



TI 2005-099/3

Tinbergen Institute Discussion Paper

# Tenure Profiles and Tenure Profiles and Efficient Separation in a Stochastic Productivity Model

*Sebastian Buhai*<sup>1</sup>

*Coen N. Teulings*<sup>2</sup>

<sup>1</sup> *Erasmus University Rotterdam, and Tinbergen Institute,*

<sup>2</sup> *SEO Economic Research, University of Amsterdam, and Tinbergen Institute.*

**Tinbergen Institute**

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

**Tinbergen Institute Amsterdam**

Roetersstraat 31

1018 WB Amsterdam

The Netherlands

Tel.: +31(0)20 551 3500

Fax: +31(0)20 551 3555

**Tinbergen Institute Rotterdam**

Burg. Oudlaan 50

3062 PA Rotterdam

The Netherlands

Tel.: +31(0)10 408 8900

Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to [driessen@tinbergen.nl](mailto:driessen@tinbergen.nl).

Most TI discussion papers can be downloaded at <http://www.tinbergen.nl>.

# Tenure Profiles and Efficient Separation in a Stochastic Productivity Model\*

Sebastian Buhai <sup>†</sup> and Coen Teulings<sup>‡</sup>

October 20, 2005

## Abstract

This paper provides a new way of analyzing tenure profiles in wages, by modelling simultaneously the evolution of wages and the distribution of tenures. Starting point is the observation that within-job log wages for an individual can be described by random walk. We develop a theoretical model based on efficient bargaining, where both log outside wage and log wage in the current job follow a random walk. This setting allows the application of real option theory. We derive the efficient separation rule, which stipulates that workers switch jobs when the difference between the outside wage and the wage in the current job reaches a threshold. The model fits well the observed distribution of job tenures. 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.

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

**JEL codes:** C51, C52, J63

---

\*This paper circulated in previous versions under the title "Individual Wage Evolution and Job Separation in a Stochastic Productivity Model". We are very grateful for useful comments and discussion to Rob Alessie, Richard Blundell, Bas van der Klaauw, Randall Olsen, Miguel Portela, Aico van Vuuren, Robert Waldmann and to audiences in seminars and conferences at the Univ. of Chicago, Northwestern Univ., Federal Reserve Bank of Chicago, London School of Economics, Univ. College London, Univ. of Copenhagen, Tinbergen Institute, ESWC 2005 in London, EEA 2005 in Amsterdam, Villa Mondragone Workshop 2005 in Rome, SOLE 2004 in San Antonio and SMYE 2004 in Warsaw. We also thank Nick Williams and Lennart Janssens for clarifying questions about the dataset. Parts of this paper were completed while Buhai was visiting the Department of Economics, University College London, visit partly funded by a generous grant from the Netherlands Organization for Scientific Research (NWO). The first co-author also gratefully acknowledges continuous financial support from the Tinbergen Institute and the Erasmus University Rotterdam. The usual disclaimers apply.

<sup>†</sup>Tinbergen Institute, Erasmus University Rotterdam and Aarhus School of Business; corresponding author: buhai@tinbergen.nl, <http://www.tinbergen.nl/~buhai>

<sup>‡</sup>SEO Economic Research, University of Amsterdam and Tinbergen Institute; c.teulings@seo.nl

# 1 Introduction

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

We take efficient bargaining as benchmark. Hence, quits and job layoffs are observationally equivalent, as in McLaughlin (1991). The model explains the observed correlation between wages and job tenure from the random evolution of wages after job start. This random evolution is due to the random evolution of the job's productivity value. Separation occurs when the value of the productivity in the job falls below the value of the outside option<sup>1</sup>, which also evolves randomly. We refer to this outside option as the reservation wage. The observed correlation between wages and tenure is caused by the fact that only jobs that evolve favorably relative to the reservation wage 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.

---

<sup>1</sup>We regard the outside option as the alternative activity to the current job. The payoff of this alternative activity, also labelled outside wage or reservation wage, is equal to the value of the individual's "productivity" in this outside activity. In the paper we shall thus use the terms "value of the outside option" and respectively, "outside wage" or "reservation wage", interchangeably.

The evolution of an individual's within-job log wage is reasonably described by a random walk with transitory shocks, as previously found in 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 productivity value in the job and log reservation wage are assumed to follow a random walk. We assume a linear sharing rule of the surplus between the log productivity value in the job and the log reservation wage. This ensures 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 reservation wage 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 combination of irreversible specific investment and productivity and reservation wages following a random walk implies that we can apply the theory of real options, see for example Dixit (1989), Bentolila and Bertola (1990), and Dixit and Pindyck (1994). Teulings and van der Ende (2000) used 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.

From the distribution of job tenures we are able to estimate the initial surplus of the job's productivity above the reservation wage 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 only observes the reservation wage at job start and at job separation. At job start, the worker chooses the best option that is available at that moment, which is by definition equal to the reservation wage. Using the same reasoning, the reservation wage at job transition is equal to the first wage in the new job<sup>2</sup>. Our estimation procedure exploits both pieces of information on the reservation wage 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

---

<sup>2</sup>In practice, we shall define the reservation wage not as the first wage in the new job, but as the lowest acceptable wage in the old job.

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 reservation wage. Observed reservation wages are positively selected, since we observe them only at the moment when workers switch jobs and since workers switch jobs only when reservation wages are high. 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 reservation wage turns out to be an empirically important phenomenon.

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 surplus of productivity above the reservation wage, is linear by assumption, this concavity is fully due to selection. One can 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 however really 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 the efficient bargaining hypothesis. The estimated tenure profile is on the high end of the spectrum, 5% per year, though two thirds of the return take the form of a declining reservation wage instead of a rising wage in the current job.

