. The risk neutrality of both players implies that there is no insurance problem. We ignore any disutility of effort, so that the worker's utility depends only on her income. There is no search cost involved from either party in finding the optimal match. A worker picks therefore that vacancy that yields the highest expected net discounted value. At the start of the job, specific investments are made, which are irreversibly lost upon a separation between the worker and the firm. However, the firm retains the property right on the vacancy. That is, it can hire at any future time provided that the new worker and the firm are prepared to pay the cost of the specific investment again. These specific investments are verifiable, so that there are no hold-up problems: the worker and the firm can always agree on a side payment at the start of their relation that offsets expected unbalances in bargaining power. We shall relax this assumption later on, when discussing the impact of unions. Furthermore, the investments are made instantaneously and do not require any time for implementation. The log productivity of the job and the log outside wage evolve over time according to a random walk. Both worker and firm are perfectly informed about their current value, but their future evolution is unknown. The worker and the firm bargain over the surplus of the inside productivity over its reservation value. This bargaining is efficient: as long as there is a surplus, the worker and the firm will agree on a sharing rule. At some moment the productivity has fallen below its reservation value so that separation becomes the efficient alternative. Then, separation occurs at mutual consent since there are no gains from trade left. Turnover is therefore efficient, and quits and layoffs are observationally identical, as in McLaughlin (1991). For the sake of

convenience, we shall refer to separations as the firm firing the worker in the rest of the paper, though separations can be both quits or layoffs. Given these assumptions (risk neutrality, no hold up problems, and efficient bargaining), wage setting and separation decisions can be analyzed separately, since matching and separation decisions maximize the joint surplus, regardless of its precise distribution. This section focuses on the separation decision, wage setting being discussed in the next section.

We assume that job's productivity P_t follows a geometric Brownian with drift. The outside productivity R_t is also a geometric Brownian with drift. This outside productivity R_t is the initial productivity in the best job available at time t . Since individuals can costlessly pick this best alternative, $P_t = R_t$ at the moment of job start. The specific investments at the moment of job start are proportional to the reservation wage: $R_t I$. One can think of I as cost of investment measured in units of labor time and of R_t as the price of one unit. Using lower cases to denote the logs of the corresponding upper cases, the law of motion between arbitrary dates s and t , with $s < t$, is characterized by a bivariate normal distribution:

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

where:

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

Let $J(p_t, r_t)$ and $V(p_t, r_t)$ be the value of a job and a vacancy respectively, both as functions of log in- and outside productivity, p_t and r_t . The Bellman equations for both value functions read (see e.g. Dixit and Pindyck, 1994) :

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

where we leave out the arguments of $J(\cdot)$ and $V(\cdot)$ for convenience and where ρ denotes the interest rate. The first term in the expression for J is the current output of the job, 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. The value matching and smooth pasting conditions read:

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

where S is the moment of job start and T is the moment of separation. The first condition states that at the moment of job start (when by definition, $p_S = r_S$), the value of a filled job must be equal to the value of the vacancy plus the cost of specific investment. The second condition states that at the moment of job separation, the value of a vacancy in the current job is equal to the value of the current job minus the value of the vacancy to which the worker is moving; the latter term reflects the fact that by leaving his current job, the worker becomes available for filling another job. The last four conditions are the smooth pasting conditions. These conditions and the Bellman equations (2) jointly determine $J(\cdot)$ and $V(\cdot)$.

This system can be simplified to a system with only one state variable. Define:

$$B_t \equiv \frac{P_t}{R_t}$$

Since P_t and R_t follow a geometric Brownian, B_t also follows a geometric Brownian:

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

where again b_t denotes $\log B_t$. By construction, $B_S = 1$, $b_t = p_t - r_t$, and $b_S = 0$. Our claim is that the value functions $J(\cdot)$ and $V(\cdot)$ can be written as:

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

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

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

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

$$\begin{aligned}
j(0) &= v(0) + I \\
v(b_T) &= j(b_T) - v(0) \\
j'(0) &= v'(0) \\
v'(b_T) &= j'(b_T) - v'(0)
\end{aligned}$$

where we use $b_S = 0$. The proof of this claim is by straightforward substitution¹ of equation (5) in the Bellman equations (2) and the value matching and smooth pasting conditions (3).

The hiring and separation rule depend therefore purely on b_t : a vacancy should be filled at the first time S that b_t rises up to $b^h = 0$, a worker and a firm should separate the worker at the first time T that b_t falls below b^s . One can prove: $b^s < 0$. The value of b^h and b^s can be expressed as a function of the model's parameters I, μ, Σ and the interest rate ρ , using the value-matching and smooth-pasting conditions above, as described in detail by Dixit and Pindyck (1994). Since we do not need these expressions for the subsequent analysis, we do not present them here, see Teulings and Van der Ende (2000) for those.

2.2 Job Tenure Distribution

Without loss of generality we normalize the moment of job start to zero, $S = 0$, in what follows. Hence, $b_0 = b^h = 0$. From that moment on, b_t evolves according to its law of motion specified in equation (4). The separation occurs at the first moment $t = T$ when $b_t = b^s$. Hence, T is the completed tenure of that job spell. Analogously to the probit model, where the variance of the error term is non-identified and can therefore be normalized to unity without loss of generality, σ , the standard deviation of b_t , is unidentified in this model. We normalize therefore all other parameters by σ .

¹We use:

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

Define:

$$\begin{aligned}\Omega_t &\equiv \frac{b_t - b^s}{\sigma} \\ \Omega &\equiv \frac{b^h - b^s}{\sigma}, \Omega > 0 \\ \pi &\equiv \frac{\mu}{\sigma}\end{aligned}$$

Thus Ω_t is a Brownian with drift π and unit variance per unit time, with $\Omega_0 = \Omega$ and $\Omega_T = 0$, where T is a stochastic variable determined by the evolution of b_t . Ω_t is the surplus of the inside productivity p_t above its reservation value at which separation becomes efficient, relative to outside productivity r_t and normalized by the standard deviation of b_t per unit of time. Hence, we refer to Ω_t as the normalized surplus of a job. The distribution of job tenures is the "First Passage Time" distribution (e.g. Cox and Miller, 1965), the distribution of durations till the random walk Ω_t passes the single absorbing barrier $\Omega_t = 0$ for the first time. The density function of Ω_t , conditional on $\Omega_0 = \Omega$, reads:

$$\frac{1}{\sqrt{t}} \phi \left(\frac{\Omega_t - \Omega - \pi t}{\sqrt{t}} \right)$$

where $\phi(\cdot)$ is the standard normal PDF. However, a realization of Ω_t is not interesting if separation has occurred before time t . This situation cannot correspond to a job-worker match since separation decisions are irreversible. We are thus interested in the density of Ω_t conditional on the fact that no separation has occurred before time t , that is $\Omega_s > 0$ for all $0 \leq s < t$. A simple methodology can be applied in this regard, often used in pricing barrier options in mathematical finance, the stochastic reflection principle: there is a one-to-one correspondence between trajectories from Ω to Ω_t having crossed $\Omega_s = 0$ at least once, and trajectories from $-\Omega$ to Ω_t . This latter group of trajectories should be subtracted from the trajectories from Ω to Ω_t for the calculation of the density of all trajectories from Ω to Ω_t that never crossed

$\Omega_s = 0$. Let $g(\omega, t, \Omega)$ be the joint density/probability of $\Omega_t = \omega$ and $T > t$, so $g(\omega, t, \Omega) \equiv \Pr(\Omega_t = \omega \wedge T > t | \Omega)$ (we add the parameter Ω as an argument for future reference). This density satisfies:

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

where the factor $e^{-2\Omega\pi}$ corrects for the differential effect of the drift on the density for upward and downward trajectories. The cumulative distribution of jobs surviving at time t , $1 - F(t, \Omega)$, is calculated by integrating $g(\omega, t, \Omega)$ over $\Omega_t \in [0, \infty)$:

$$\begin{aligned} 1 - F(t, \Omega) &\equiv \Pr(T > t | \Omega) = \int_0^\infty g(\omega, t, \Omega) d\omega \\ &= \Phi \left(\frac{\Omega + \pi t}{\sqrt{t}} \right) - e^{-2\Omega\pi} \Phi \left(\frac{-\Omega + \pi t}{\sqrt{t}} \right) \end{aligned} \quad (7)$$

where $\Phi(\cdot)$ is the standard normal CDF. The distribution of completed job tenures is therefore fully specified by two parameters, the distance from the separation threshold at job separation Ω and the drift π . The corresponding density function is the derivative of $F(t, \Omega)$ with respect to t :

$$f(t, \Omega) = \frac{\Omega}{t\sqrt{t}} \phi \left(\frac{\Omega + \pi t}{\sqrt{t}} \right) \quad (8)$$

where we have used $\phi \left(\frac{\Omega + \pi t}{\sqrt{t}} \right) = e^{-2\Omega\pi} \phi \left(\frac{-\Omega + \pi t}{\sqrt{t}} \right)$.

The job exit rate is then given by $f(t, \Omega) / [1 - F(t, \Omega)]$. It is straightforward to check that the exit rate is hump shaped, starting from 0, reaching a peak at t^* , $0 < t^* < 2/3\Omega^2$, and afterwards either declining monotonically to 0 for positive drift $\pi > 0$ or to $1/2\pi^2$ for negative drift $\pi < 0$. 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 survivor function (7) when $t \rightarrow \infty$, hence by $1 - e^{-2\Omega\pi}$. We plot the exit rates for pairs

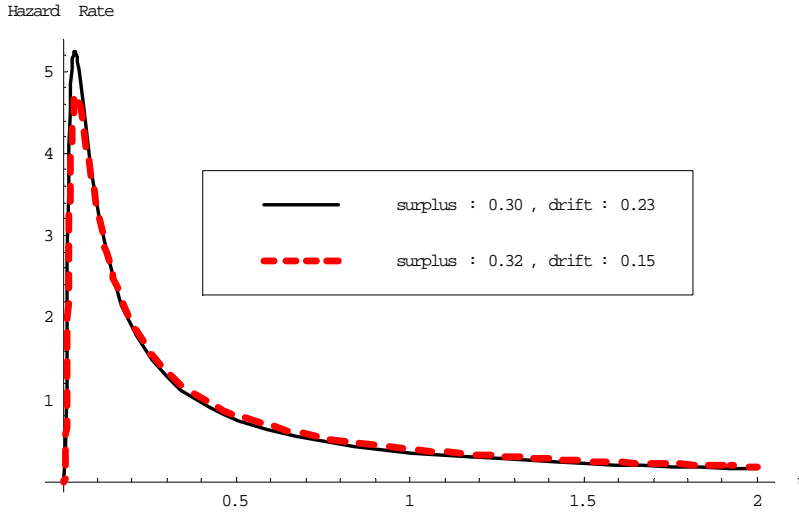


Figure 1: Predicted Job Hazards

$\Omega = 0.32$ and $\pi = 0.15$ and respectively $\Omega = 0.30$ and $\pi = 0.23$ (these are mean values for Ω and π from our estimations of the tenure distribution parameters, see Section 3 below) in Figure 1. In both cases the peak is reached at $t \simeq 0.04$ years. Since $\pi > 0$, the hazard rate converges to zero and a positive fraction of the jobs will never end. For mean values of parameters of $\Omega = 0.32$ and $\pi = 0.15$ about 10 % of the jobs never end. We conclude that parameters Ω and π can be identified from data on the distribution of job tenures, but parameter σ cannot.

