A joint Initiative of Ludwig-Maximilians-Universität and Ifo Institute for Economic Research

CESifo Working Papers

A STRUCTURAL MODEL OF TENURE AND SPECIFIC INVESTMENTS

Coen N. Teulings
Martin A. van der Ende*

CESifo Working Paper No. 532

August 2001

CESifo
Center for Economic Studies & Ifo Institute for Economic Research
Poschingerstr. 5, 81679 Munich, Germany
Phone: +49 (89) 9224-1410 - Fax: +49 (89) 9224-1409
e-mail: office@CESifo.de
ISSN 1617-9595

An electronic version of the paper may be downloaded
- from the SSRN website: www.SSRN.com
- from the CESifo website: www.CESifo.de

* The authors thank Nicholas Bloom, Pieter Gautier, Dan Hamermesh, Lennart Janssens, Maarten Lindeboom, Audra Bowlus, Joop Hartog, Jan van Ours, Guiseppe Bertola and anonymous referees for helpful comments.

CESifo *Working Paper No. 532*
August 2001

A STRUCTURAL MODEL OF TENURE AND SPECIFIC INVESTMENTS

Abstract

Though a lot of work has been done on the distribution of job tenures, we are still uncertain about its main determinants. In this paper, we stress random shocks to match productivity after the start of an employment relation. The specificity of investment makes hiring and separation decisions irreversible. These decisions therefore have an option value. Assumptions on risk neutrality, efficient bargaining, and the efficient resolution of hold up problems allow investment and separation decisions to be analyzed separately from wage setting. The tenure profiles in wages implied by the model fit the observed pattern quite well. The model yields a hump shaped pattern in separation rates, similar to learning models, but with a slower decline after the peak. Estimation results using job tenure data from the NLSY support this humped shaped pattern and favor this model above the learning model. We develop a methodology to analyze the decomposition of shocks to match productivity into idiosyncratic and macro-level shocks.
When assuming a Last-In-First-Out (LIFO) separation rule, this model of individualemployment relations is embedded in a model of firm level employment, that satisfies Gibrat's law. The LIFO rule is interpreted as an institution protecting the property rights on specific investments of incumbent workers against hiring new workers by the firm.

Keywords: option value, job tenure, tenure profiles.

JEL Classification: J63

Coen N. Teulings
Erasmus University Rotterdam and Tinbergen Institute
P O Box 1738
NL-3000 DR Rotterdam
The Netherlands
Teulings@few.eur.nl

Martin A. van der Ende
Nederlands Economisch Instituut
P O Box 4175
NL-3006 AD Rotterdam
The Netherlands

1 Introduction

Though a lot of work has been done on the analysis of the job tenure distribution, we are still uncertain about the nature of the process triggering job separation. The literature provides a number of explanations, where search and learning models are two prominent lines of thought. The present paper focuses on the unpredictable evolution of future match productivity after the date of job start. A job start requires specific investments. When the match productivity evolves unfavorably, these investments lose their value and separation becomes the efficient alternative. We refer to this model as the random growth model. The model is easily embedded in a firm-level model of employment, when random shocks to future productivity are interpreted as shocks in the firm's demand curve.

For a better understanding of the relation of the random growth model to other strands of the literature in this field, it is useful to contrast the type of stochastic processes triggering separation, in search model on the hand and learning models on the other hand. In search models (e.g. Jovanovic (1979b), Burdett and Mortensen (1998)), a worker receives alternative job offers every now and then. Usually, the arrival of new job offers is modeled as a Poisson process. Offers are drawn randomly from an offer distribution. When the value of a newly arrived offer exceeds the value of the present job, there is separation. Two types of stochastic shocks contribute to the separation decision: first, the arrival process of new offers, and second, conditional on arrival, the value of that offer. A critical feature of these shocks is that the effect of both types of shocks is transitory. The arrival of a job offer today does not affect the probability of the arrival of an offer tomorrow, neither does the value of this job offer affects the value of the next offer. A major achievement of search models is that they provide a simple explanation for the empirical regularity that separation rates decline with the accumulation of experience. The longer a worker has been around on the labor market, the more job offers she has received. Where the present job offer is the best of these offers, this maximum will –in expectation- move up with experience. Hence, the probability of receiving an even better job declines over time.

In learning models, the workers and the firm have only imperfect information about the quality of their match (Jovanovic, 1979a; Miller, 1984). Each period, the match produces a random output with a constant mean. The better match quality, the higher is the expected value of output. The worker and the firm gradually learn about match quality by observing realizations of this random output. They separate whenever the match quality they infer from all past realizations of output is below a certain threshold. The (Bayesian) implication is that the better the worker and the firm are informed about match quality, the smaller will be the impact of new information. Beyond a certain point in time, new information has hardly any impact on beliefs and future separation becomes increasingly unlikely. The separation rate converges to zero rapidly. Unlike the search model, random shocks to output have a permanent effect on the probability of future separation, since beliefs about the expected value of output are a function of all past realizations.

The learning model generates a hump shaped rate. Directly after the start of an employment relation, the worker and firm have not yet collected sufficient information to form an accurate belief about match quality. In the second stage, beliefs become more accurate and bad matches are eliminated. Finally, beliefs are almost exact, but all bad matches have been eliminated previously, so that the separation rate drops to zero. The analyses of Lancaster, Imbens, and Dolton (1987) and Miller (1984) indicate that there is indeed a hump shape, but that the learning model predicts the separation rates to decline more quickly than is observed empirically.

The information assumptions in the random growth model considered in this paper are the mirror image of those in the learning model. The worker and the firm are perfectly informed about today's match productivity, but they do not know its future evolution. Future productivity is supposed to follow a geometric random walk. When the productivity of the match falls below a threshold, separation becomes the efficient alternative. Hence, shocks to productivity have a permanent effect on future separations: a downward shock today moves productivity in the direction of the separation threshold, increasing the probability of separation at some time in the future. The uncertainty about future productivity relates the model to the literature on firm's labor demand under uncertainty, see Bentolila and Bertola (1990). The stochastic process underlying the random growth model fits in the strand of learning models, in the sense that stochastic shocks have a cumulative effect on future separations.[1] However, unlike the learning model, the effect of new shocks does not decline in the course of time.