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 employment relationship 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 reservation 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 productivity and the reservation wage. 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 the reservation wage so that separation becomes the efficient alternative. Then, separation occurs at mutual consent since there are no gains from trade left. The turnover is therefore efficient, and quits and layoffs observationally identical (although behaviorally distinct), 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. 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 worker's reservation wage  $R_t$  is also a geometric Brownian with drift.  $R_t$  is defined such that separation is efficient as soon as  $P_t \leq R_t$ . The specific investments at the moment of job start are proportional to the reservation wage:  $R_t I$ . We 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  $V(P_t, R_t)$  be the option value of a vacancy and  $J(P_t, R_t)$  the expected present value of a job, both as functions of productivity  $P_t$  and the reservation wage  $R_t$ . Hence,  $J(P_t, R_t)$  is the net discounted value of  $P_s - R_s$ , for  $s$  running from  $t$  till the expected future moment of separation  $T$ , plus the option value  $V(P_T, R_T)$  of filling the vacancy at a later time  $s > T$ . The current values of  $P_t$  and  $R_t$  are sufficient statistics for the calculation of  $J(\cdot)$  and  $V(\cdot)$  since both  $P_t$  and  $R_t$  are martingales. A firm will hire a worker at the first moment  $S$  that the value of a filled job exceeds the value of a vacancy plus the cost of investment:

$$J(P_S, R_S) = V(P_S, R_S) + R_S I \quad (2)$$

Similarly, a firm will fire the worker at the first moment  $T$  that  $P_t = R_t$ , or equivalently, when the value of a filled job is less than that of vacancy:

$$V(P_T, R_T) = J(P_T, R_T) \quad (3)$$

Therefore we shall refer to  $P_t - R_t$  as the absolute current surplus of a job. 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:

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

where again  $b_t$  denotes  $\log B_t$ . By construction,  $B_T = 1$ ,  $b_t = p_t - r_t$ , and  $b_T = 0$ . We shall refer to  $b_t$  as the relative current surplus of a job, or easier, the job surplus. Using this definition we can write the absolute current surplus of a filled job as:

$$P_t - R_t = R_t (B_t - 1)$$

Thus, all potential future investment cost  $R_s I$  and all future benefits  $R_s (B_s - 1)$  for future dates  $s > t$  are proportional to  $R_s$ . Since  $R_s$  follows a geometric Brownian,  $E[R_s | t] = R_t E[\exp(r_s - r_t)] = R_t \exp[(\mu + \frac{1}{2}\sigma^2)(t-s)]$ . Hence, the expectation of future values of  $R_s$  is proportional to  $R_t$ . It follows that  $J(\cdot)$



and  $V(\cdot)$  are proportional to  $R_t$ . These value functions can therefore be written as:

$$\begin{aligned} J(P_t, R_t) &= R_t j(B_t) \\ V(P_t, R_t) &= R_t v(B_t) \end{aligned}$$

Hence, we can write the hiring and separation conditions (2) and (3) as:

$$\begin{aligned} j(B_S) &= v(B_S) + I \\ v(B_T) &= j(B_T) \end{aligned}$$

The optimal hiring and separation rules depend therefore purely on  $B_t$ : fill a vacancy at the first time that  $B_t$  rises up to  $B^h$ , fire the worker at the first time  $B_t$  falls below  $B^s$ . The value of  $B^h$  and  $B^s$  can be expressed as a function of the model's parameters  $I, \mu, \Sigma$  and the interest rate  $\rho$ , using techniques described in 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 that.

## 2.2 Job Tenure Distribution

Without loss of generality we normalize the moment of job start to zero,  $S = 0$ , so that  $b_0 = b^h$ . 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,  $\sigma$ , the standard deviation of  $b_t$ , is unidentified in this model. We normalize thus all other parameters by  $\sigma$ . Define:

$$\begin{aligned} \Omega_t &\equiv \frac{b_t - b^s}{\sigma} \\ \Omega &\equiv \frac{b^h - b^s}{\sigma} \\ \pi &\equiv \frac{\mu}{\sigma} \end{aligned} \tag{5}$$

Hence,  $\Omega_t$  is a Brownian with drift  $\pi$  and unit variance per unit time, with  $\Omega_0 = \Omega$  and  $\Omega_T = 0$ . The distribution of job tenures is the "First Passage time" distribution (eg. Cox and Miller, 1965), i.e. the distribution of durations till the random walk  $\Omega_t$  passes the single absorbing barrier  $\Omega_t = 0$  for the first time. The density function of  $\Omega_t$  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 density. However, a realization of  $\Omega_t$  is not interesting if separation has occurred before time  $t$ . This situation cannot correspond to a job-worker match since separation decisions are by assumption

irreversible. We are thus interested in the density of  $\Omega_t$  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  $\Omega$  to  $\Omega_t$  having crossed  $\Omega_s = 0$  at least once, and trajectories from  $-\Omega$  to  $\Omega_t$ . This latter group of trajectories should thus be subtracted from the trajectories from  $\Omega$  to  $\Omega_t$  for the calculation of the density of all trajectories from  $\Omega$  to  $\Omega_t$  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  $\Omega$  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  $\Omega_t \in [0, \infty)$  (since negative values of  $\Omega_t$  are inconsistent with  $T > t$ ):

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

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 start  $\Omega$  and the drift  $\pi$ . 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 declining monotonically to 0 for positive drift  $\pi > 0$  and respectively 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. We plot the exit rates for pairs  $\Omega = 0.32$  and  $\pi = 0.15$  and respectively  $\Omega = 0.30$  and  $\pi = 0.23$  (these are mean values for  $\Omega$  and  $\pi$  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. The fraction of surviving job spells for  $\pi > 0$  is given by the survivor function (7) when  $t \rightarrow \infty$ , hence by  $1 - e^{-2\Omega\pi}$ . For

mean values of parameters of  $\Omega = 0.32$  and  $\pi = 0.15$  about 10 % of the jobs never end. We conclude that the parameters  $\Omega$  and  $\pi$  can be identified from data on the distribution of job tenures, but the parameter  $\sigma$  cannot.