2.3 Tenure Profile in Wages

2.3.1 Sharing Rule of Surpluses

We extend the model with an explicit sharing rule of surpluses during the course of the job spell. We use a rule stipulating that the surplus $b_t - b^s$ is shared in fixed

proportions between the worker and the firm². The worker's log wage w_t then satisfies:

$$w_t = r_t + b^s + \beta(b_t - b^s) = r_t + b^s + \bar{\sigma}\Omega_t \quad (9)$$

where $\bar{\sigma} \equiv \beta\sigma$. We interpret β as the worker's bargaining power. Equation (9) implies efficient separation, since at the moment of separation $t = T$, $b_T = b^s$, and hence, $w_T = r_T + b^s$. Since $b^s = p_T - r_T$, this implies $w_T = p_T$: log wages are equal to log inside productivity. For a slightly lower p_t , $w_t > p_t$ and firms prefer separation above continuation of the employment relation. Log wages within a job follow a Brownian with drift $\mu_r + \bar{\sigma}\pi$. The term $\bar{\sigma}\pi$ is the tenure profile. At the start of each job, $t = 0$ and $\Omega_t = \Omega$. After job start, Ω_t goes up in expectation with π every period, conditional on the fact that Ω_t remains positive. Were job separations independent of the realization of Ω_t (and of r_t), we would have:

$$\mathbb{E}(w_t | t < T) = \mathbb{E}(r_t) + b^s + \bar{\sigma}\pi t$$

In that case, the tenure profile would be estimated easily, for instance in discrete time by first differencing the equation above and then comparing log wage growth for job stayers and job changers:

$$\begin{aligned} \text{stayers} & : \quad \Delta \mathbb{E}(w_t | 1 < t < T) = \mu_r + \bar{\sigma}\pi \\ \text{changers} & : \quad \Delta \mathbb{E}(w_t^* | t = T) = \mu_r - \bar{\sigma}\pi (T - 1) \end{aligned}$$

where Δ is the first difference operator and where the superscript $*$ indicates that we compare log wages in the new and the old job; hence, Δw_T^* compares the starting wage in the new job to the wage one year before separation in the old job.

However, in completed job spells Ω_t is correlated to T for three reasons: (i) $\Omega_0 = \Omega$,

²This is more pragmatic than what is commonly used, ie. a sharing rule of the return on the expected discounted value of all future absolute surpluses, instead of the instantaneous relative surplus. Our approach can be viewed as a first order expansion of that rule.

(ii) $\Omega_T = 0$, and (iii) $\Omega_t > 0$ for $0 \leq t < T$. Empirically, r_t is only observed at the beginning of a job: $r_0 = w_0 - b^s - \bar{\sigma}\Omega$, and at the end of a job: $r_T = w_T - b_s$. In between, the researcher has no information on r_t , but only on w_t . Our strategy is to calculate $E(\Omega_t)$ conditional on the three pieces of information available, (i), (ii), and (iii), and to enter this expectation as a regressor in a regression of within job log wage growth. *Mutatis mutandis* the same applies to job spells that do not end before the end of the time span covered by the data, the incomplete job spells. Let L be the last date on which data are available. What we know about an incomplete job spell is that it is still running at L . Hence, there are again three pieces of information: (i) $\Omega_0 = \Omega$, (ii) $T > L > t$, and hence (iii) $\Omega_t > 0$ for $0 \leq t < L$. And again we calculate the conditional expectation $E(\Omega_t)$. Below we discuss the conditional expectation, first for completed and then for incomplete job spells.

2.3.2 Conditional Expectation of Ω_t for Completed Spells

Let $h(\omega, t, \tau)$ be the density of $\Omega_t = \omega$ conditional on (i) $\Omega_0 = \Omega$, (ii) $\Omega_\tau = 0$, and (iii) $\Omega_t > 0$ for $0 \leq t < \tau$. Comparing this density to $g(\omega, t, \Omega)$, there is one additional condition: $\Omega_\tau = 0$, or equivalently, $T = \tau$. When we want to apply Bayes's rule, we need the distribution of T conditional on $\Omega_t = \omega$. Since Ω_t is a martingale, the distribution of T conditional on $\Omega_t = \omega$ is equal to the distribution of $T = \tau - t$ conditional on $\Omega = \omega$. Hence, its density is $f(\tau - t, \omega)$, see equation (8). Then $h(\omega, t, \tau)$ can be calculated from $f(\cdot)$ and $g(\cdot)$ by Bayes's rule, see equations (6) and (8):

$$h(\omega, t, \tau) = \frac{f(\tau - t, \omega)g(\omega, t, \Omega)}{\int_0^\infty f(\tau - t, x)g(x, t, \Omega) dx} \quad (10)$$

The conditional expectation reads:

$$\begin{aligned}
\mathbf{E}(\Omega_t|0 < t < T = \tau) &= \int_0^\infty \omega h(\omega, t, \tau) d\omega & (11) \\
&= 2\sqrt{f(t)} t \phi\left(\sqrt{f(t)}/t\Omega\right) - \left(\frac{t}{\Omega} + f(t)\Omega\right) \left[1 - 2\Phi\left(\sqrt{f(t)}/t\Omega\right)\right] \\
f(t) &\equiv \frac{\tau - t}{\tau}
\end{aligned}$$

where τ denotes the realization of T . The derivation is discussed in Appendix A1. The remarkable feature of this expression is that it does not depend on the tenure profile in wages, π . Hence, conditional on the model that we specified, the evolution of wages in completed job spells does not provide any information at all on the tenure profile in wages. Given the many papers that have tried to do so, this is a staggering conclusion.

$\mathbf{E}(\Omega_t|0 < t < T = \tau)$ satisfies the following conditions:

$$\begin{aligned}
\lim_{t \rightarrow 0} \mathbf{E}(\Omega_t|0 < t < T = \tau) &= \Omega \\
\lim_{t \rightarrow T} \mathbf{E}(\Omega_t|0 < t < T = \tau) &= 0 \\
\lim_{t \rightarrow 0} \frac{d\mathbf{E}(\Omega_t|0 < t < T = \tau)}{dt} &= \frac{1}{\Omega} - \frac{\Omega}{\tau} \\
\lim_{t \rightarrow \tau} \frac{d\mathbf{E}(\Omega_t|0 < t < T = \tau)}{dt} &= -\infty \\
\frac{d^2\mathbf{E}(\Omega_t|0 < t < T = \tau)}{dt^2} &< 0
\end{aligned}$$

The full expression of derivatives as a function of t can be found in the Appendix A1. The first two lines above fit our assumptions that a job starts at $\Omega_0 = \Omega$ and ends at $\Omega_T = 0$. The third line says that the initial slope is negative for short spells, $T < \Omega^2$, and positive for longer spells. For short spells the expected surplus must decline immediately to reach $\Omega_T = 0$ in time. Therefore these spells are a selective sample of trajectories for which the expected surplus declines right from the start of the job spell. Correspondingly, the long spells $T > \Omega^2$ are the selective sample for which the

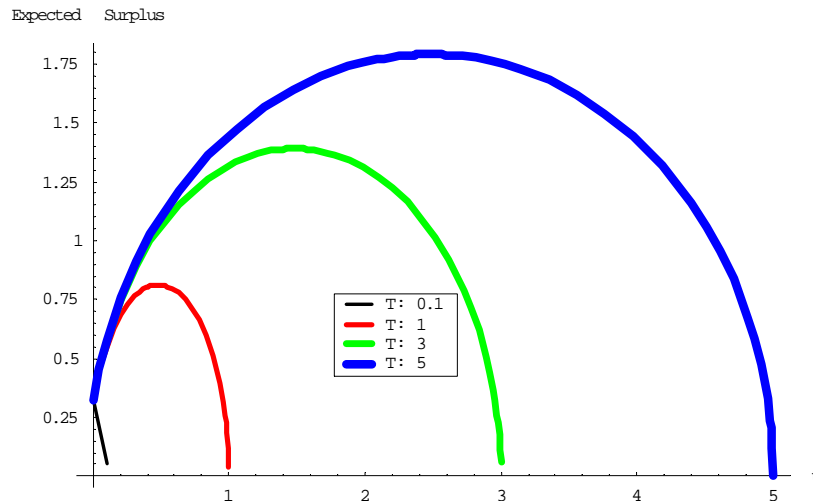


Figure 2: Expected Surplus in Completed Job Spells

opposite holds. The fourth line shows that the expected surplus declines infinitely fast just before separation. Trajectories that separate the next minute are therefore a highly selective sample. This result is consistent with the empirical evidence by Jacobson, LaLonde and Sullivan (1993) on the decline in the wage profile in the period just before the moment of separation. The final line above shows that the second derivative is always negative. Hence, the expected surplus is concave in t ; it is monotonically decreasing for short spells $T < \Omega^2$ and it is hump shaped for longer spells. The tenure profile is plotted for the estimated mean value $\Omega = 0.32$ and for various values of T in Figure 2. For $T \leq 0.1$ years the tenure profile is monotonically decreasing, while for larger T it is increasingly concave. The top of the profile is increasing in T , showing the importance of conditioning on the eventual tenure.