Like the learning model, this model yields a hump shaped separation rate. At first, the separation rate is low. Players would not have made the specific investment if match productivity were close to the separation threshold. In due time, some matches have accumulated negative shocks, leading to an increase in the separation rate. Eventually, separation rates decline again, since matches with accumulated negative shocks will have been eliminated previously. However, contrary to the learning model, the effect of new shocks does not diminish over time. Hence, the separation rate declines more slowly than in learning models. When the drift of the Brownian is negative, the separation rate does not even converge to zero.

Workers and firms are required to make specific investments at the start of the match. These investments can either be hiring cost, or firm-specific formal training programs, or more generally, the time that is needed to get acquainted to the type of work that the firm expects the worker to do. We do not take a stance on the weights of these components. These investments lose their value upon separation. This sets our model apart from the literature on temporary lay-offs, where workers expect to be rehired by their previous employer, see for example Feldstein (1976).

Our model is formally equivalent to that of Dixit (1989). Since specific investments are required, hiring and separation decisions are irreversible. Hence, these investments have an option value. We make three assumptions. First, we assume efficient bargaining on the distribution of surpluses from specific investment, so that separation decisions are always efficient. Second, we assume risk neutrality, so that the allocation of the uncertainty about the future evolution of productivity is irrelevant. These assumptions allow us to analyze separation decisions and wage setting separately. Finally, we assume holdup problems to be resolved efficiently, so that the surplus value of a job is equal to the cost of investment at job start.

For most of the paper, we benefit from the option to analyze wage and separation decisions separately, by ignoring wages. However, we do explore the implications for the tenure

[1] Keane and Wolpin (1997) offer a model with randomness in productivity after job start that is driven by transitory shocks.

profile. For this purpose, we consider the implications of a simple Nash bargaining rule. Tenure profiles in wages emerge naturally from our probability law for the path of productivity. Even when there is no deterministic trend in the within-job-productivity of the worker relative to her outside market option, the model generates a simple rationale for a tenure profile. Random walks, which develop unfavorable, are eliminated from the stock ongoing matches. The remaining stock is therefore a selective sample of random walks. This selection mechanism generates a tenure profile in wages, which is consistent with what is reported by for example Topel (1991).

We consider two versions of the model, a general version where both the outside option of the worker and the productivity in the firm follow a geometric Brownian and a more restricted version, where the outside option is deterministic. We consider the latter to be a limiting case, following the evidence of Davis and Haltiwanger (1992) that aggregate shocks to productivity (affecting the outside options) have a much smaller variance than firm specific shocks (affecting the productivity of a match). The restricted version of the model is applied to the tenure distribution of individual employment relations, using the NLSY. From this distribution, we can identify all structural parameters up to the variance of shocks in productivity. For the identification of this final parameter, we apply the first order condition for the optimal hours of on-the-job training. On-the–job training is just one of the components of the specific investments required at the start of a match. Workers and firms set their level of investment for each component as such that the marginal cost of this component is equal to the increase in the value of job. Since hours of training are observed and since their effect on the value of the job can be estimated, the first order condition for this component provides an additional constraint that can be applied for the identification of the final parameter.

The empirical analysis of the tenure distribution in Section 3 shows that the model describes the data well. The hump shape pattern predicted by both the learning and the random growth model is indeed observed in the data, but the random growth model does a much better job in explaining the separation rates at higher tenures. However, we would like to have more direct evidence on the crucial issue that separations are driven by accumulated past negative shocks to productivity, not by the most recent shock, as follows from search models. Here, we face the problem that we do not directly observe the surplus of match productivity above the separation threshold most of the time. We only know that productivity must be equal to this threshold at the date of separation. However, the value of productivity at the separation date is obviously a highly selective sample from all conceivable evolutions of productivity from the start of the job till the moment of separation. Hence, we cannot apply standard techniques to analyze these data.

In Section 4, we work out a simple methodology to deal with this problem. We apply this methodology to one particular type of shocks, the aggregate shocks that underlie the evolution of unemployment, using monthly state data. Obviously, the aggregate shocks reflected in the unemployment rate are only a small fraction of all shocks that drive the evolution of productivity in individual matches. Individual and firm specific shocks will probably by more important, see Davis and Haltiwanger (1992). The aggregate shocks reflected in unemployment can therefore offer only a partial explanation. We estimate a VAR(2) model to recover the innovations in these unemployment histories and use these innovations to construct a Brownian, which we then apply to analyze the evolution of productivity of individual matches. Our analysis shows a strong correlation between this constructed Brownian and the productivity as revealed at the moment of separations. This offers strong evidence in favor of the random growth model. It is not a transitory shock -the state of the business cycle at this point in time- that triggers separation. It is the accumulated history of shocks during the course of the match that determines whether or not its continuation is profitable.

In Section 5 we show how our model of individual employment relations is embedded in Bentolila and Bertola's (1990) model on the evolution of employment at the level of the firm. The hiring cost in their model can be identified as the specific investments in our model. Bentolila and Bertola assume an iso-elastic demand function evolving over time according to a geometric Brownian. With these assumptions, there is a one-to-one correspondence between our model of individual matches and the model of Bentolila and Bertola, if a Last-In-First-Out (LIFO) separation rule applies. Kuhn (1988) provides a rationale for the use of LIFO lay-off rules in the context of a unionized firm. We draw on this idea and argue that a LIFO lay-off rule serves to protect the property rights of the specific investments of incumbent workers against the claims of workers that are hired later on.