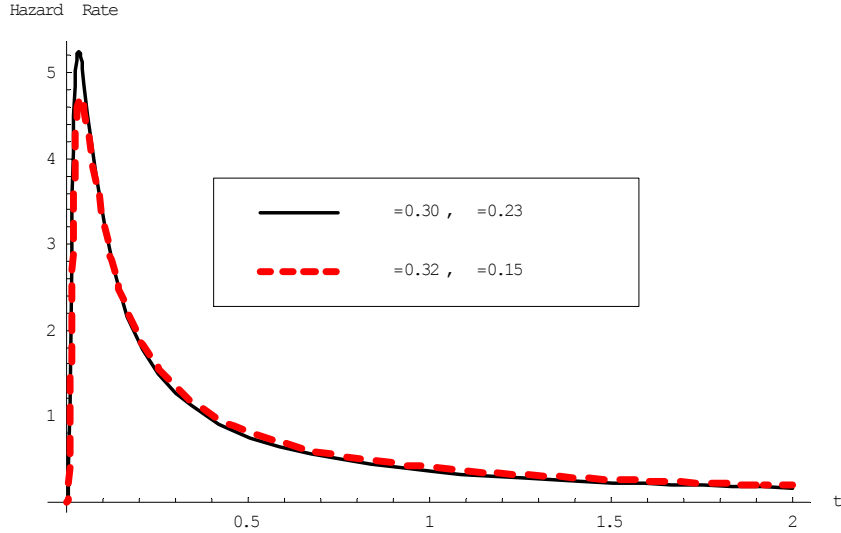


Figure 1: Predicted Job Hazards

## 2.3 Tenure Profile in Wages

### 2.3.1 Sharing Rule of Surpluses

We now extend the model with an explicit sharing rule of surpluses during the course of the job spell. We use a rule stipulating that log instantaneous surplus  $p_t - r_t$  is shared in fixed proportions between the worker and the firm.<sup>3</sup> The worker's log wage  $w_t$  then satisfies:

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

where  $\bar{\sigma} \equiv \beta\sigma$ . We interpret  $\beta$  as the worker's bargaining power. 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,  $\Omega_t$  goes up in expectation with  $\pi$  every period, conditional on the fact that  $\Omega_t$  remains positive. Were job separations independent of the realization of  $\Omega_t$  (and of  $r_t$ ), we would have:

$$E(w_t | t < T) = E(r_t) + \bar{\sigma}(\Omega + \pi t)$$

<sup>3</sup>This approach is more pragmatic than what is commonly used in the literature, where the wage setting is done by Nash bargaining leading to a sharing rule of the return on the expected discounted value of all future absolute surpluses instead of the instantaneous relative surplus. One can view our approach as a first order expansion of this rule.

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:

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

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

However, in completed job spells  $\Omega_t$  is correlated to  $T$  for three reasons: (i)  $\Omega_0 = \Omega$ , (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:  $r_0 = w_0 - \bar{\sigma}\Omega$ , and at the end of a job:  $r_T = w_T$ . 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 the last date on which data are available. What we know about an incomplete job spell 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 $\Omega_t$ 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  $\Omega_t$  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:

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

where  $\tau$  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,  $\pi$ . 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.

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

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

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 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 $\Omega_t$ 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  $0 \leq t < L$ , and (iii)  $T > L$ . The application of the Bayes rule leads

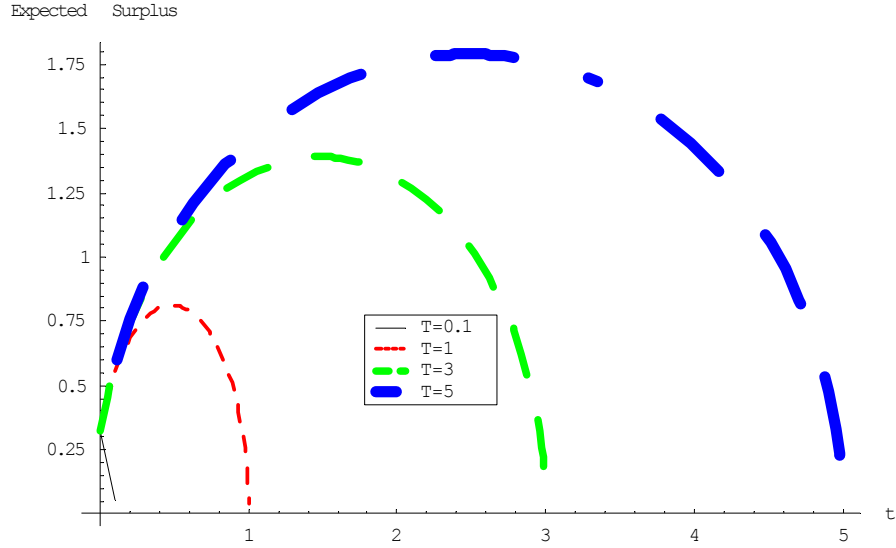


Figure 2: Expected Surplus in Completed Job Spells

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,  $E(\Omega_t | t < L < T) = E(\Omega_t) = \Omega + \pi t$ . The trajectories are strongly concave, implying that selection plays an important role. Contrary to the completed spells case, incomplete spells do provide information on the drift  $\pi$ . 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.