2.3.3 Conditional Expectation of Ω_t for Incomplete Spells

The conditional expectation for incomplete job spells, $E(\Omega_t | t < L < T)$, is calculated by using the same methodology as in the case of the completed job spells. Let $h^*(\omega, t, L)$ be the density of $\Omega_t = \omega$ conditional on (i) $\Omega_0 = \Omega$, (ii) $\Omega_t > 0$ for

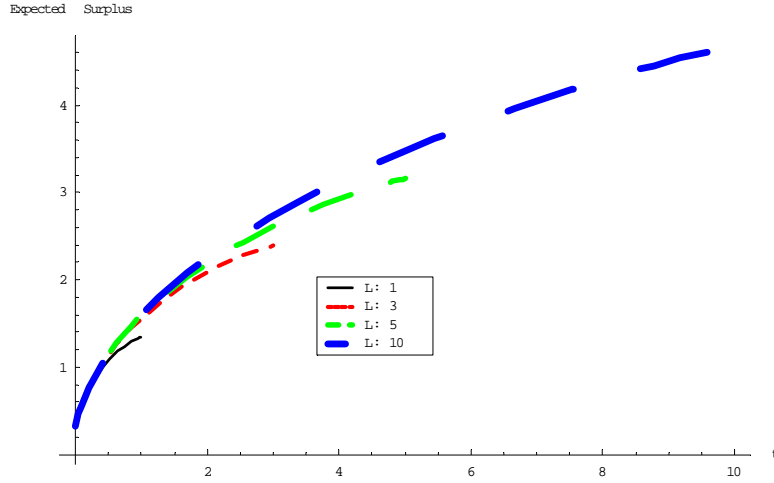


Figure 3: Expected Surplus in Incomplete Job Spells

$0 \leq t < L$, and (iii) $T > L$. The application of the Bayes rule leads to the following expression for the conditional density:

$$h^*(\omega, t, L) = \frac{[1 - F(L - t, \omega)] g(\omega, t, \Omega)}{\int_0^\infty [1 - F(L - t, x)] g(x, t, \Omega) dx} \quad (12)$$

This density can be used for the calculation of the conditional expectation in the same way as equation (11). Contrary to the case of completed spells, there is however no explicit expression for the conditional expectation in this case, cf. Appendix A2 for the final expressions on which numerical integration is performed. Figure 3 presents the trajectory of $E(\Omega_t | t < L < T)$ for $\Omega = 0.32$, $\pi = 0.15$ and $L = 1, 3, 5, 10$. The higher L , the more information on T is available, since $T > L$.

$E(\Omega_t | t < L < T)$ for a fixed t is increasing in L . The reason is that higher values of L imply a greater selectivity, since more and more trajectories leading to a separation have been selected out. Were there no selectivity, then the trajectory would be linear, $E(\Omega_t | t < L < T) = E(\Omega_t) = \pi t$. The trajectories are strongly concave, implying that selection plays an important role. Contrary to the completed spells case, incomplete

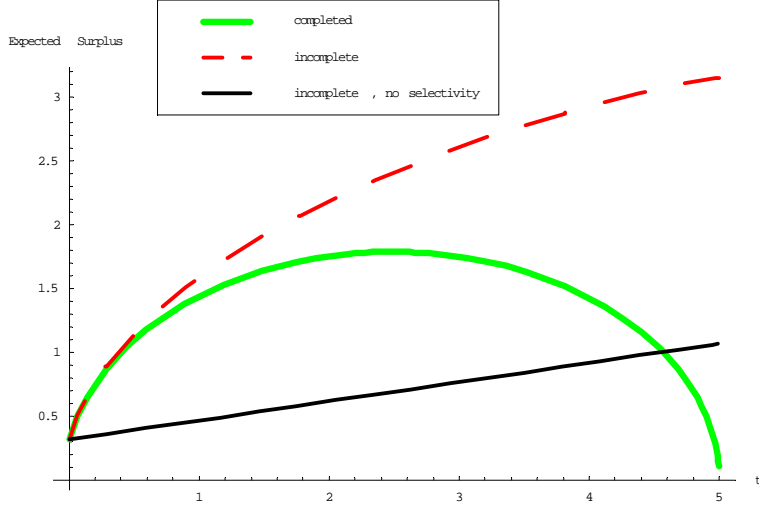


Figure 4: Selectivity versus Drift in the Expected Surplus

spells do provide information on the drift π . Nevertheless the impact of the drift is negligible compared to selectivity, as documented by Figure 4, which compares the trajectories of the conditional expectation in completed spells, in incomplete spells and in absence of any selectivity. The concavity outweighs the linear trajectory by far, at least for the first five years. We also plot conditional expectations of the surplus for mean values of the tenure distribution parameters in both completed and incomplete job spells, for very long job durations. Trajectories of the expected surplus for $T = 10, 20$ and respectively $L = 10, 20$ are plotted in Figure 5. One notices that the difference between the expected surplus in completed job spells and incomplete job spells increases with the time span. At the same time the strong concavity due to selection is clearly visible in both cases.

2.3.4 Expected Within-Job and Between-Job Wage Growth

We can apply the conditional expectations of Ω_t in incomplete and completed job spells for the analysis of the expected wage growth Δw_t within a job and respectively Δw_T^* between jobs. For this purpose, we decompose the random variables $[\Delta p_t, \Delta r_t]$

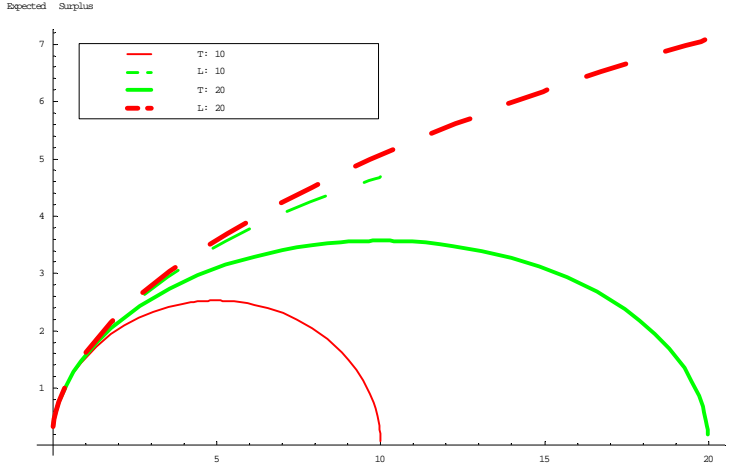


Figure 5: Expected Surplus in Long Spells

in two orthogonal components Δb_t and Δz_t , such that $\text{Cor}(\Delta b_t, \Delta z_t) = 0$. Given the previous assumptions on the joint normality of Δp_t and Δr_t , such a decomposition is always feasible. Hence, for $1 < t < T$,

$$\begin{aligned}\Delta r_t &= \Delta z_t - \gamma\beta\Delta b_t = \Delta z_t - \gamma\bar{\sigma}\Delta\Omega_t \\ \Delta w_t &= \Delta z_t + (1 - \gamma)\beta\Delta b_t = \Delta z_t + (1 - \gamma)\bar{\sigma}\Delta\Omega_t\end{aligned}\tag{13}$$

with $\Delta z_t \sim N(\mu_z, \sigma_z^2)$ and $\text{Cov}(\Delta z_t, \Delta b_t) = 0$. Obviously, the parameter γ can be expressed in terms of the covariance matrix Σ and the bargaining power β , but that is of little help here. It is more useful to interpret it as a reflection of the correlation between the match surplus and the reservation wage. In the one extreme case $\gamma = 0$, we can write $\Delta p_t = \Delta r_t + \Delta b_t$, with both right-hand side variables being uncorrelated. Then Δr_t reflects the evolution of the general human capital of the worker in this job as well as in all other jobs, which evolves independently of the value of the specific capital in the present job, Δb_t . Hence, the duration of the actual job is fully determined by its own (mis)fortune. Though the distinction between quits and layoffs makes little sense in this model, separations look like layoffs in this case:

the firm fires the worker since she is no longer productive. In the opposite extreme case $\gamma = 1$, we can write $\Delta r_t = \Delta p_t - \Delta b_t$, again with both right-hand side variables being uncorrelated. Now, Δp_t reflects the evolution of the general human capital of the worker in this job as well as in all other jobs; 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. We can use equation (13) to specify four OLS regressions, which are discussed below.

Taking expectations in the second equation of (13) yields:

$$\begin{aligned} \text{E}(\Delta w_t | 1 < t < T = \tau) &= \mu_z + (1 - \gamma)\bar{\sigma}\text{E}(\Delta\Omega_t | 1 < t < T = \tau) \quad (14) \\ \text{E}(\Delta w_t | 1 < t < L < T) &= \mu_z + (1 - \gamma)\bar{\sigma}\text{E}(\Delta\Omega_t | 1 < t < L < T) \\ \text{Var}(\Delta w_t | t < T) &\equiv \sigma_w^2 \end{aligned}$$

The first equation from (14) applies for completed spells, where we observe $T = \tau$; the second equation applies for incomplete spells, where we only know that the job ends beyond the period covered by the data, $T > L$. Since separation decisions are fully determined by the evolution of b_t (or, equivalently, Ω_t) and since Δb_t and Δz_t are uncorrelated, there is no selectivity in Δz_t . Hence, the conditioning $1 < t < L < T$ can be omitted in $\text{E}(\Delta z_t)$. As discussed before, Ω and π can be estimated from the distribution of observed job spells. These parameters are sufficient statistics for the calculation of the conditional expectations of $\Delta\Omega_t$. These expectations can be used as explanatory variables in a regression of within job log wage growth, Δw_t .³ These equations identify $(1 - \gamma)\bar{\sigma}$, but not γ and $\bar{\sigma}$ separately. Hence, we can only infer the part of the tenure profile that is associated with the selectivity in w_t , not in r_t .

³There is an alternative estimation strategy for within job log wage changes for completed spells. We observe r_t at the beginning and at the end of the job spell: $r_0 = w_0 - b^s - \bar{\sigma}\Omega$ and $r_T = w_0^* - b^s - \bar{\sigma}\Omega$, where w_0^* is the starting wage in the new job and Ω is kept constant across jobs. We can add these conditions to our regression analysis. Hence:

$$\text{E}(\Delta r_t | 1 < t < T, r_0, r_T) = \frac{r_T - r_0}{T} = \frac{w_0^* - w_0}{T}$$

For job changers, we can write a similar equation:

$$\mathbb{E}(\Delta w_t^* | t = T) = \mu_z + (1 - \gamma)\bar{\sigma}\mathbb{E}(\Delta\Omega_t | t = T) + \bar{\sigma}\Omega^* \quad (15)$$

The term $\mathbb{E}(\Delta\Omega_t | t = T)$ reflects the wage decline in the old job the year before separation.⁴ It is always negative, see Figure 2. The term Ω^* reflects the wage increase due to entering the new job, where the log productivity p_t is again substantially above the log reservation wage r_t . Equation (15) allows the separate identification of γ and $\bar{\sigma}$, by the final term, the upward jump in wages as a return to the specific investment in the new job, $\bar{\sigma}\Omega^*$. Hence, we are able to identify γ , and hence the selectivity in r_t , only by comparing the concavity of the tenure profile within job spells to the jump in wages when changing jobs. Otherwise, the data do not provide another way to estimate this part of the selectivity.

Finally, the change in starting wages from one job to another provides information on σ_z^2 :

$$\begin{aligned} \mathbb{E}(w_0^* - w_0) &= \mu_z T + \bar{\sigma}(\Omega^* - \Omega) \\ \text{Var}(w_0^* - w_0) &= \sigma_z^2 T = [\sigma_w^2 - (1 - \gamma)^2 \bar{\sigma}^2] T \end{aligned} \quad (16)$$

The first equation above follows immediately from the second equation of (13), since $\Omega_T = 0$. The first term reflects the general drift in log wages. Though till sofar the

Then, taking first differences and expectations in equation (9) yields:

$$\mathbb{E}(\Delta w_t | 1 < t < T, r_0, r_T) = \frac{w_T^* - w_0}{T} + \bar{\sigma}\mathbb{E}(\Delta\Omega_t | 1 < t < T)$$

We do not use this methodology here since it can only be applied to completed job spells, which, moreover, start within the observation period of the data (for otherwise, we do not observe w_0). This is a small subset of the total number of job spells.