The analogy of our model of an individual match and a firm level model also provides a rationale for the assumption that productivity evolves according to a geometric Brownian. Bentolila and Bertola's model implies firm size to evolve (almost) according to Gibrat's law. The excess productivity of matches above their separation threshold is shown to correspond to the productivity of the intramarginal workers in the firm. With adequate data, we would be able to analyze the share of firm level shocks in the total variance of the shocks that drive the evolution of match productivity, using the methodology that has been set out in Section 4.

The set-up of the paper is as follows. The model is derived in Section 2. Section 3 discusses the estimation results for the tenure distribution. Section 4 sets out a method for analyzing the decomposition of shocks to match productivity into, for example, idiosyncratic and macro-level shocks. Estimation results on the impact of macro-level shocks are presented. In Section 5, we analyze the relation of our model with that of Bentolila and Bertola (1990). Section 6 concludes.

2 The Model

2.1 Assumptions

Consider an economy with risk neutral workers and firms. In the market, workers can collect their market value or outside option R_t per unit of time. This outside option is simply the maximum of all alternatives that are available. Firms own vacancies. A filled vacancy (a job) produces a particular type of output with market value P_t per unit of time. At a particular time t, a firm can decide to hire a worker to fill the vacancy and to start producing output. At the start of an employment relationship, relation specific investments $R_t I$ have to be paid; I can be seen as the physical amount of investment required and R_t as the price per unit of investment. The value of the specific investments is therefore proportional to the outside option of the worker. As far as these investments are made up of non-productive hours of the worker, R_t is obviously its adequate price. However, also in the case that these investments require other, non-labor inputs, R_t is reasonable proxy to their price, since R_t will be closely correlated to the general price level[2].

[2] Van der Ende (1997) allows for a mixture of investments at prices P_t and R_t.

Specific investments will be lost upon separation between the worker and the firm. However, the firm retains the property rights of the vacancy, that is, it holds the option to hire another worker at a later date, for example when the market value of output improves. In that case, specific investments have to be made again. All specific investments are made at the start of the employment relationship. This assumption is obviously restrictive but it is a reasonable first approximation. Both the worker and the firm are perfectly informed about the present value of R_t and P_t. However, the future evolution of both variables is subject to uncertainty. Both R_t and P_t follow a geometric Brownian.

Firms and workers are assumed to bargain efficiently. That is, as long as it is efficient to continue the relationship, they reach agreement on the distribution of the surplus. Furthermore, we assume that the firm and the worker are able to resolve potential hold-up problems and to share future surpluses according to their share in specific investments. Together with the risk neutrality, these assumptions imply that the pattern of job duration and specific investments can be analyzed separately from the distribution of the surplus arising from these investments between the worker and the firm. We shall discuss our model as if the firm pays for all investments and accordingly receives the full surplus of the relationship, while the worker gets her outside option. However, any other sharing rule is consistent with our results, as long as investment cost and surpluses are shared in the same way. Since separation decisions are made efficiently, there is no surplus left at that point in time and separation is therefore in the mutual interest of both players. Hence, it does not make sense to distinguish between quits and lay-offs, compare McLaughlin (1991). Clearly, the assumptions on risk neutrality, efficient bargaining and the resolution of hold up problems are unlikely to be met completely in practice. For example, the evidence presented Jacobson, LaLonde, and Sullivan (1993) supports the idea that there are at least some gains of trade left at the date of separation. However, we feel that it is better first to have an idea about the parameters that are consistent with first-best before entering the fog of a second-best real world.

As discussed in Section 1, we consider a general model, where both R_t and P_t follow a geometric Brownian, and a restricted model, where R_t is deterministic. The next subsection investigates the general model and derives its implications for the tenure distribution. The restricted model will be discussed in Section 2.5. There, we derive the relation between the structural parameters and the hiring and separation thresholds.

2.2 The General Model and its Implications for the Tenure Distribution

Using lower cases for logarithms, the law of motion for the market value of log output and the log outside option of the worker between arbitrary dates *s* and *t* which are subject to permanent, Normally distributed shocks is:

$$[p_t - p_s, r_t - r_s] \sim \mathrm{N}\,[(t - s)\,[\mu_p, \mu_r], (t - s)\,\Sigma] \tag{1}$$

The value of a vacancy, denoted $\underline{V}(P_t,R_t)$ and a filled job, $\underline{J}(P_t,R_t)$, the latter net of the outside option of the worker, both measured at date t are given by:

$$\begin{aligned} \underline{V}(P_t, R_t) &= \mathrm{E}_t\left[\mathrm{e}^{-\rho(X-t)}\left(\underline{J}(P_X, R_X) - R_X I\right)\right] \\ \underline{J}(P_t, R_t) &= \mathrm{E}_t\left[\mathrm{e}^{-\rho(T-t)}\ \underline{V}(P_T, R_T)\right] + \mathrm{E}_t\left[\int_t^T \mathrm{e}^{-\rho(s-t)}(P_s - R_s)ds\right] \end{aligned} \tag{2}$$

where ρ denotes the interest rate, T is the efficient separation date, and X is the efficient hiring date. Both T and X are random variables depending on the 'filter' P_t, R_t, $t \leq X,T$ and where $E_t[.]$ is the expectation operator for the filter P_s, R_s, $s > t$. The value of a vacancy is fully determined by the option of filling the vacancy by making the specific investment $R_X I$ at some unknown future date X. In return for this investment, the firm obtains the value of a filled job. In the second equation, the value of filled job is made up of two parts. The first part is the option to fire the worker at some unknown future date X. In that case, the firm holds the value of a vacancy. The second part is the expected value of the productivity of the worker, net of what she would earn on the outside market (her outside option R_t). At the moments of hiring and separation, when the firm switches back and forth between the value of a job and a vacancy, the firm is indifferent between the two alternatives. When switching from a vacancy to a filled job, we have to account for the cost of specific investments. Hence, letting $t=X$ and $t=T$ respectively in equation (2):

$$\underline{J}(P_X,R_X) = \underline{V}(P_X,R_X) + R_X I,$$
$$\underline{V}(P_T,R_T) = \underline{J}(P_T,R_T).$$