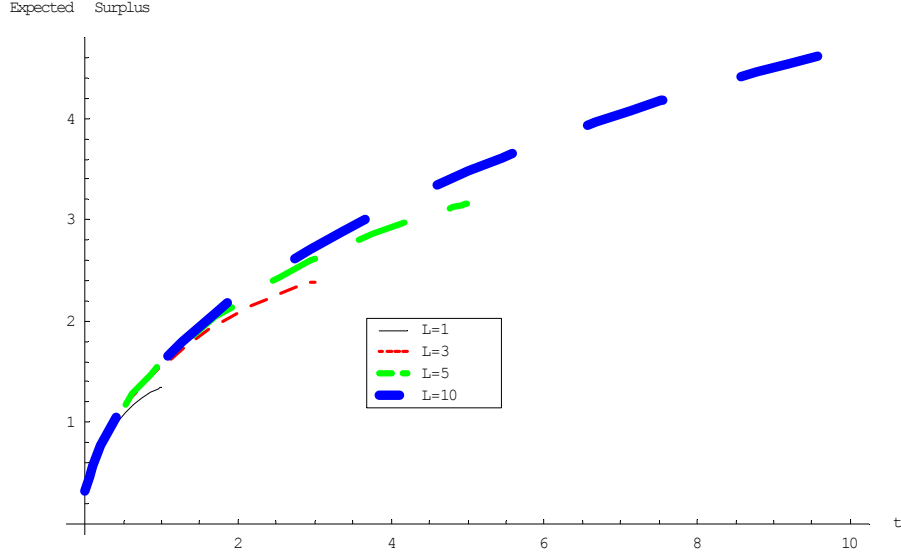


Figure 3: Expected Surplus in Incomplete Job Spells

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  $\Omega_t$  in incomplete and completed job spells for the analysis of the expected wage growth  $\Delta w_t$  within a job and respectively  $\Delta w_T^*$  between jobs. For this purpose, we decompose the random variables  $[\Delta p_t, \Delta r_t]$  in two orthogonal components  $\Delta b_t$  and  $\Delta z_t$ , such that  $\text{Cor}(\Delta b_t, \Delta z_t) = 0$ . Given the previous assumptions on the joint normality of  $\Delta p_t$  and  $\Delta r_t$ , such a decomposition is always feasible. Hence, for  $1 < t < T$ ,

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

with  $\Delta z_t \sim N(\mu_z, \sigma_z^2)$  and  $\text{Cov}(\Delta z_t, \Delta b_t) = 0$ . Obviously, the parameter  $\gamma$  can be expressed in terms of the covariance matrix  $\Sigma$  and the bargaining power  $\beta$ , but that is of little help here. It is more useful to interpret it as a reflection

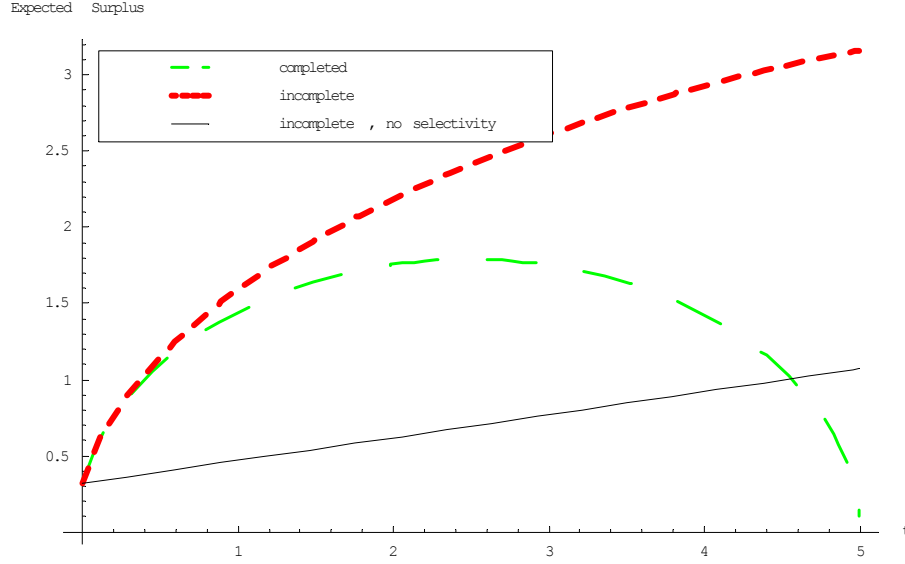


Figure 4: Selectivity versus Drift in the Expected Surplus

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  $\Delta r_t$  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,  $\Delta b_t$ . 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,  $\Delta p_t$  reflects the evolution of the general human capital of the worker in this job as well as in all other jobs;  $\Delta b_t$  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:

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

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



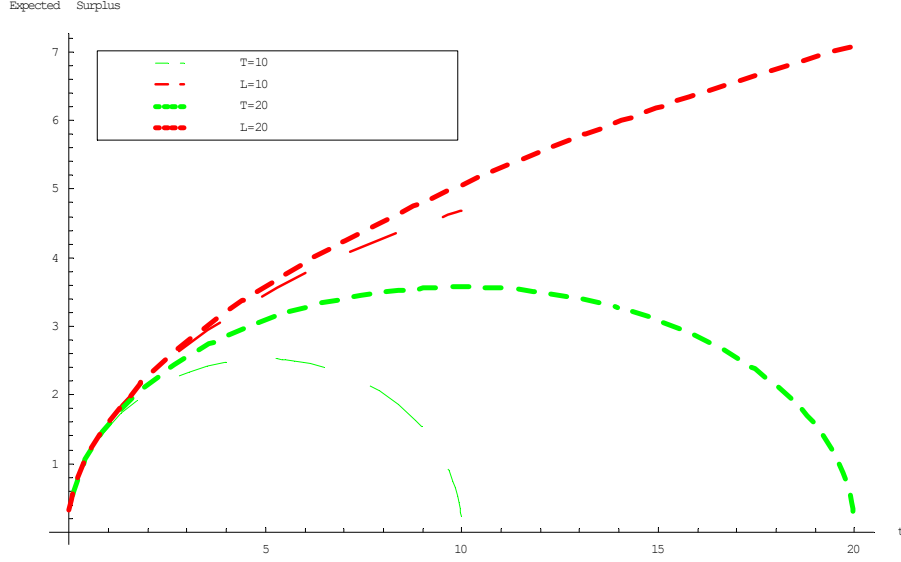


Figure 5: Comparison Exp. Surplus in Long Completed and Incomplete Spells

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,  $\Omega_t$ ) and since  $\Delta b_t$  and  $\Delta z_t$  are uncorrelated, there is no selectivity in  $\Delta z_t$ . Hence, the conditioning  $1 < t < L < T$  can be omitted in  $E(\Delta z_t)$ . As discussed before,  $\Omega$  and  $\pi$  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,  $\Delta w_t$ .<sup>4</sup> These equations identify  $(1 - \gamma)\bar{\sigma}$ , but not  $\gamma$  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  $r_t$ .

<sup>4</sup>There 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:  $r_0 = w_0 - \beta\Omega$  and  $r_T = w_T^* - \beta\Omega$ , where  $w_T^*$  is the starting wage in the new job. We can add these conditions to our regression analysis. Hence:

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

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

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

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.

For job changers, we can write a similar equation:

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

The term  $E(\Delta\Omega_t | t = T)$  reflects the wage decline in the old job the year before separation.<sup>5</sup> It is always negative, see Figure 2. The term  $\Omega$  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  $\gamma$  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  $\gamma$ , 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  $\sigma_z^2$ :

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

The first equation above follows immediately from the first equation of (13), since  $\Omega_0 = \Omega$ . The first term reflects the general drift in log wages. Though till so far the value of  $\Omega$  was held constant across jobs, so that the second term drops out since the difference between the value of  $\Omega$  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  $\Delta z_t$  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  $\Delta w_t$  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.

---

<sup>5</sup>Implicitly, 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. We leave this complication aside in what follows.

## 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. 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 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 cointegration of residuals from this regression, from which we infer the process

Table 1: Summary Statistics Dataset

Variable	Mean	Std. Dev.	Min.	Max.	Observations
logwage <sup>(1)</sup>	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

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

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);  $\eta_i$  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  $\beta_1$  and  $\gamma_1$  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  $\beta_1$  and  $\gamma_1$ . 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)
$\Delta$ Tenure $_{ijt}$			-0.001	(0.009)
$\Delta$ Tenure $_{ijt}^2$	-8e-4*	(4e-4)	7e-4	(0.001)
$\Delta$ Tenure $_{ijt}^3$	e-5 <sup>†</sup>	(8e-6)	-e-5	(2e-5)
$\Delta$ Experience $_{ijt}^2$	-0.002**	(4e-4)	-0.003	(0.002)
$\Delta$ Experience $_{ijt}^3$	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)
Observations	11834		1448	
SER <sup>(1)</sup>	0.179		0.361	
R <sup>2</sup>	0.0219		0.039	
Wald time <sup>(2)</sup>	$\chi^2_{(16)}$ : 144.96**		$\chi^2_{(13)}$ : 20.02 <sup>†</sup>	
Wald joint <sup>(3)</sup>	$\chi^2_{(24)}$ : 264.72**		$\chi^2_{(22)}$ : 57.86**	

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

The regressions include time effects

<sup>(1)</sup> SER= standard error of the regression (root mean square error)

<sup>(2)</sup> Wald test for joint significance of time dummy coefficients

<sup>(3)</sup> 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  $\gamma_1$  and  $\gamma_2$  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. It is inconsistent with a model of efficient union bargaining (the *labor demand curve* model), because according to that model the union and the firm should have renegotiated the wage contract instead of the firm firing the worker.

Table 3 presents the covariogram of the residuals of the within-job wage change regression. Residuals are strongly negatively correlated to their first lag,

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

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) 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.<sup>6</sup> 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  $\eta_{it}$ :

$$\begin{aligned} u_{it} &= e_{it} + \eta_{it} \\ \Delta e_{it} &= \varepsilon_{it} \end{aligned} \tag{21}$$

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

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

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

<sup>6</sup>We 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.

$\sigma_\eta^2 = 0.005$ . Our result for the variance of persistent income shocks is similar to that reported by Meghir and Pistaferri (2004) for the PSID<sup>7</sup>. 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  $\eta_{it}$  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)
$\Delta$ Tenure <sub>ijt</sub> <sup>2</sup>			-0.003	(0.009)
$\Delta$ Tenure <sub>ijt</sub> <sup>2</sup>	-2e-4	(3e-4)	6e-4	(0.001)
$\Delta$ Tenure <sub>ijt</sub> <sup>3</sup>	-3e-6	(7e-6)	-2e-5	(3e-5)
$\Delta$ Experience <sub>ijt</sub> <sup>2</sup>	e-4	(4e-4)	0.001	(0.002)
$\Delta$ Experience <sub>ijt</sub> <sup>3</sup>	2e-6	( 6e-6)	-2e-5	(4e-5)
Education <sub>ijt</sub>	0.003**	(9e-4)	-7e-4	(0.006)
Metropolis <sub>ijt</sub>	0.004	(0.003)	0.018	(0.020)
Union member <sub>ijt</sub>	-0.017**	(0.004)	0.033	(0.030)
Married <sub>ijt</sub>	0.004	(0.005)	-0.027	(0.024)
Observations	11834		1448	
SER	0.166		0.357	
R <sup>2</sup>	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 <sup>(1)</sup>	$\chi^2_{(24)}$ : N*R <sup>2</sup> =87.57**		$\chi^2_{(22)}$ : N*R <sup>2</sup> =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

<sup>7</sup>For 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.