⁴Implicitly, we assume here that separation takes place exactly at the end of the year of observation. Taking the model literally, this is an important assumption, since wages decline steeply in the last year before separation, see Figure 2. If separation occurs earlier on during the year of observation, part of the fall in wages during the last year before separation is captured by the previous observation. In the empirical section we also use the "full fall" in wages.

value of Ω was held constant across jobs, so that the second term drops out since the difference between the value of Ω for subsequent jobs vanishes, $\Omega^* - \Omega = 0$, we retained this term for future reference. The second equation in (16) follows from the second equation of (13), using the orthogonality of Δz_t and $(1 - \gamma)\beta\Delta b_t$ and $\text{Var}[(1 - \gamma)\beta\Delta b_t] = (1 - \gamma)^2 \bar{\sigma}^2$. The relation between the regression coefficients on $E(\Delta\Omega_t|\cdot)$ in equations (14) and (15) on the one hand, and the variances of Δw_t and $w_0^* - w_0$ on the other hand, provides a strong test for the model:

$$(1 - \gamma)\bar{\sigma} = \sqrt{\sigma_w^2 - \frac{1}{T}\text{Var}(w_0^* - w_0)} \quad (17)$$

This test relates the observed variance of wage changes within job spells, net of the variance of the overall shock z , to the degree of concavity in wages. If the model survives this test, it shows that the concavity in the tenure profile can be fully explained by selectivity.

3 Empirical Analysis

3.1 The Data

We use a dataset based on a PSID extract of 18 waves, covering the years 1975 through 1992, same as the one used by Altonji and Williams (1997, 1999). Our model does not work very well when employed people consider other alternatives than switching to another job, like retirement, leaving the labor force or taking up full time education. The availability of these other alternatives yields two problems. First, we do not observe the reservation wage at the point of separation when people do not accept another job. Second, with only one alternative to the present job, the decision problem is simply whether a particular indicator switches signs. With more alternatives, that choice process becomes far more complicated. Therefore we restrict the sample to people who do not switch in and out the labor force regularly and for

whom retirement is not a relevant option: white male heads of household with more than 12 years of education and less than 60 years of age. Our reasoning is similar to the one used in Mincer and Jovanovic (1981), who also use job separation synonymous to job change, thereby also defining labor mobility as change of employer and excluding other alternatives, which are minor phenomena in the case of the full-time male working force. Furthermore, we restrict the attention to those individuals that were employed, temporarily laid off, or unemployed at the time of the survey, and were not from Alaska or Hawaii. We use the tenure and experience measures constructed with the algorithm described by Altonji and Williams (1999 and previous working versions). Table 1 presents summary statistics of the variables of interest. There are missing values for all variables. Most missing values are reported for the log wage variable. However, we do not need this variable in the tenure distribution analysis and thus we include also the observations for which the log wage is not reported. One can distinguish four types of job spells. Apart from the distinction between completed and incomplete spells (right censoring), one can also make a distinction between spells that start before the time span covered by the data, and spells that start afterwards (left censoring). The table provides the number of spells of each of these four types. There is, however, a fundamental difference between these types of censoring. While right censoring implies that we do not know when a job spell has ended, left censoring does not imply that we do not know when a job has started because at the start of the observation period workers are asked for how long they hold their present job.

3.2 Test of the Random Walk Hypothesis

To prepare the ground for our formal analysis we document some stylized facts on wages. In particular, we verify that log wages follow a random walk, as assumed in our theoretical modelling. For this purpose we repeat the analysis of wage dynamics by MaCurdy (1982), Abowd and Card (1989) and Topel and Ward (1992). In the

Table 1: Summary Statistics Dataset

Variable	Mean	Std. Dev.	Min.	Max.	Observations
logwage ⁽¹⁾	2.44	0.49	0.06	4.82	18151
tenure (years)	7.45	7.88	0.08	43.69	20175
experience (years)	15.07	9.4	0.08	43.69	21099
year of observation	84.01	5.06	75	92	21099
job per individual	1.85	1.42	1	14	21099
age	34.56	9.68	18	60	21099
education (years)	13.69	1.77	12	17	20857
metropolis	0.61	0.49	0	1	21099
union member	0.21	0.41	0	1	20725
married	0.85	0.36	0	1	21099

Dataset for Estimating the Tenure Distribution Parameters	
Number observations discarded from AW (1997)	5431
Number of individuals	2837
Total number job spells	5484
- started before the observed range	1924
- started within the observed range	3560
Completed job spells	1911
- started before the observed range	434
- started within the observed range	1477
Incomplete job spells	3573

⁽¹⁾reported average hourly wage deflated using the implicit price deflator with 1982 base year

process of this analysis we document some further stylized facts that are useful in the subsequent analysis. First, we run a regression of within-job log wage differentials on a number of controls. Next, we construct a covariogram of residuals from this regression, from which we infer the process driving the wage dynamics. Finally, we show that the variance of the innovations in wages does not depend on experience and tenure.

Consider the following model with a very simple tenure profile:

$$w_{it} = \alpha + \beta_1 E_{ijt} + \beta_2 E_{ijt}^2 + \gamma_1 T_{ijt} + \gamma_2 T_{ijt}^2 + \eta_i + v_j + u_{it} \quad (18)$$

where $j(i, t)$ is the job j where worker i is employed at time t (we leave out the arguments of $j(\cdot)$ for convenience); η_i is a random individual effect (e.g. ability), v_j is a random job effect, and u_{it} is a time-varying stochastic component of wages. E stands for labor market experience and T for job tenure. For the sake of the argument,

let us assume that T_{ijt} is orthogonal to u_{it} . First-differencing equation (18) for job stayers yields:

$$\Delta w_{it} = (\beta_1 + \gamma_1) + \beta_2 \Delta E_{it}^2 + \gamma_2 \Delta T_{ijt}^2 + \Delta u_{it} \quad (19)$$

Since E_{it} and T_{ijt} increase at the same pace within jobs, we are not able to identify β_1 and γ_1 separately. This is one of the main problems in the identification of the tenure profile in wages: only workers who switch jobs allow us to distinguish β_1 and γ_1 . However, $\beta_1 + \gamma_1$ is estimated consistently in an OLS regression. We allow $\beta_1 + \gamma_1$ to vary by education level, for union members, for married people and for people living in a metropolitan area. Furthermore, we add a full set of year dummies to account for general variation in real wage growth and inflation, so that we do not obtain a single estimate for $\beta_1 + \gamma_1$. First-differencing for job movers yields:

$$\Delta w_{it}^* = \beta_1 - \gamma_1 (T_{ijt} - 1) + \beta_2 \Delta E_{it}^2 - \gamma_2 (T_{ijt}^2 - 1) + \Delta u_{it} \quad (20)$$

where the superfix * indicates that a job change has taken place and where T_{ijt} is the tenure in the old job. In this simple model, we can estimate the first order effect of the tenure profile by the effect of tenure on the change in wages for job movers. The regression results are displayed in Table 2.

Table 2: Within and Between-Jobs Wage Change Regressions

Variable	Within-Jobs		Between-Jobs	
	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
Intercept	0.055**	(0.016)	-0.058	(0.091)
$\Delta \text{Tenure}_{ijt}$			-0.001	(0.009)
$\Delta \text{Tenure}_{ijt}^2$	-8e-4*	(4e-4)	7e-4	(0.001)
$\Delta \text{Tenure}_{ijt}^3$	e-5 [†]	(8e-6)	-e-5	(2e-5)
$\Delta \text{Experience}_{it}^2$	-0.002**	(4e-4)	-0.003	(0.002)
$\Delta \text{Experience}_{it}^3$	2e-5**	(7e-6)	3e-5	(4e-5)
Education _{ijt}	9e-4	(9e-4)	0.008	(0.006)
Metropolis _{ijt}	0.007*	(0.003)	-0.011	(0.020)
Union member _{ijt}	0.001	(0.004)	-0.133**	(0.030)
Married _{ijt}	-0.006	(0.005)	-4e-4	(0.024)

Continued on next page...

... table 2 continued

Variable	Within-Jobs		Between-Jobs	
	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
Observations	11834		1448	
SER ⁽¹⁾	0.179		0.361	
R ²	0.0219		0.039	
Wald time ⁽²⁾	$\chi^2_{(16)}$: 144.96**		$\chi^2_{(13)}$: 20.02 [†]	
Wald joint ⁽³⁾	$\chi^2_{(24)}$: 264.72**		$\chi^2_{(22)}$: 57.86**	

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

The regressions include time effects

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

(2) Wald test for joint significance of time dummy coefficients

(3) Wald test for joint significance of all covariates

We run separate regressions for job stayers and changers. In the regression for stayers, we find evidence for concavity in the tenure and experience profiles, though the higher order terms of the tenure profile are smaller and less significant than those for the experience profile. Other variables do not matter, except for living in a metropolitan area, which positively influences wage growth. The results for job movers do not fit the simplest theory of a deterministic tenure profile that is unrelated to u_{it} , since the coefficients γ_1 and γ_2 are both insignificant. The only factor that matters is union membership. This result is consistent with the *right to manage* model of the union, see for example MaCurdy and Pencavel (1986), where unions negotiate wages above the going market rate, but where firms decide on employment. Moreover, the result is also consistent with some versions of the efficient bargaining model, where the union forces the firm to retain workers even when their wage is above marginal productivity. It is inconsistent with Nash bargaining at an individual level.

Table 3 presents the covariogram of the residuals of the within-job wage change regression. Residuals are strongly negatively correlated to their first lag, while autocorrelations for longer lags are small and statistically insignificant beyond lag 3. Lag 1 and 2 are strongly significant, lag 3 is only marginally significant at 5%. This outcome is very similar to results obtained by MaCurdy(1982), Abowd and Card (1989)

Table 3: Residual Autocovariances for Within-Job Wage Innovations

Lag	Autocovariance	Autocorrelation	Sig. Level	Obs.
0	.03213	1	0	11834
1	-.01056	-.35175	0	8992
2	-.00098	-.03342	.004	7385
3	.00072	.02484	.049	6293
4	.00025	.00866	.525	5381
5	-.00048	-.01781	.227	4590
6	-.00024	-.00913	.570	3875
7	-.00042	-.01657	.346	3238
8	.00015	.00591	.759	2691
9	-.00024	-.00917	.666	2209
10	.00066	.02358	.323	1755
11	-.00036	-.01402	.604	1370
12	.00007	.00323	.917	1034
13	.00061	.02944	.422	746
14	-.00061	-.02618	.553	516
15	-.00087	-.04970	.399	289
16	.00054	.04069	.681	104

and Topel and Ward (1992). Our covariogram is thus typical of an MA(2) process or even an MA(1) once we note that the second order lag autocovariance is close to 0.⁵ For simplicity, we focus on the MA(1) case.