The stopping rules for hiring and separation depend only on P_t and R_t because the law of motion formula (1) implies the strong Markov property for $[P_t,R_t]$ and because the time horizon is infinite, see McDonald and Siegel (1986, page 712-713). Since the law of motion for $k \cdot [P_t,R_t]$ is equal to that of $[P_t,R_t]$, it can be shown that the stopping rules depend only on the ratio of the market value of output P_t to the outside option of the worker R_t at time t, defined as B_t. Its logarithm b_t is a Brownian with drift μ_p-μ_r and variance rate $\sigma^2 = \sigma^2_p + \sigma^2_r - 2\sigma_{pr}$. The firm hires a worker at the moment X when B_t hits an upper bound B_X and the worker separates from the firm at the moment T when this ratio hits a lower bound B_T. Heuristically, R_t can be divided out of both equations in formula (2) so that:

$$\underline{V}(P_t,R_t) = R_t\underline{V}(B_t,1);$$
$$\underline{J}(P_t,R_t) = R_t\underline{J}(B_t,1).$$

An employment relation ends when the worker and the firm no longer consider its continuation beneficial. This happens at the first time after the start of the job that b_t is at the separation level b_T. Since the employment relation has started at time t such that b_t is at the hiring level b_X, the duration until separation is determined by the time that is required for the random walk to travel down the distance b_X-b_T. Whether this distance is large or small depends on the standard deviation of shocks to b_t per unit of time, σ. The probability of this distance being traveled remains unaffected by a proportional variation of this distance, the drift and the standard deviation. Distance and drift can therefore be normalized by the standard deviation. Define: $\Delta_t \equiv (b_t - b_T)/\sigma$, $\Delta \equiv (b_X - b_T)/\sigma$, and $\pi \equiv (\mu_p - \mu_r)/\sigma$. Δ_t is the normalized distance between the actual log productivity and the separation threshold, Δ is the normalized distance between the hiring and separation threshold, π is the normalized drift. At the moment of hiring $\Delta_X = \Delta$, while at the moment of separation $\Delta_T = 0$. Hence, the distribution of job tenures is fully determined by two parameters, Δ and π. Note that Δ is not a structural parameter. It depends on the optimal hiring and separation thresholds b_X and b_T, which will be determined in Section 2.5.

The distribution of Δ_t - Δ conditional on the hiring time X is $N[(t-X)\pi,(t-X)\sigma]$. When $\Delta_t < 0$, separation has occurred at some time $X < T < t$. However, not all realizations of $\Delta_t > 0$ correspond to an ongoing employment relation. It might be the case that $\Delta_s < 0$ for some s, $X < s$

$< t$, but that Δ_t traveled back to a positive value since then. However, since the separation decision is irreversible, these realizations do not correspond to ongoing employment relations. The probability that no endogenous separation occurs before time t is the probability that $\Delta_s > 0$ for all $X \leq s < t$. This conditional density can be calculated from the reflection principle. This principle is illustrated in Figure 1. There is a one-to-one correspondence between the trajectories starting at Δ and ending at Δ_t but having crossed the line $\Delta_s = 0$ at least once on the one hand, and the trajectories starting from $-\Delta$ and ending in Δ_t on the other hand. Hence, these trajectories should be subtracted when calculating the density of all trajectories that never crossed the line $\Delta_s = 0$. Hence, the density of Δ_t and the relationship still going on (that is: $T > t$) conditional on the starting time of the job, X=0, and the moment of observation t is:

$$\Pr[\Delta_t, T > t \mid t, X = 0] = \frac{1}{\sqrt{t}}\left[\phi\left(\frac{\Delta_t - \Delta - \pi t}{\sqrt{t}}\right) - \Theta\,\phi\left(\frac{\Delta_t + \Delta - \pi t}{\sqrt{t}}\right)\right] \tag{3}$$

where $\Theta = e^{-2\Delta\pi}$. The parameter Θ accounts for the effect of the drift π. It cancels when $\pi = 0$.[3] The distribution function of completed job tenures follows from integrating out Δ_t:

$$\Pr[T > t \mid t, X = 0] \equiv 1 - F(t) = \Phi_t^+ - \Theta\Phi_t^-$$

where $\Phi_t^i = \Phi(x_t^i)$, $x_t^+ = (\Delta + \pi t)/\sqrt{t}$ and $x_t^- = (-\Delta + \pi t)/\sqrt{t}$. Other authors have applied this statistical model for the description of duration data before. The first paper that we are aware of is that by Lancaster (1972), who applies the model to the duration of strikes, with considerable success.

The economic model discussed above is not the only model that yields this distribution for completed job tenures characterized by Δ and π. For example, a model where productivity follows a Brownian instead of a geometric Brownian, leads to exactly the same statistical model for completed job tenures. Hence, the fact that observed tenure distribution matches the predicted distribution closely does not necessarily imply that the random growth model is the only model that can explain the data.

The exit rate from employment is given by $\lambda(t) \equiv f(t)/(1-F(t))$. The pattern of this exit rate has the following characteristics[4]:

i) $\lambda(0) = 0$ and increases from then on;
ii) $\lambda(t)$ reaches a peak at t_0, where $0 < t_0 < {}^2/_3\Delta^2$;
iii) after t_0, $\lambda(t)$ declines monotonically to:
for $\pi > 0$: $\lambda(\infty) = 0$;
for $\pi < 0$: $\lambda(\infty) = -\frac{1}{2}\pi^2$.

Stated roughly, Δ locates the peak of $\lambda(t)$ in time and π determines its final level. There is a clear intuition for this pattern. The firm only hires a worker when the productivity of the job is way above the outside option of the worker, since it will have to pay this outside option to

[3] The intuition for Θ is that $\Pr[\Delta_s=0, \Delta_t|\Delta_0=\Delta, s,t,s<t] = \Theta \Pr[\Delta_s=0, \Delta_t|\Delta_0= -\Delta, s,t,s<t]$. Note that the factor Θ is independent of s.