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  $\eta_{it}$  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 efficient bargaining, at least 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 initial surplus  $\Omega$  and the drift  $\pi$  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,  $\Omega$  and  $\pi$  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 otherwise than from functional form assumptions. Taking into account that  $\Omega$  has to be positive, the following specification for  $\Omega$  and  $\pi$  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  $\sigma_\Omega$  and  $\sigma_\pi$ , 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 married. 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  $\Omega$  and  $\pi$  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  $j^{\text{th}}$  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 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:

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

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 question<sup>8</sup>, but we do not know whether there has been another job spell between the job observed a year ago and the job observed now. Since the hazard rate implied by our model is hump shaped, with the hump likely to be within the first year, cf. Farber (1994), this phenomenon is expected to have a large impact on the estimation results. We are likely to overestimate  $\Omega$  and  $\pi$ , 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 itself<sup>9</sup>. 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:

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

Then, the total contribution to the likelihood of a spell with initial tenure  $t$  and completed tenure  $T$  reads:

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

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 wave of the PSID, can therefore 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  $D_{ij} = \begin{cases} \text{if spell starts after start PSID} & \int_0^1 [1 - F(x)] dx \\ \text{if spell starts before start PSID} & 1 - F(t_{ij}) \end{cases}$

<sup>8</sup>Initial tenures are either reported or inferred by making them consistent with the latest reported tenures- see appendix for more details on this process of constructing tenures in the PSID

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

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 PSID<sup>10</sup>.

Table 5: MLE Tenure Distribution Parameters

Variable	Small Sample <sup>(1)</sup>		Big Sample <sup>(2)</sup>	
	Drift $\pi$	Dist $\Omega$	Drift $\pi$	Dist $\Omega$
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)
Random worker effects $\sigma$	0.277**	0.0001	1.34e-8	1.75e-7
(st. errors)	(0.050)	(5.51)	(0.001)	(0.010)
Number Observations (job spells)		1911	5484	

<sup>(1)</sup>Small sample= sample of job spells starting within the range of the PSID sample

<sup>(2)</sup>Big sample= sample of all job spells

All covariates are taken in deviations from their means over jobs

Standard errors in brackets under each coefficient

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

The estimation results are presented in Table 5. Theoretically, the results for both likelihood functions should be identical, but that turns out not to be the case. Including the long spells that started before the first wave of the PSID matters substantially, suggesting that there is some degree of misspecification left. Though the estimated intercepts of  $\Omega$  and respectively  $\pi$  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  $\Omega$  and  $\pi$ ), the coefficients for other variables differ substantially. However, the difference is small, as shown by Figure 1, which shows the predicted hazard rate for both sets of estimation results from Table 5.

Two observations are in place. First, as stated before, though both sets of estimation results are statistically different, the theoretical hazards for both

<sup>10</sup>In order to estimate the log-likelihood functions above, we used simulated maximum likelihood techniques, cf. Stern (1997). Sampling from a joint normal distribution with mean 0 and variances  $\sigma_\pi^2$  and  $\sigma_\Omega^2$  and using a sampling size of 500 sampling points (the results are in fact 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.

models look very similar, the only difference being the height of the peak (lower for the case where we use all job spells). Secondly, a number of results in Table 5 are consistent across both sets of estimation results. First, spells started at higher age and spells at unionized firms have a larger drift  $\pi_{ij}$ . When we include the spells started before the first wave of the PSID, these variables also have a positive effect on the initial surplus  $\Omega_{ij}$ . 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 comparative advantages best. For the positive effect of union membership, economic theory provides two explanations, one in which union membership causes a high  $\Omega_{ij}$  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 initial surplus  $\Omega_{ij}$  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  $\Omega_{ij}$ , 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  $\pi_{ij}$  is positive and large. In both cases, there are hardly observations for which  $\pi_{ij}$  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.

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.<sup>11</sup> 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

---

<sup>11</sup>This density is calculated by a recursive scheme. We divide the 32 years time period in  $32 \times 256$  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  $\Omega$  and  $\pi$ .

shapes of the two densities are identical. Note the small peak in the density for short incomplete spells, which is due to the hump shape pattern in the hazard: if your job ends for instance in the last five years before the end of the observation period, there is a substantial probability that you experience further separations afterwards due to the peak in the hazard rate, leading to a peak of short incomplete tenures. Close alignment of the predicted and the empirical densities suggests that our model works well.