We decompose the stochastic time-variant component of the wage equation from (18) in a martingale persistent component e_{it} and a transitory component η_{it} :

$$u_{it} = e_{it} + \eta_{it} \quad (21)$$

$$\Delta e_{it} = \varepsilon_{it}$$

where η_{it} and ε_{it} are i.i.d. with $\text{Var}(\eta_{it}) = \sigma_{\eta}^2$ and $\text{Var}(\varepsilon_{it}) = \sigma_w^2$, as defined before.

⁵We also considered the partial autocorrelation function (PACF) for the lagged residuals. The PACF values, lag 1 to 16, are the following, with starred values statistically significant at 10% or better: -0.35^{***} , -0.20^{***} , -0.07^{***} , -0.04^{***} , -0.02 , -0.015 , -0.04^{**} , -0.011 , -0.03 , 0.03 , 0.008 , 0.02 , 0.06 , -0.06 , -0.08 , 0.29^* . Hence, the PACF pattern further supports the pure MA specification versus a mixed ARMA type process.

Then:

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

This is a reasonable description of the pattern of autocorrelations in the covariogram in Table 2. A back-of-the-envelope calculation yields $\sigma_w^2 = 0.022$ and $\sigma_\eta^2 = 0.005$. Our result for the variance of persistent income shocks is similar to that reported by Meghir and Pistaferri (2004) for the PSID⁶. Hence, a random walk with transitory shocks provides a fairly accurate description of the dynamics of log wages. The standard deviation of yearly permanent innovations is substantial: 15% of the wage level per year. The transitory shocks η_{it} might reflect measurement error.

Table 4 present results for the Koenker (1981) "Studentized LM" version of the Breusch-Pagan (1979) test for homoskedasticity of u_{it} for both stayers and movers: the squared residuals are regressed on a constant term and on all control variables.

Table 4: BP Heteroskedasticity Test for Wage Changes Within and Between-Jobs

Variable	Within-Jobs		Between-Jobs	
	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
Intercept	-0.018	(0.015)	0.114	(0.090)
$\Delta \text{Tenure}_{ijt}$			-0.003	(0.009)
$\Delta \text{Tenure}_{ijt}^2$	-2e-4	(3e-4)	6e-4	(0.001)
$\Delta \text{Tenure}_{ijt}^3$	-3e-6	(7e-6)	-2e-5	(3e-5)
$\Delta \text{Experience}_{it}^2$	e-4	(4e-4)	0.001	(0.002)
$\Delta \text{Experience}_{it}^3$	2e-6	(6e-6)	-2e-5	(4e-5)
Education $_{ijt}$	0.003**	(9e-4)	-7e-4	(0.006)
Metropolis $_{ijt}$	0.004	(0.003)	0.018	(0.020)
Union member $_{ijt}$	-0.017**	(0.004)	0.033	(0.030)

Continued on next page...

⁶For their whole PSID sample of males heads of household, Meghir and Pistaferri (2004) obtain 0.031 as variance of the permanent income shocks. For highschool graduates without college degree (the largest part of individuals in our PSID extract) they report an even closer result to ours, 0.027.

... table 4 continued

Variable	Within-Jobs		Between-Jobs	
	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
Married _{ijt}	0.004	(0.005)	-0.027	(0.024)
Observations	11834		1448	
SER	0.166		0.357	
R ²	0.007		0.014	
Wald time	$\chi^2_{(16)}$: 35.20**		$\chi^2_{(13)}$: 16.90	
Wald joint	$\chi^2_{(24)}$: 89.28**		$\chi^2_{(22)}$: 21.34	
Breusch-Pagan ⁽¹⁾	$\chi^2_{(24)}$: N*R ² =87.57**		$\chi^2_{(22)}$: N*R ² =21.43	

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

The regressions include time effects

The dependent variable is the square residual from Table 2.

We use Koenker's (1981) modified version of the Breusch-Pagan test

The null hypothesis of homoskedasticity is rejected for stayers, but only because of education and union membership. The evolution of wages in unionized jobs is largely governed by predetermined wage scales, explaining the low variance of wage changes. The higher variance for better educated workers squares with the conclusions from studies analyzing risk in educational choice and is confirmed also by other empirical applications on the PSID. In the regression for job movers homoskedasticity is not rejected. However, the variance in wages for movers is much higher than for stayers, so there is clearly heteroskedasticity between movers and stayers. At first sight this result is hard to square with our theoretical framework, where wages follow a random walk and where there is a perfect frictionless market for alternative job opportunities. However, essential for our purpose is that there is no heteroskedasticity with respect to either job tenure or experience. For example, a simple learning model would imply a higher variance early on in the career, when people still have to learn their capabilities and comparative advantages before finding their optimal profession, see Jovanovic (1979) and Topel and Ward (1992). The results reported in Table 4 do not confirm this idea. Instead they provide support for the ideas put forward in Section 2, where w_t is assumed to follow a random walk. When we assume that workers are able to disentangle permanent and transitory shocks (which is certainly true if the

transitory component η_{it} reflects measurement error or variation in hours worked), then the transitory shock will not have an effect on job changes because changing jobs permanently is not a useful response to a shock that has only a transitory effect: as soon as you have changed jobs, the cause for changing has faded away. Hence, we feel safe to ignore the effect of transitory shocks of job relocation in the subsequent analysis.

Our preparatory ground work provides support for the main ingredient of our model, wages following a random walk, where neither the variance of the innovations nor the wage loss of moving to another jobs depend on tenure and experience. There are two aspects which do not fit our theoretical framework well: (i) the larger wage loss for union members upon job change, which seems to be *prima facie* evidence against Nash bargaining for unionized firms, and (ii) the larger variance of wages upon job change, which suggests that there are substantial search frictions on the market for alternative job opportunities.

3.3 The Parameters of the Tenure Distribution

The parameters of the tenure distribution for the surplus Ω and the drift π can be estimated by maximum likelihood, using the density function (8). Till sofar, we have treated both parameters as constants which do not depend on worker characteristics. However, one can expect that workers choose their optimal job type according to their characteristics. Hence, Ω and π are likely to differ according to both observed and unobserved worker characteristics. Since we deal with longitudinal data we can take into account random worker effects. We do not consider random job effects for both theoretical and empirical reasons. From a theoretical point of view, our assumption of a frictionless market for alternative job opportunities, where the only constraint on instantaneous mobility is the specific investment in the present job and not the cost of getting another job offer, each worker type will choose that job type that fits best her comparative advantages, like in Sherwin Rosen's famous hedonic world of kissing

curves. Hence, job characteristics are implied by worker characteristics. The only job characteristic that we allow for is union membership. From an empirical point of view, we observe each job only once, so that we have no basis for identifying random job effects other than from functional form assumptions. Taking into account that Ω has to be positive, the following specification for Ω and π is adequate:

$$\begin{aligned}\Omega_i &= \exp(x'_{ij}\beta_\Omega + u_{\Omega i}) \\ \pi_i &= x'_{ij}\beta_\pi + u_{\pi i}\end{aligned}\tag{22}$$

where $u_{\Omega i}$ and $u_{\pi i}$ are normally distributed random worker effects with mean 0 and standard deviations σ_Ω and σ_π , and where x_{ij} is a vector of observed worker characteristics, i.e. education, experience at the start of the job spell, a dummy for union membership, a dummy for living in a metropolitan city and a dummy for being married⁷. We take all covariates (including the dummy variables) in deviations from their means over jobs. Hence, the intercept can be interpreted as the mean value for Ω and π respectively. We assume both random effects to be uncorrelated. Then, the log likelihood function reads:

$$\log L = \sum_i^N \ln \int \int \prod_{j=1}^{J_i} [1 - F(T_{ij})]^{1-d_{ij}} \cdot f(T_{ij})^{d_{ij}} d\Phi\left(\frac{u_{\Omega i}}{\sigma_\Omega}\right) d\Phi\left(\frac{u_{\pi i}}{\sigma_\pi}\right)\tag{23}$$

where $j(i)$ is the j^{th} job held by worker i (we leave out the argument i of $j(i)$ for the sake of convenience), where d_{ij} is a dummy variable, taking the value $d_{ij} = 1$ if the job spell is completed and the value $d_{ij} = 0$ otherwise, where T_{ij} is the completed tenure if $d_{ij} = 1$ and respectively the tenure at time L , the last moment of observation in the panel, otherwise, and where N is the number of individuals and J_i is the number

⁷Since we do not observe prior mobility for most workers in our PSID sample we cannot include it as covariate, as done for instance by Mincer and Jovanovic (1981) or Farber (1994). Although prior mobility is often found to increase the probability of separation, whether this is due to unobserved characteristics of the workers or it has a direct effect on current tenure is unclear; including random worker effects we should be able to control for most unobserved worker characteristics.

of jobs for individual i .

There are two reasons why we have to make amendments to the simple likelihood function in equation (23). First, we could restrict the estimation to job spells that have started within the observation range of our PSID extract. However this means that we do not consider jobs starting before the beginning of the observation period. 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 spells starting before the first observation period of the PSID. We know all x_{ij} 's for these spells since there are neither dummy indicator changes, nor education changes during the course of a job and since 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:

$$\log L = \sum_i^N \ln \int \int \prod_{j=1}^{J_i} \frac{[1 - F(T_{ij})]^{1-d_{ij}} \cdot f(T_{ij})^{d_{ij}}}{1 - F(t_{ij})} d\Phi\left(\frac{u_{\Omega i}}{\sigma_{\Omega}}\right) d\Phi\left(\frac{u_{\pi i}}{\sigma_{\pi}}\right) \quad (24)$$

where t_{ij} is the tenure of individual i in job j at the start of the PSID. Note that for spells started after the start of the PSID, $t_{ij} = 0$, so $F(t_{ij}) = 0$, meaning that we are back in the simple case from (23).