[4] The proposition of the peak follows from inserting $\lambda' = 0$ into $\lambda'' < 0$.

the worker. Hence, initially, the chance that this surplus is dissipated by random shocks is negligible. After some time, a sufficient number of shocks have been accumulated, pushing up the separation rate. Later on, the separation rate declines by a selection mechanism. Trajectories of the Brownian that started with a large number of negative shocks have been eliminated by previous separation, so the probability mass of remaining jobs shifts upward. When the drift is positive, the drift (increasing linearly with time) will dominate the random shocks (depending with the standard deviation of the Brownian, which increases with the square root of time) in the long run. When the drift is negative, there will be a constant force pressing the surviving jobs towards the separation threshold. The hump-shaped pattern is indeed a feature of empirically observed job-exit rates, see Farber (1994). We shall return to the issue of the shape of the hazard rate when discussing the estimation results.

2.3 A Comparison with the Learning Model

Learning models (Jovanovic (1979a), Lancaster, Imbens, and Dolton (1987) and Miller (1984)) yield the same hump shaped pattern in separation rates as the random growth model. However, the random growth model generate much higher separation rates at longer tenures. For the sake of comparison, we offer a short discussion of a simplified version of the learning model, using a notation that highlights its similarity with the random growth model. Match productivity x_0 is a match specific constant, which is however unknown to the firm and the worker. The distribution of x_0 across jobs is normal with mean b_0 and variance σ_0^2. Actual output in period t, x_t is equal to x_0 plus normally distributed white noise ε_t with variance σ^2. Unlike the random growth model, the model is in discrete time, but we can approximate a continuous time model arbitrarily close by choosing an ever-shorter unit of time and decreasing the value of σ^2 proportionally. The firm and the worker have to infer match productivity x_0 from realizations of output x_t. They form beliefs b_t about match productivity by Bayes' rule:

$$b_t = \frac{1}{1+\delta t}\left[\sum_{s=1}^{t} \delta x_s + b_0\right]$$

where $\delta \equiv \sigma_0^2/\sigma^2$. Beliefs are a weighted average of expected match productivity and past realizations of output. In this simplified version, separation occurs the first time that the belief b_t is below a certain threshold, b_T. In the full model, the separation threshold is time dependent since the variance of the beliefs b_t decreases over time. The smaller the variance, the lower the option value of continuing the employment relation. Hence, the separation threshold increases over time. However, this effect is strongly non-linear and cannot be characterized analytically, see Lancaster, Imbens, and Dolton (1987).[5] We shall therefore ignore this effect in the subsequent analysis. The simplified version of the learning model can be respecified in terms of two parameters, $\Delta \equiv (b_0 - b_T)/\sigma_0$ and δ. The first is an analogue of the parameter Δ in the random growth model. By ignoring the upward trend in the separation threshold we have implicitly set the drift equal to zero. Like in the random growth model, separation occurs whenever standardized beliefs $\Delta_t \equiv (b_t - b_0)/\sigma_0$ have traveled down the distance Δ.

[5] Hence, contrary to the optimal separation rule in the random growth model, the optimal separation rule in the learning model does not have a complete analytical characterization. This complicates a precise comparison between both models.

The crucial difference between the random growth model and the learning model comes to surface when comparing the evolution of the variance of Δ in both models. For the learning model, we have:

$$Var[\Delta_t] = \frac{\delta t}{1+\delta t}$$

while the random growth model yields $Var[\Delta_t] = t$. Hence, the accumulation of shocks leading to the gradual evolution of Δ_t proceeds in a time scale in the learning model that is transformed compared to the time scale that applies in the random growth model. Eventually, for $t \to \infty$, the variance of beliefs in the learning model converge to: $Var[b_t - b_0] = Var[x_0 - b_0] = \sigma_0^2$, since beliefs converge to actual match productivity x_0.

When we set: $\Delta_{learning} = \sqrt{\delta}\,\Delta_{random\ growth}$, the accumulation of shocks relative to the size of the initial surplus, starts at the same rate per unit of time in both models. Hence, the hazard rates at $t \to 0$ are equal in both models. After a while, the accumulation of shocks in the learning model starts lagging behind that in the random growth model, and so does the hazard rate. The variance of accumulated shocks in the learning model will never get above the value that is achieved at $t = \delta^{-1}$ in the random growth model. The lower δ, the longer it takes before the hazard rate in the learning model starts lagging behind that of the zero-drift random growth model. The pattern of separation rates of a zero-drift random growth model is therefore a special case of the pattern generated by the simplified version of the learning model for $\delta = 0$.

This previous thought experiment aims at setting equal the hazard rates of both models at $t \to 0$. One can also choose to set equal the hazard rate at later points in time. In each case, the hazard rate of the learning model will decline relative to that of the random growth model from that time onwards. This is illustrated in Figure 2. The thin line represents the separation rate for a random growth model with $\Delta_{random\ growth} = 4$, while the fat line represents the separation rate for a learning model with $\Delta_{learning}/\sqrt{\delta} = 3.8$ and $\delta = 0.01$, implying that the variance of transitory shocks in productivity per unit of time is equal to 100 times the variance of x_0 around b_0. The parameters of the learning model are set as such that the location of the peak is similar in both models. At that point in time both models have produced about the same number of separations. Even for this low value of δ, the random growth model produces a substantially higher number of separations than the learning model for $t > 30$. For higher values of δ the difference between both models is even more pronounced.

Where the pattern of separation rates of a zero-drift random growth model is a special case of the simplified version of the learning model, this correspondence does not extend to the random growth model with a negative drift. This follows from a simple argument. In the random growth model, a negative drift implies that eventually all matches will be broken up, since $\lambda(\infty) = -½\,\pi^2 > 0$. In the learning model, there is always some fraction of the matches that survive forever, since learning the match quality x_0 makes sense only if there is some fraction for which productivity is above the separation threshold. These matches survive forever, or more precisely, end only for reasons that are exogenous to the model. Note that this argument does not rely on our simplification of the learning model, by disregarding the non-linear negative drift. Hence, the pattern of separation rates of a random growth model with negative drift can never be generated by a pure learning model.