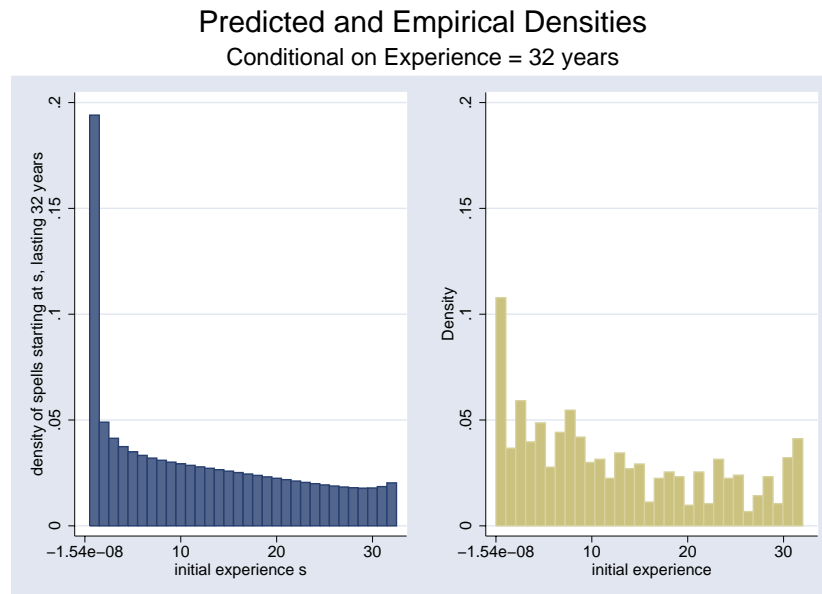


Figure 6: Density of Incomplete Job Spells with Exit Option

### 3.4 Wage Regressions

Table 6 presents estimation results on wage changes with heteroskedasticity-robust standard errors<sup>12</sup>.

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

Variable	1: Completed Spells		2: Incomplete Spells		3: Job Transition	
	Coef	St Err	Coef	St Err	Coef	St Err
A: Unrestricted Wage Changes Regression Estimates						

*Continued on next page...*

<sup>12</sup>We tested for the absence of individual specific effects. We left out being married as a regressor since it was insignificant in all regressions.

... table 6 continued

Intercept	0.060**	(0.015)	0.054**	(0.010)	0.216*	(0.102)
$\Delta E(\Omega_{ijt})$	0.008	(0.006)	0.039**	(0.013)	0.137*	(0.062)
$\Omega_{i,j+1,0}$					0.308*	(0.145)
$\text{Exp}_{ijt}$	-0.004*	(0.002)	-0.005**	(9e-4)	-0.007	(0.007)
$\text{Exp}_{ijt}^2$	8e-5	(5e-5)	9e-5**	(2e-5)	6e-5	(e-4)
$\text{Educ}_{ijt}^{(1)}$	0.002	(0.002)	e-4	(0.001)	0.031**	(0.009)
$\text{Union}_{ijt}$	0.006	(0.010)	-0.004	(0.003)	-0.163**	(0.062)
$\text{Metropolis}_{ijt}$	0.007	(0.008)	0.011**	(0.003)	-0.017	(0.031)
Observations	2022		9653		618	
SER	0.177		0.180		0.362	
R <sup>2</sup>	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 Wage Changes Regression Estimates**

Intercept	0.049**	(0.017)		0.147 <sup>†</sup>	(0.079)
$+\Delta E(\Omega_{ijt})^{(2)}$	0.051*	(0.025)		0.075**	(0.026)
$-\Delta E(\Omega_{ijt})^{(2)}$	-6e-4	(0.007)			
$\Omega_{i,j+1,0}$				0.288*	(0.142)
$\text{Exp}_{ijt}$	-0.003 <sup>†</sup>	(0.002)		-0.005	(0.007)
$\text{Exp}_{ijt}^2$	6e-5	(5e-5)		7e-5	(e-4)
$\text{Educ}_{ijt}^{(1)}$	0.002	(0.002)		0.030**	(0.009)
$\text{Union}_{ijt}$	0.007	(0.010)		-0.161**	(0.061)
$\text{Metropolis}_{ijt}$	0.007	(0.008)		-0.016	(0.031)
Observations	2022			618	
SER	0.177			0.360	
R <sup>2</sup>	0.009			0.067	
Wald joint	$\chi_{(7)}^2$ : 20.09**			$\chi_{(7)}^2$ : 45.78**	

**C: General Wage Change Reg. for All Spells**

Variable	Coef	St Err			
Intercept	0.053**	(0.008)			
$+\Delta E(\Omega_{ijt})$	0.041**	(0.009)			
$-\Delta E(\Omega_{ijt})$	-0.006	(0.006)			
$\Omega_{i,j+1,0}$	0.332**	(0.065)			
$\text{Exp}_{ijt}$	-0.005**	(7e-4)			
$\text{Exp}_{ijt}^2$	8e-5**	(2e-5)			
$\text{Educ}_{ijt}^{(3)(1)}$	0.031**	(0.008)			
$\text{Union}_{ijt}^{(3)}$	-0.174**	(0.053)			
$\text{Metropolis}_{ijt}$	0.010**	(0.003)			
Observations	12293				
SER	0.99981				
R <sup>2</sup>	0.0286				
Wald joint	$\chi_{(8)}^2$ : 166.48**				

**D: Initial Wage Changes Regression at Job Transitions<sup>(4)</sup>**

1: Mean			2: Variance <sup>(5)</sup>		
Variable	Coef	St Err	Variable	Coef	St Err
Intercept			Intercept	0.092**	(0.024)
$\text{Ten}_{i,j-1,T}$	0.057**	(0.011)	$\text{Ten}_{i,j-1,T-1}$	0.019**	(0.007)
$\Delta\Omega_{ij0}$	0.075	(0.251)			
$\Delta\text{Exp}_{ij0}^2$ <sup>(6)</sup>	-0.001*	(4e-4)			
Observations	336			336	
SER	0.379			0.297	
R <sup>2</sup>	0.14			0.017	
Wald joint	$\chi_{(3)}^2$ : 41.16**			$\chi_{(1)}^2$ : 7.44**	

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

(1) Education is taken in deviation from the means over all jobs of an individual

(2) The +/- split of  $\Delta E(\Omega_{ijt})$  is only relevant for completed spells

(3) In the general regression union and education are interacted with job change

(4) We only keep completed jobs that last more than 1 year ( $\text{Ten}_{i,j-1,T} > 1$ )

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

(6)  $\Delta\text{Exp}_{ij0}^2 = \text{Exp}_{ij0}^2 - \text{Exp}_{i,j-1,0}^2 = (\text{Exp}_{i,j-1,0} + \text{Ten}_{i,j-1,T})^2 - \text{Exp}_{i,j-1,0}^2$

Panel (A) 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  $\Omega_{i,j+1,0}$ , 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. Panel C 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. 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  $\Omega_{ij0}$ , see Table 5).