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 question⁸, but we do not know whether there has been another job spell between the job observed a year ago and the job observed now. Since the hazard rate implied by our model is hump shaped, with the hump likely to be within the first year, cf. Farber (1994), this

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

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 (24), where t_{ij} is the initial tenure in months as measured at the first observation after the start of the spell. However, this approach does not use the distribution of these t_{ij} 's itself⁹ 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:

$$q(t) = \frac{1 - F(t)}{\int_0^1 [1 - F(x)] dx}$$

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

$$\frac{f(T)}{1 - F(t)} q(t) = \frac{f(T)}{\int_0^1 [1 - F(x)] dx}$$

Hence, the log likelihood reads:

$$\log L = \sum_i^N \ln \int \int \prod_{j=1}^{J_i} \frac{[1 - F(T_{ij})]^{1-d_{ij}} \cdot f(T_{ij})^{d_{ij}}}{\int_0^1 [1 - F(x)] dx} d\Phi\left(\frac{u_{\Omega i}}{\sigma_{\Omega}}\right) d\Phi\left(\frac{u_{\pi i}}{\sigma_{\pi}}\right) \quad (25)$$

The log likelihood that accounts both for jobs starting before the first wave of the PSID and for the left censoring for spells shorter than a year started after the first

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

wave of the PSID, can thus be written as:

$$\log L = \sum_i^N \ln \int \int \prod_{j=1}^{J_i} \frac{[1 - F(T_{ij})]^{1-d_{ij}} \cdot f(T_{ij})^{d_{ij}}}{D_{ij}} d\Phi\left(\frac{u_{\Omega i}}{\sigma_{\Omega}}\right) d\Phi\left(\frac{u_{\pi i}}{\sigma_{\pi}}\right) \quad (26)$$

where $D_{ij} = \begin{cases} \text{if spell starts after start PSID: } \int_0^1 [1 - F(x)] dx \\ \text{if spell starts before start PSID: } 1 - F(t_{ij}) \end{cases}$

We report results for (25), where we use only the jobs that start within the observation period of the PSID, and for (26), where we use the sample including job spells starting before the first wave of the PSID¹⁰. The estimation results are presented in Table 5.

Theoretically, the results for both likelihood functions should be identical. Two observations are in place. First, though the two sets of estimation results are statistically different, the theoretical hazards for both models look very similar (cf. Figure 1 above), the only difference being the height of the peak (lower for the case where we use all job spells). The estimated intercepts of Ω and respectively π are more or less identical (remember that all other x_{ij} are taken in deviation from the mean, so that the intercept is something like the mean value for Ω and π), but the coefficients for other variables differ substantially, in particular those that are less precisely measured in the small sample. Second, a number of results in Table 5 are consistent across both estimations. Spells started at higher age and spells at unionized firms have a larger drift π_{ij} . When we include the spells started before the first wave of the PSID, these variables also have a positive effect on the surplus Ω_{ij} . The positive effect of experience would be consistent with the idea that workers start their career with some initial job hopping, before settling down in a job that fits one's compara-

¹⁰In 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.

tive advantages best. For the positive effect of union membership, economic theory provides two explanations, one in which union membership causes a high Ω_{ij} and one in which the causality runs the other way around. The first argument relies on a hold up problem. Unions extract part of the firm’s compensation for the specific investment at job start. Firms respond by postponing job openings till the surplus Ω_{ij} is so large that even their smaller share in it provides sufficient compensation for their investment. This type of argument is supported by empirical findings in Bertrand and Mullainathan (2003) where if firms are insulated from takeovers, wages rise, while the rates of destruction of old plants and creation of new plants fall. The second argument relies on the idea that the larger Ω_{ij} , the larger the workers’ incentives to get a proper share in it, and hence the greater the payoff of setting up a union. Our estimation results do not allow distinguishing between these two scenarios. The second result that is consistent between both set of estimation results is that the intercept for π_{ij} is positive and large. In both cases, there are hardly observations for which π_{ij} 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%, as calculated in Section 2. This observation will play an important role in the rest of the analysis.

Table 5: MLE Tenure Distribution Parameters

Variable	Small Sample ⁽¹⁾		Large Sample ⁽²⁾	
	Drift π	Dist Ω	Drift π	Dist Ω
Intercept	0.228**	-1.208 **	0.158**	-1.132**
(st. errors)	(0.021)	(0.155)	(0.0015)	(0.014)
Education	0.017	0.040	-0.0005	-0.032**
(st. errors)	(0.012)	(0.039)	(0.0008)	(0.007)
Initial experience	0.008**	-0.009	0.011**	0.007**
(st. errors)	(0.002)	(0.009)	(0.0002)	(0.001)
Union member	0.233**	0.271	0.138**	0.795**
(st. errors)	(0.056)	(0.175)	(0.003)	(0.025)
Metropolis	0.015	0.070	0.009**	-0.133**
(st. errors)	(0.043)	(0.140)	(0.002)	(0.023)
Married	0.091 [†]	0.121	-0.033**	-0.012
(st. errors)	(0.049)	(0.162)	(0.003)	(0.030)

Continued on next page...

... table 5 continued

Variable	Small Sample		Large Sample	
	Drift	Dist	Drift	Dist
Random worker effects σ (st. errors)	0.277** (0.050)	0.0001 (5.51)	1.34e-8 (0.001)	1.75e-7 (0.010)
Number Observations (job spells)	1911		5484	

⁽¹⁾Small sample= sample of job spells starting within the range of the PSID sample

⁽²⁾Large sample= sample of all job spells

All covariates are taken in deviations from their means over jobs

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

One remarkable conclusion is that there are no unobserved worker effects when we use the sample including the spells started before the first wave of the PSID, while there is unobserved heterogeneity in the drift for the sample without these spells. Since the long spells started before the first wave contain crucial information, we focus on the estimation results obtained from the full sample of job spells in what follows.

For future reference and as a test of the goodness of fit of the model, we compute the density of incomplete job spells after a fixed working experience, in this case $L = 32$ years.¹¹ Figure 6 depicts both the predicted and the empirical density of incomplete job spells. There is a reasonable correspondence between both densities. The peak in the first year is overestimated, but otherwise the shapes of the two densities are identical. Note the small peak in the density for short incomplete spells, which is due to the hump shape pattern in the hazard: if your job ends for instance in the last five years before the end of the observation period, there is a substantial

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

Predicted and Empirical Densities
Conditional on Experience = 32 years

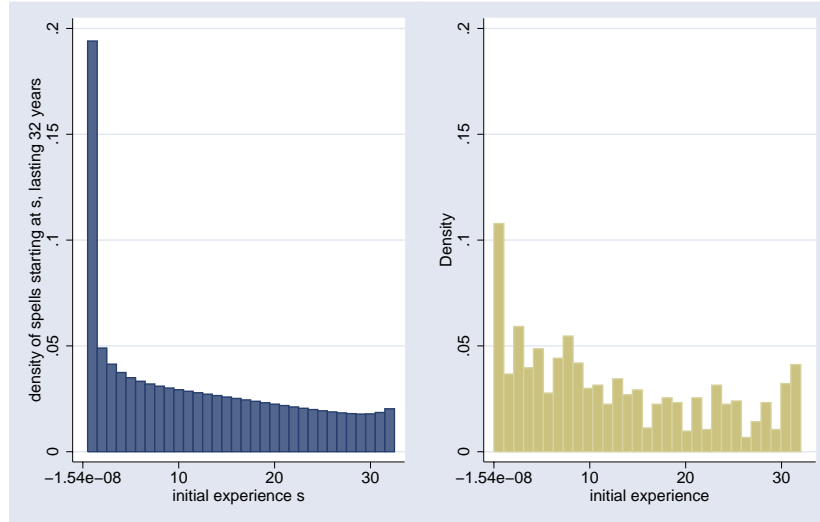


Figure 6: Density of Incomplete Job Spells with Exit Option

probability that you experience further separations afterwards due to the peak in the hazard rate, leading to a peak of short incomplete tenures. Close alignment of the predicted and the empirical densities suggests that our model works well.

3.4 Wage Dynamics

Table 6 presents estimation results on wage changes, separately for completed job spells, incomplete job spells and at job transitions, with heteroskedasticity-robust standard errors¹².

Table 6: Wage Changes in Completed Spells, Incomplete Spells and at Job Transitions

1: Completed	2: Incomplete	3: Job Switch
A: Unrestricted Regression Estimates		

Continued on next page...

¹²We tested for the absence of individual specific effects. We left out marital status as a regressor since it was insignificant in all regressions.

... table 6 continued

Intercept	0.060**	0.054**	0.216*
(st. errors)	(0.015)	(0.010)	(0.102)
$\Delta E(\Omega_{ijt})$	0.008	0.039**	0.137*
(st. errors)	(0.006)	(0.013)	(0.062)
$\Omega_{i,j+1,0}$			0.308*
(st. errors)			(0.145)
Exp_{ijt}	-0.004*	-0.005**	-0.007
(st. errors)	(0.002)	(9e-4)	(0.007)
Exp_{ijt}^2	8e-5	9e-5**	6e-5
(st. errors)	(5e-5)	(2e-5)	(e-4)
Educ_{ijt}	0.002	e-4	0.031**
(st. errors)	(0.002)	(0.001)	(0.009)
Union_{ijt}	0.006	-0.004	-0.163**
(st. errors)	(0.010)	(0.003)	(0.062)
Metropolis_{ijt}	0.007	0.011**	-0.017
(st. errors)	(0.008)	(0.003)	(0.031)
Observations	2022	9653	618
SER	0.177	0.180	0.362
R ²	0.007	0.011	0.061
Wald joint	$\chi_{(6)}^2: 15.84^*$	$\chi_{(6)}^2: 110.76^{**}$	$\chi_{(7)}^2: 43.05^{**}$

B: Restricted Regression Estimates

Intercept	0.049**	0.147 [†]
(st. errors)	(0.017)	(0.079)
$+\Delta E(\Omega_{ijt})^{(a)}$	0.051*	0.075** ^(b)
(st. errors)	(0.025)	(0.026)
$-\Delta E(\Omega_{ijt})^{(a)}$	-6e-4	
(st. errors)	(0.007)	
$\Omega_{i,j+1,0}$		0.288*
(st. errors)		(0.142)
Exp_{ijt}	-0.003 [†]	-0.005
(st. errors)	(0.002)	(0.007)
Exp_{ijt}^2	6e-5	7e-5
(st. errors)	(5e-5)	(e-4)
Educ_{ijt}	0.002	0.030**
(st. errors)	(0.002)	(0.009)
Union_{ijt}	0.007	-0.161**
(st. errors)	(0.010)	(0.061)
Metropolis_{ijt}	0.007	-0.016
(st. errors)	(0.008)	(0.031)
SER	0.177	0.360
R ²	0.009	0.067
Wald joint	$\chi_{(7)}^2: 20.09^{**}$	$\chi_{(7)}^2: 45.78^{**}$