2.4 Tenure Profiles in Wages

Simple cross section wage regressions tend to show substantial returns to tenure. There is an extensive literature on the measurement of these tenure profiles, see for example Abraham and Farber (1987), Altonji and Shakotko (1987) and Topel (1991). This literature takes into account all kind of biases introduced by the self-selection of workers into particular types of jobs. The question asked by this literature can be summarized as: do high wages cause long tenures, or is it the other way around? In this subsection, we address the potential implications of the random growth model for observed tenure profiles. For this purpose, we extend the model with a simple sharing rule for the distribution between the worker and the firm of surpluses from specific investments. Let w_t be the log wage of the worker. Our sharing rule simply distributes instantaneous surpluses proportional to the worker and the firm[6]:

$$w_t = r_t + \ln\{ 1 + \beta[\exp(b_t) - 1] \} \cong r_t + \beta\, b_t = r_t + \beta\sigma\Delta_t \tag{4}$$

where β, $0 < \beta < 1$, is the worker's share in specific investments. This sharing rule is consistent with the previous assumption that hold up problems are resolved efficiently if workers' share in the specific investments is equal to β. By letting the worker and the firm share the instantaneous surplus in this way, the random growth model generates a natural explanation for the positive correlation of job-tenure and wages that is observed in cross section data, even when the normalized drift π is negative. To see this, consider the expected value of Δ_t for ongoing employment relations. This expectation can be calculated from the density of Δ_t conditional on survival at $X < t < T$ and the date of job start X, see equation (3):

$$\mathrm{E}[\Delta_t \mid t, X < t < T, X = 0] = \sqrt{t}\,\frac{\varphi_t^+ - \Theta\varphi_t^- + \Phi_t^+ x_t^+ - \Theta\Phi_t^- x_t^-}{\Phi_t^+ - \Theta\Phi_t^-} = \pi t + \Delta\frac{\Phi_t^+ + \Theta\Phi_t^-}{\Phi_t^+ - \Theta\Phi_t^-} \tag{5}$$

For the second equality, we use $\varphi_t^+ = \Theta\,\varphi_t^-$. The slope of the tenure profile in a cross-section regression on log wages is equal to $\sigma\beta \times$ the derivative of this expectation with respect to t. The latter reads:

$$\frac{\mathrm{d}\,\mathrm{E}[\Delta_t \mid t, X < t < T, X = 0]}{\mathrm{d}t} = \pi + \frac{\Delta\phi_t^+}{t\sqrt{t}\left[\Phi_t^+ - \Theta\Phi_t^-\right]}\left[\pi t + \Delta\frac{\Phi_t^+ + \Theta\Phi_t^-}{\Phi_t^+ - \Theta\Phi_t^-}\right]$$

For a positive drift, this derivative is always positive, leading to a tenure profile. The interesting case is that of a negative drift, $\pi < 0$, which we discuss below. The derivative consists of two terms. The first term measures the direct effect of the drift, which is negative. The second term measures the effect of the elimination of unfavorable trajectories of the Brownian by separations prior to time t. In the short run, the first term dominates, because there is not yet much selection going on, see the discussion on the initial value of $\lambda(t)$. For $t = 0$, the second term even vanishes

[6] In using this sharing rule we apply a pragmatic approach compared to what is most common in the literature, where wage setting is done not by sharing the instantaneous surplus but by sharing the return on the expected discounted value of future surpluses, yielding: $w_t = r_t + \ln\{ 1 + \rho\beta(V_t - J_t) \}$.
The second approximation in equation (4) follows from a first order Taylor expansion that applies for small values of p_t. For larger values of p_t, w_t converges to $p_t + \ln\beta$. The advantage of this first order expansion of a sharing rule based on the instantenaous surplus is its linearity in Δ_t.

due to the factor ϕ_t^+. In the long run, the selection effect and the effect of the drift cancel, as follows from taking limit for $t \to \infty$ of equation (5)[7]:

$$\lim_{t\to\infty} \mathrm{E}\left[\Delta_t \mid t, X < t < T, X = 0\right] = -\frac{2}{\pi}$$

Since the conditional expectation of Δ_t is equal to Δ for $t = 0$ and is equal to $-2/\pi$ for $t \to \infty$, its slope has to be positive in the intermediate run if: $-2/\pi > \Delta$. In that case, the selection effect dominates the drift, so that observed wages exhibit a tenure profile even when there is no inherent job specific productivity gain. We shall apply these formulas when discussing the implications of our estimation results.

The previous analysis is useful for cross section data, where we observe the starting date X but not the stopping time T. However, in panel data, we also observe T for completed spells. The random growth model with Nash bargaining implies that w_t-r_t follows a Brownian with drift. Topel (1991), Topel and Ward (1992) and Dustman and Meghir (2001) find indeed strong evidence that log wages within a job follow a Brownian, although their evidence regards w_t, and not w_t-r_t. Then, a simple cross-section regression would overestimate the drift, since log wages in surviving jobs are a selective sample of jobs, where wages are above workers' separation threshold. The eventual tenure captures the information that at the moment of separation, the wage rate is equal to workers' separation threshold, while it is above this threshold before the moment of separation. Hence, the model implies that the growth in w_t-r_t should be lower just before separation. This prediction gets support in Topel and Ward (1992, Table VI, model (v)), but not in Topel (1991, Table 4). However, their evidence regards w_t, and not w_t-r_t. This difference matters in particular when worker's human capital is not fully job specific so that the productivity in the job will be correlated positively to the value of the outside option. Jacobson, LaLonde, and Sullivan (1993) find strong support for a declining profile in w_t-r_t in the period before separation, though their evidence suggests that there are still substantial gains from trade at the moment of separation. However, their evidence refers to lay-offs only, not to quits. If our model is correct, there is no such thing as 'the' earnings loss or 'the' tenure profile in wages. Tenure profiles depend on the evolution of a match.[8]