The estimation results from Panel C 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  $\pi$ ). However, the high value of  $\gamma$  implies that most of the return to tenure takes the form of the log reservation wage  $r_t$  falling, instead of the inside wage  $w_t$  rising, cf. equation (13). Both estimates, that of  $\gamma$  and of the tenure profile are fully driven by the estimate of  $\bar{\sigma}$ , which itself is driven by the variation in the surplus in the next job,  $\Omega_{ij0}$ , and its effect on the wage change for job movers, see equation (15). This might be a thin line for identification, but it is hard to think of an alternative route to estimate this parameter.

Equation (17) runs the test of comparing the concavity in the tenure profile to the variance in wage changes. Panel D of Table 6 presents estimation results for equation (16). We restrict the sample to jobs lasting more than 1 year, since wage changes for jobs the last only one year are noisy anyway<sup>13</sup>. 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,  $\eta_{ijt}$ , see equation (21). The coefficient for  $T$  is a consistent estimator of  $\frac{1}{T}\text{Var}(w_0^* - w_0)$ . An estimator for  $\sigma_w^2$  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 Panel C, 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 job, 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 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 after  $L$  years of experience, computed and plotted for  $L = 32$  in Figure 6, and the density of  $\Omega_t = b_t/\sigma$  conditional on the

<sup>13</sup>Using the whole sample, the coefficient of  $T$  in the variance of wages regression is 0.010.



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 job compresses the variance of  $b_t$  considerably: by about 65% after 32 years of experience<sup>14</sup>.

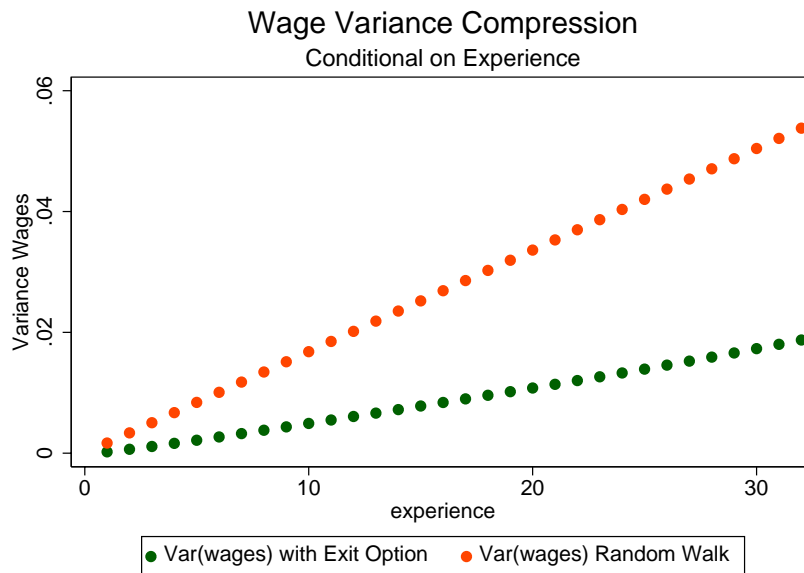


Figure 7: Compression Effect of the Exit Option on the Wage Variance

## 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 frictionsless labor market at the moment of job start (which enables workers to pick the best job alternative straightaway), specific investment and hence lock in the current job subsequently, 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. This puts into question a large literature on the estimation of

<sup>14</sup>We use  $\sigma = 0.041$ , our coefficient from Panel C. Hence, after  $L = 32$  years of experience, without the exit option the variance of  $b_t$  would be about 0.054. With the option to exit the wage variance is compressed after the same period to about 0.019.

tenure profiles. Hence, the tenure profile can only be estimated from either the distribution of tenure or 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 if the underlying tenure profile is linear. In general, the selection effect tends to be much more important than the deterministic trend. Remarkably, job separation is driven more by the selectivity in the reservation wage  $r_t$  than by shocks to the productivity in the job  $p_t$ . 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  $\Omega_t$  becomes negative.  $\Psi_t$  is the maximum of its value in the previous period and the current value of  $\Omega_t$ . Then, we would observe more wage rigidity in job spells that are expected to end shortly (since there  $\Omega_t$  is declining), which explains that  $\Delta\Omega_t$  does not affect  $\Delta w_t$  for  $\Delta\Omega_t < 0$ . Hence, the standard error of  $\Delta w_t$  should be smaller in the period just before separation. This model needs to be analyzed more formally, but the analysis presented here already suggests that it provides an excellent description of the data.

## 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.
- Buhai, I.S. & K. Tytt (2005), ‘Article rtr’, *gdfgfd* **5**(7), 23–67.
- 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.
- 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.

## 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  $\Omega$  by  $\omega$ :

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

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

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

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

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

The expectation in (11) can now be calculated:

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

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:

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

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

## A.2 Incomplete job spells

The first factor in the numerator of (12) is given by substituting  $t$  with  $L$  and  $\Omega$  with  $\omega$  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:

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

and

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