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

^(a) The +/- split of $\Delta E(\Omega_{ijt})$ is used for completed spells

^(b) For the job to job transitions we use the "full fall" in $E(\Omega_{ijt})$

Panel (A) of the table above presents unrestricted regression results for job stayers, cf. equation (14), for completed spells in column 1 and for incomplete spells in column 2, and for job movers, cf. equation (15), in column 3. The theoretically relevant regressors, $\Delta E(\Omega_{ijt}|T, L)$ and $\Omega_{i,j+1}$, have the right sign. The other variables tend to have limited effect. However, the coefficient on $\Delta E(\Omega_{ijt}|T, L)$ should be the same in all three regressions, namely $(1 - \gamma)\bar{\sigma}$. Clearly, this prediction is not confirmed by the data. Panel (B) therefore considers a slightly amended version of the model. Consider Figures 2 to 5 on the trajectories of $E(\Omega_{ijt}|T)$ for completed spells. These trajectories are falling in the period just before separation. Suppose that there is downward rigidity that prevents wages from actually falling. This is consistent, for example, with the study by Beaudry and DiNardo (1991), who show that within a job spell wages go up when unemployment falls, but do not go down when unemployment rises. We could test this idea in a crude way by separating the initial part of the job spell for which the surplus is increasing, $\Delta E(\Omega_{ijt}|T) > 0$, and the final part for which the surplus is decreasing, $\Delta E(\Omega_{ijt}|T) < 0$, and enter both as separate regressors. The estimation results for this model are presented in column 1. The results strongly confirm our hypothesis: the upward part comes in with a coefficient which is very similar in size with the coefficient for $\Delta E(\Omega_{ijt}|T > L)$, from Panel (A), column 2. Furthermore, if this downward rigidity model really applies, one would expect that "missing wage declines" in the years before separation are actually compensated at the moment of separation by an additional fall. Hence, we include in the regression for job changers not the change in the surplus during the year of separation, $\Delta E(\Omega_{ij,T-1})$, but the full decline, starting from the maximum of $E(\Omega_{ijt}|T)$ during the job spell, till its minimum value at the moment of separation, $E(\Omega_{ijT}) = 0$. The results for this model are presented in column 3 of Panel B. The relevant coefficient comes down from 0.137 to 0.075, a number that is close to the coefficients found for the wage changes in completed and incomplete spells. Both results provide support for our amended model.

Table 7 combines the regressions for completed spells, incomplete spells and job changers for this amended model, that is, column 2 of Panel A and columns 1 and 3 of Panel B from Table 6 above.

Table 7: Wage Change Regressions for All Job Spells

Intercept	0.053**
(st. errors)	(0.008)
$+\Delta E(\Omega_{ijt})$	0.041**
(st. errors)	(0.009)
$-\Delta E(\Omega_{ijt})$	-0.006
(st. errors)	(0.006)
$\Omega_{i,j+1,0}$	0.332**
(st. errors)	(0.065)
Exp_{ijt}	-0.005**
(st. errors)	(7e-4)
Exp_{ijt}^2	8e-5**
(st. errors)	(2e-5)
$\text{Educ}_{ijt}^{(a)}$	0.031**
(st. errors)	(0.008)
$\text{Union}_{ijt}^{(a)}$	-0.174**
(st. errors)	(0.053)
Metropolis_{ijt}	0.010**
(st. errors)	(0.003)
Observations	12293
SER	0.99981
R ²	0.0286
Wald joint	$\chi_{(8)}^2: 166.48^{**}$

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

^(a) Union and education are interacted with job change

We make one further amendment to the model, by interacting union membership and education level with job change, so that both variables affect only the wage change for job movers. We weight the three samples by their respective standard errors of the regression (SER) obtained in each of the three separate restricted regression models. The $F(14, 12270)$ statistic of the restriction that the coefficients of these 3 regressions are equal is 0.667, so that restrictions are accepted. Hence, the amended model provides a good description of the data. The only exception are union members, who face a 17% additional wage loss upon separation. Furthermore, the positive

effect of education on the wage change of movers is not predicted by the model (note that education enters the regression via its effect on Ω_{ij} , see Table 5). The estimation results yield $(1 - \gamma)\bar{\sigma} = 0.041$ and $\bar{\sigma} = 0.332$, implying $\gamma = 0.877$. This is a remarkable result. Apparently, separation is not so much driven by shocks to the job's productivity p_t , as it is by shocks to the worker's reservation wage r_t , which seems somewhat counter-intuitive. Similarly, we can calculate the return to tenure, $\bar{\sigma}\pi = 0.332 * 0.15 = 5\%$ (taking the estimated mean value of π). However, the high value of γ implies that most of the return to tenure, more than 85%, takes the form of the log outside wage r_t falling, instead of the inside wage w_t rising, cf. equation (13). Without this selectivity part the tenure profile is really small, $(1 - \gamma)\bar{\sigma}\pi = 0.041 * 0.15 = 0.6\%$. Both estimates, that of γ and of the tenure profile are fully driven by the estimate of $\bar{\sigma}$, which is driven by the effect of Ω_{ijt} on the wage change for job movers, see equation (15). This is a thin line of identification. Although the part of the tenure profile due to selectivity is weakly identified, this is the first research to actually account for selectivity in the observed outside wages.

Equation (17) runs the test of comparing the concavity in the tenure profile to the variance in wage changes. Table 8 presents estimation results for equation (16).

Table 8: Initial Wage Changes Regressions at Job Transition

1: Mean^(a)		2: Variance^{(a),(b)}	
Intercept		Intercept	0.092**
(st. errors)		(st. errors)	(0.024)
Ten _{i,j-1,T} ^(c)	0.057**	Ten _{i,j-1,T-1}	0.019**
(st. errors)	(0.011)	(st. errors)	(0.007)
$\Delta\Omega_{ij0}$	0.075		
(st. errors)	(0.251)		
ΔExp_{ij0}^2	-0.001*		
(st. errors)	(4e-4)		
Observations	336	336	
SER	0.379	0.297	
R ²	0.14	0.017	
Wald joint	$\chi_{(3)}^2: 41.16^{**}$	$\chi_{(1)}^2: 7.44^{**}$	
Significance levels : † : 10% * : 5% ** : 1%			
^(a) We use completed jobs that last more than 1 year (Ten _{i,j-1,T} > 1)			

Continued on next page...

... table 8 continued

(b) The dependent variable is the squared residual from the left column estimation

(c) $Ten_{i,j-1,T} = Exp_{ij0} - Exp_{i,j-1,0}$

We restrict the sample to jobs lasting more than 1 year, since wage changes for jobs that last only one year are noisy anyway¹³. Column 1 reports the coefficients of the regression. Column 2 takes the squared residuals from the first regression, and regresses them on elapsed tenure, T . The intercept captures the excess variance for job movers, see Table 4, and the transitory shocks in wages, η_{ijt} , see equation (21). The coefficient for T is a consistent estimator of $\frac{1}{T}\text{Var}(w_0^* - w_0)$. An estimator for σ_w^2 is derived from the regression for within-job wage changes in Table 4, see Section 3.2. Hence:

$$(1 - \gamma)\bar{\sigma} = \sqrt{0.022 - 0.019} = 0.054$$

which is very similar to the estimated value in Table 7, 0.041. Thus the amended model provides a good description of the data. The concavity of the tenure profile is fully captured by the term $\Delta E(\Omega_{ijt}|T, L)$ and the size of its regression coefficient is consistent with the variance of yearly wage innovations and the variance of initial wages over job spells.

A final question we ask is to what extent the option to switch jobs limits the growth of the variance in log wages over time. Without the option to switch jobs, the variance of log wages would increase linearly over time, due to the fact that z_t and b_t follow a random walk. However, the option to switch jobs allows the worker to eliminate bad trajectories of b_t , thereby compressing its variance. This can be seen from the distribution of incomplete tenures, see Figure 6, showing that a substantial fraction of the jobs has an incomplete tenure of less than 32 years. There are two mechanisms that lead to compression. First, many jobs have an incomplete tenure of less than 32 years and hence a smaller variance, since the variance increases proportional to

¹³Using the whole sample, the coefficient of T in the variance of wages regression is 0.010.

incomplete tenure. Second, those jobs that are still going on after some period are a selective sample of all the trajectories that have started initially, namely those which never crossed the separation threshold. This selection process compresses the variance. We use the density of incomplete tenures in Figure 6, and the density of $\Omega_t = b_t/\sigma$ conditional on the incomplete tenure T , $g(\Omega_t, t, \Omega_0) / [1 - F(t, \Omega_0)]$, see equations (6) and (7). In Figure 7, we plot the evolution of the variance of b_t without the option to switch jobs, the line $\sigma^2 T$, and the evolution of the variance with that option. The plots reveal that the option to switch jobs compresses the variance of b_t considerably: by about 65% after 32 years of experience¹⁴. Note however that the variance of z_t remains unaffected by this process, while this accounts for the main share of the total variance of w_t .

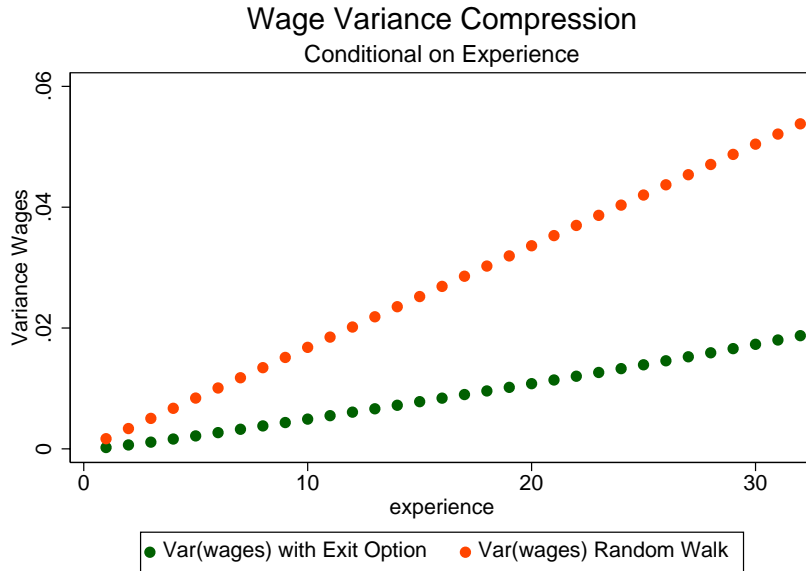


Figure 7: Compression Effect of the Exit Option on the Wage Variance

¹⁴We use $\sigma = 0.041$, our estimated coefficient from Table 7. Hence, after 32 years of experience, without the exit option the variance of b_t would be $32 * 0.41^2 = 0.054$. With the option to exit the wage variance is compressed after the same period to about 0.019.

4 Concluding Remarks

Our simple model for the stochastic evolution of productivity explains the data on the job tenure distribution and wages for the USA surprisingly well. This model features frictionless labor market at the moment of job start (which enables workers to pick the best job alternative straightaway), specific investment and hence subsequent lock-in on the current job, and efficient bargaining over the match surplus. 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 tenure 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 observed 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 . More than 5/6 of our estimated tenure profile is accounted for by this selectivity in the outside option. However, identification of the part of the variance due to variation in r_t is fragile, since we observe r_t only at the moment of job switching. We find excess variance of wages at job transition, indicating that our assumption of frictionless market for alternatives is incorrect. Apparently, there is a great deal more randomness in the wage in a new job than is to be expected from a model with frictionless market.