2.5 The Restricted Version of the Model and the Level of Specific Investments

In the restricted version of the model, the outside option is deterministic and hence the covariance matrix Σ has only a single non-zero element, which is equal to σ^2 by previous definitions. Since the efficient hiring and separation dates are given in terms of B_t, it is convenient to normalize the value functions by dividing through R_t and using B_t as the only argument. By applying Ito's lemma, these value functions are defined by two Bellman equations, see Dixit (1989):

$$\begin{aligned}\rho_r V(B_t) &= (\sigma\pi + \tfrac{1}{2}\sigma^2)V'(B_t)B_t + \tfrac{1}{2}\sigma^2 V''(B_t)B_t^2 \\ \rho_r J(B_t) &= (\sigma\pi + \tfrac{1}{2}\sigma^2)J'(B_t)B_t + \tfrac{1}{2}\sigma^2 J''(B_t)B_t^2 + B_t - 1\end{aligned} \tag{6}$$

[7] We use: $\lim_{x\to-\infty} \Phi(x) = -(1-x^{-2})\, x^{-1}\, \phi(x) + O(x^{-4})$ and: $\varphi_t^+ = \Theta\, \varphi_t^-$

[8] This observation puts into question much of the literature on the estimation of the tenure profile, see e.g. Altonji and Shakotko (1987), which tries to estimate a tenure profile that is independent of the future perspectives of the job. Our model also offers a natural explanation why the variance of the error term in a wage equation increases with tenure.

where $R_tV(B_t) \equiv \underline{V}(B_t,1)$, $R_tJ(B_t) \equiv \underline{J}(B_t,1)$, and $\rho_r \equiv \rho - \mu_r$; ρ_r can be interpreted as a modified discount rate, accounting for the drift in the outside option μ_r. The first term on right hand side of both equations takes care of the drift in B_t. The second term accounts for the non-vanishing second order effect of shocks to B_t. The final terms in the second equation measure the current output of a filled job, net of the outside option of the worker, which equals unity due to the normalization by R_t. The hiring threshold B_X and the normalized distance between the hiring and firing threshold Δ are implicitly defined by the following relations, see the Appendix for their derivation:

$$\begin{aligned} I &= \frac{1}{\sigma\rho_r}\frac{\sigma\alpha_1\left(1-D^{\sigma-\alpha_1}\right)\left(D^{\alpha_1}-D^{\alpha_1-\alpha_2}\right)+\sigma\alpha_2\left(D^{\sigma-\alpha_2}-1\right)\left(D^{\alpha_1}-1\right)+2\rho_r\left(D^{\sigma}-1\right)\left(D^{\alpha_1-\alpha_2}-1\right)}{\alpha_1\left(1-D^{\sigma-\alpha_1}\right)D^{\alpha_1-\alpha_2}+\alpha_2\left(D^{\sigma-\alpha_2}-1\right)+\frac{2\rho_r}{\sigma}\left(D^{\alpha_1-\alpha_2}-1\right)} \\ B_X &= \frac{1}{\sigma}\frac{2\rho_p\left(D^{\alpha_1-\alpha_2}-1\right)D^{1/\sigma}}{\alpha_1\left(1-D^{\sigma-\alpha_1}\right)D^{\alpha_1-\alpha_2}+\alpha_2\left(D^{\sigma-\alpha_2}-1\right)-\frac{2\rho_r}{\sigma}\left(D^{\alpha_1-\alpha_2}-1\right)} \end{aligned} \qquad (7)$$

where $D \equiv e^{\Delta}$, $\alpha_1 \equiv -\pi + \sqrt{(\pi^2+2\rho_r)}$, and $\alpha_2 \equiv -\pi - \sqrt{(\pi^2+2\rho_r)}$. The value of a job is finite if and only if $\rho_p \equiv \rho_r - \pi\,\sigma - ½\,\sigma^2 > 0$ and $\rho_r > 0$, which we assume in the sequel; ρ_p can be interpreted as a modified discount rate for the flow of payoffs from the match, accounting for the deterministic and stochastic drift in p_t. Equation (7) implies that b_X is positive and b_T is negative. Immediately after the hiring decision, the productivity in the job P_X exceeds the outside option R_X by at least the interest payments on the specific investment, $\rho\, R_X\, I$. If not, the firm would be able to increase its profits by postponing the investment. At the moment of separation, P_T is below R_T. When the productivity is just slightly below the outside option, it is better to retain the worker since separation decisions are irreversible. The worker and the firm can only benefit from the opportunity that productivity might pick up later on as long as they have not separated, for otherwise they would have to incur the cost of specific investment again.

The drift parameters μ_r and μ_p, and the discount rate ρ do not appear independently of π and ρ_r in equation (7). We can therefore respecify the model in terms of the latter two parameters and eliminate μ_r, μ_p, and ρ. With this reparametrization, the restricted model has four parameters: π, σ, ρ_r, and I. We assume exogenous information on the modified discount rate ρ_r to be available; 10 % per year seems to be a reasonable value. The distribution of job tenures is determined by two (composite) parameters, see Section 2.2: π and Δ. These parameters can therefore be estimated from tenure data. The first of the pair of equations (7) provides an implicit relationship between the composite parameter Δ and the underlying structural parameters, in particular the level of investment, I. Hence, till so far, the model is identified up to a single parameter, the standard deviation of shocks per unit of time, σ. This final parameter will be identified from the contribution of the marginal hour of investment in job-specific training, H, to the total of specific investments, I. Time spent on specific training is only one of the components of job specific investments. There are other components, for example the time of experienced workers spent on the training of their colleagues, and hiring cost. All these components contribute to total investment, $R_X\, I$. However, the interesting feature of the time spent on training is that our model generates its marginal price. The cost of a marginal time unit of training is equal to P_X. This cost exceeds the opportunity cost of untrained workers ($R_X<P_X$), since intramarginal specific investments have already been made, which raise the productivity of the worker above her outside market option. Workers and/or firms will set marginal cost equal to the marginal revenues of specific training. Since I is equal to the expected discounted value of a job, we have: $R_X\, dI/dH = P_X$. Variation in the actual amount of on job specific training can be related to the

observed tenure distribution, which allows us to estimate dΔ/dH. I is a (non-linear) function of Δ, denoted $I(\Delta)$, see equation (7). Hence:

$$I'(\Delta)\, d\Delta/dH = B_X \qquad (8)$$

Expressions for B_X and $I'(\Delta)$ can be obtained from equation (7). Given the availability of estimates of dΔ/dH, this equation can be solved numerically for its only unknown parameter, σ.