Our model implies that, on average, wages should be falling relative to the wages in other jobs in the second half of a completed spell. The data do not support this prediction. Apparently, there is some downward rigidity in wages. This is in particular true for unionized jobs, suggesting that there is inefficient separation in the unionized sector. Efficient bargaining over the match surplus is clearly rejected for unionized jobs. A slightly adjusted model, allowing for downward rigidity in wages, can explain the data very well:

$$\begin{aligned}\Psi_t &= \max(\Omega_t, \Psi_{t-1}) \\ w_t &= r_t + \bar{\sigma}\Psi_t\end{aligned}$$

where job separation is still governed by the rule that a job ends at the first moment in time that Ω_t reaches 0. Ψ_t is the maximum of its value in the previous period and the current value of Ω_t . Then, we would observe more wage rigidity in job spells that are expected to end shortly (since there Ω_t is declining), which explains that $\Delta\Omega_t$ does not affect Δw_t for $\Delta\Omega_t < 0$. Hence, the standard error of Δw_t should be smaller in the period just before separation. This model needs to be analyzed more formally, but the analysis presented here suggests that it provides an adequate description of the data.

A Computational Appendix

A.1 Completed job spells

The first factor in the numerator of (10) is given by the equation of the density function in (8) substituting t by $\tau - t$ and Ω by ω :

$$f(\tau - t, \omega) = \frac{\omega}{(\tau - t)\sqrt{\tau - t}} \phi\left(\frac{\omega + (\tau - t)\pi}{\sqrt{\tau - t}}\right) \quad (27)$$

The second part of the numerator in (10) was given in (6).

Substitution of (27) and (6) in (10), and some simplification yields:

$$\begin{aligned}
& h(\omega, t, \tau) \tag{28} \\
= & \frac{\omega\tau}{\Omega(\tau-t)} \sqrt{\frac{\tau}{(\tau-t)t}} \\
& \times \left[\phi \left(\sqrt{\frac{\tau}{(\tau-t)t}} \left(\omega - \frac{\tau-t}{\tau} \Omega \right) \right) - \phi \left(\sqrt{\frac{\tau}{(\tau-t)t}} \left(\omega + \frac{\tau-t}{\tau} \Omega \right) \right) \right]
\end{aligned}$$

Interestingly, this probability density does not depend at all on the drift π . The drift affects therefore the distribution of completed job tenures (see expression 7 above), but not the distribution of Ω_t conditional on the completed tenure.

The expectation in (11) can now be calculated:

$$\begin{aligned}
\mathbf{E}(\Omega_t | 0 < t < T) &= \int_0^\infty \omega h(\omega, t, T) d\omega \tag{29} \\
&= 2\sqrt{\frac{(T-t)t}{T}} \phi \left(\sqrt{\frac{T-t}{Tt}} \Omega \right) \\
&\quad + \left(\frac{t}{\Omega} + \frac{T-t}{T} \Omega \right) \left[2\Phi \left(\sqrt{\frac{T-t}{Tt}} \Omega \right) - 1 \right]
\end{aligned}$$

If we denote $f(t) \equiv \frac{T-t}{T}$ in (29) above we obtain exactly (11).

We compute the first and second derivatives of (29) below:

$$\begin{aligned}
\frac{d\mathbf{E}(\Omega_t | 0 < t < T)}{dt} &= -2 \left(\frac{\sqrt{t}}{\sqrt{T}\sqrt{T-t}} \right) \phi \left(\frac{\sqrt{T-t}\Omega}{\sqrt{t}\sqrt{T}} \right) \tag{30} \\
&\quad + \left(\frac{1}{\Omega} - \frac{\Omega}{T} \right) \left(2\Phi \left(\frac{\sqrt{T-t}\Omega}{\sqrt{t}\sqrt{T}} \right) - 1 \right)
\end{aligned}$$

$$\frac{d^2\mathbf{E}(\Omega_t | 0 < t < T)}{dt^2} = - \left(\sqrt{\frac{T}{(T-t)t}} \right)^3 \phi \left(\sqrt{\frac{T-t}{Tt}} \Omega \right) \tag{31}$$

A.2 Incomplete job spells

The first factor in the numerator of (12) is given by substituting t with L and Ω with ω in (7):

$$\begin{aligned} & 1 - F(L - t, \omega) \\ &= \Phi\left(\frac{\omega + \pi(L - t)}{\sqrt{L - t}}\right) - e^{-2\omega\pi}\Phi\left(\frac{-\omega + \pi(L - t)}{\sqrt{L - t}}\right) \end{aligned} \quad (32)$$

$g(\omega, t, \Omega)$ was given in (6).

The complicated part in (12) resides in the integral from the denominator of (12). The task would involve computing integrals of the form:

$$\int_0^\infty \phi(ax + b)\Phi(cx + d) dx \quad (33)$$

with a, b, c, d given. Since this form does not have an analytic solution, we compute numerically $E(\Omega_t | t < L < T)$:

$$E(\Omega_t | t < L < T) = \int_0^\infty \omega h^*(\omega, t, L) d\omega = \frac{A}{B} \quad (34)$$

where A and B are computed by numerical integration, as follows:

$$\begin{aligned} A &\equiv \int_0^\infty \frac{\omega}{\sqrt{t}} \left[\phi\left(\frac{\omega - \Omega - \pi t}{\sqrt{t}}\right) - e^{-2\Omega\pi} \phi\left(\frac{\omega + \Omega - \pi t}{\sqrt{t}}\right) \right] \\ &\quad \times \left[\Phi\left(\frac{\omega + \pi(L - t)}{\sqrt{L - t}}\right) - e^{-2\omega\pi} \Phi\left(\frac{-\omega + \pi(L - t)}{\sqrt{L - t}}\right) \right] d\omega \end{aligned} \quad (35)$$

and

$$\begin{aligned} B &\equiv \int_0^\infty \frac{1}{\sqrt{t}} \left[\phi\left(\frac{\omega - \Omega - \pi t}{\sqrt{t}}\right) - e^{-2\Omega\pi} \phi\left(\frac{\omega + \Omega - \pi t}{\sqrt{t}}\right) \right] \\ &\quad \times \left[\Phi\left(\frac{\omega + \pi(L - t)}{\sqrt{L - t}}\right) - e^{-2\omega\pi} \Phi\left(\frac{-\omega + \pi(L - t)}{\sqrt{L - t}}\right) \right] d\omega \end{aligned} \quad (36)$$

References

- Abowd, J., F. Kramarz & D. Margolis (1999), 'High wage workers and high wage firms', *Econometrica* **67**(2), 251–333.
- Abowd, J.M. & D. Card (1989), 'On the covariance structure of earnings and hours changes', *Econometrica* **57**(2), 411–45.
- Abraham, K.G. & H.S. Farber (1987), 'Job duration, seniority, and earnings', *The American Economic Review* **7**(3), 278–97.
- Altonji, J. G. & N. Williams (1997), Do wages rise with job seniority? a reassessment. NBER Working Paper No. 6010.
- Altonji, J. G. & 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. & N. Williams (2005), 'Do wages rise with job seniority? a reassessment', *Industrial and Labor Relations Review* **58**(3).
- Altonji, J. G. & R.A. Shakotko (1987), 'Do wages rise with job seniority?', *Review of Economic Studies* **54**(3), 437–59.
- Beaudry, P. & 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), 665–688.
- Bentolila, S. & G. Bertola (1990), 'Firing costs and labour demand: How bad is eurosclerosis?', *Review of Economic Studies* **57**(3), 381–402.
- Bertrand, M. & S. Mullainathan (2003), 'Enjoying the quiet life? corporate governance and managerial preferences', *Journal of Political Economy* **III**(5), 1043–1075.
- Breusch, T.S. & A.R. Pagan (1979), 'A simple test for heteroskedasticity and random coefficient variation', *Econometrica* **47**, 1287–1294.
- Brown, J.N. & A. Light (1992), 'Interpreting panel data on job tenure', *Journal of Labor Economics* **10**, 219–57.
- Buchinsky, M., D. Fougere, F. Kramarz & R. Tchernis (2005), 'Interfirm mobility, wages, and the returns to seniority and experience in the u.s.'. IZA Discussion Paper No. 1521, March 2005.
- Cox, D. R. & 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**(3), 620–38.
- Dixit, A.K. & R.S. Pindyck (1994), *Investment under Uncertainty*, Princeton University Press, Princeton, USA.

- Dustmann, C. & C. Meghir (2005), 'Wages, experience and seniority', *Review of Economic Studies* **72**, 77–108.
- Farber, H. (1994), 'The analysis of interfirm worker mobility', *Journal of Labor Economics* **12**(4), 554–593.
- Farber, H.S. (1999), Mobility and strategy: The dynamics of job change in labor markets, in O.Ashenfelter & D.Card, eds, 'Handbook on Labour Economics', Vol. 3-B, Amsterdam, North-Holland.
- Horowitz, J.L. & 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 & D.G. Sullivan (1993), 'Earnings losses of displaced workers', *American Economic Review* **84**(4), 685–709.
- Jovanovic, B. (1979b), 'Firm-specific capital and turnover', *Journal of Political Economy* **87**(6), 1246–60.
- Koenker, R. (1981), 'A note on studentizing a test for heteroskedasticity', *Journal of Econometrics* **17**, 107–112.
- Lancaster, T., G. Imbens & P. Dolton (1987), *Job Separation and Job Matching*, Martinus Nijhoff Publishers, pp. 31–43.
- MaCurdy, T. E. & J. H. Pencavel (1986), 'Testing between competing models of wage and employment determination in unionized markets', *Journal of Political Economy* **94**(3), S3–S39.
- 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. (1991), 'A theory of quits and layoffs with efficient turnover', *Journal of Political Economy* **99**(1), 1–29.
- Meghir, C. & L. Pistaferri (2004), 'Income variance dynamics and heterogeneity', *Econometrica* **72**(1), 1–32.
- Mincer, J. & B. Jovanovic (1981), 'Labor mobility and wages'. in S. Rosen (Ed), "Studies in Labor Markets", Chapter 1, pp. 21-63.
- Mortensen, D.T. (1988), 'Wages, separations, and job tenure: On-the-job specific training or matching?', *Journal of Labor Economics* **6**(4), 445–471.
- Neal, D. (1995), 'Industry-specific human capital: Evidence from displaced workers', *Journal of Labor Economics* **14**(4), 653–677.
- Stern, S. (1997), 'Simulation-based estimation', *Journal of Economic Literature* **35**(4), 2006–39.

Teulings, C.N. & M. Van der Ende (2000), A structural model of tenure and specific investments. Tinbergen Institute Discussion Paper 2000-009/3.

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. & M.P. Ward (1992), 'Job mobility and the careers of young men', *Quarterly Journal of Economics* **107**(2), 439–79.