3 The Empirical Implementation

3.1 Specification and Likelihood

A full structural estimation of the model would require us to specify the likelihood directly in terms of the three structural parameters to be estimated, π, σ, and I. We pursue a simpler, semi-structural approach, where we specify the likelihood in terms of (composite) parameters Δ, π, and dΔ/dH, and then use the equations (7) and (8) to recover the structural parameters. We apply the following specification for Δ and π:

$$\Delta_{ij} = x_{ij}\,\beta_{\Delta} + u_{\Delta i}$$
$$\pi_{ij} = x_{ij}\,\beta_{\pi} + u_{\pi i}$$

where x_{ij} denotes the characteristics of job j for worker i. Since our estimation is based on a panel for workers, random worker effects are included in both Δ_{ij} and π_{ij}. We refrain from including random job effects, because we observe each job only once, so its identification would rely strongly on the functional forms.[9] Since $\Delta > 0$ for all jobs, we impose the constraint $u_{\Delta i} > -\delta_i$ where $\delta_i \equiv \min_j[x_{ij}\beta_{\Delta}]$.[10] We assume that $u_{\Delta i}$ and $u_{\pi i}$ are independent and normally distributed. Hence, the log likelihood reads:

$$\sum_i \ln \int_{-\delta_i}^{\infty}\int_{-\infty}^{\infty} \prod_{j=1}^{j_i} [1 - F(t_{ij})]^{1-d_{ij}} \left[f\,(t_{ij})^{d_{ij}}\, d\,\Phi\left(\frac{u_{\pi i}}{\sigma_\pi}\right) d\,\Phi\left(\frac{u_{\Delta i}}{\sigma_\Delta}\right)\right] - \ln\Phi\left(\frac{\delta_i}{\sigma_\Delta}\right) \qquad (9)$$

where $d_{ij} = 1$ if a job is uncensored and $d_{ij} = 0$ otherwise and where j_1 is the number of jobs of individual i. We programmed both first and second derivatives of the likelihood. Convergence was quick, requiring only a few iterations. The same estimates were achieved for different starting values.

3.2 The Data

The data were taken from the National Longitudinal Surveys of Youth (NLSY), provided by the U.S. Center of Human Resource Research. We apply 14 waves in the period 1979-1992. All respondents were interviewed in 1979 and were then aged 14 through 22. We selected full-time jobs of white males since the start of the career. We discarded jobs with missing occupations. The career of the respondent is said to have started at the beginning of the first paid full-time job

---

[9] See Van der Ende (1997) for estimation results including a random job effect.

[10] The alternative would be to use an exponential specification for Δ: $\Delta = \exp.(x\beta_\Delta + u_\delta)$ However, the disadvantage of this specification is that the additive structure $\Delta_t = x\beta_\Delta + t\, x\beta_\pi$ is lost.

with a known occupation (occupation is missing for jobs lasting less than 8 weeks that were uncensored at the first interview in which the job is recorded). Furthermore, the respondent must have been working for at least 22 weeks and at least 440 hours in the next three consecutive years. A paid full-time job is any job for which at least 30 hours per week is recorded at a positive wage rate. We end up with a dataset of 8,339 jobs held by 2,352 workers.

The data are summarized in Table 1. Tenure increases with calendar time as long as the respondent reports himself associated with the same employer. When an individual is rehired by the previous employer without himself reporting still associated to this employer in the intermediate period, we reset tenure to zero.[11] The NLSY has the advantage of low attrition and accurately measured tenures, in weeks. This is a crucial feature since aggregation over time tends to hide the hump shape in the hazard rate, which plays a central role in the model. We shall use a week as the unit of time when reporting our empirical results. All explanatory variables are measured at the start of a job. We shall use deviations from their means over the selected jobs. Experience is the sum of working and non-working experience since the start of a career.[12] It is defined as the calendar time since the start of the career, regardless of the employment status. Prior unemployment includes spells in which the respondent held only part-time or military jobs.

On-the-job-training is defined as the first non-governmental program that is attended parallel to a job. Later on-the-job training programs do not fit the assumptions of the random growth model, where all specific investments are decided upon at the start of the job. Happily, these programs are infrequent and little harm will be done in excluding them from the analysis. The total hours of a program are the hours per week from the first record of the program, times the duration of the program in weeks.[13] The last column gives the number of programs that survive the associated jobs by at least one and a half weeks. This would contradict our interpretation of these training programs as being job specific; 93 % of the 1,089 training programs are completed before the end of the job for which they are started. We censor surviving training programs at the recorded end date of a job. Thirteen percent of the jobs take a training program. Table 2 presents some statistics.

3.3 Estimation Results

Table 3 gives the Maximum Likelihood estimates of (9). We included dummies for 30 occupations (see Van der Ende (1997) for their classification) in the equation for Δ. Recall from Section 2 that Δ situates the peak of the job exit rates, and that in the limit exit rates depend on the normalized drift π only. Because highly specialized jobs, like that of a lawyer, typically start at the end of our young-worker survey, their effects on the drift π are weakly identified. Hence, we exclude the occupational effects of the drift. The occupation with the highest value of Δ is medical specialists, and the lowest is that of agriculture workers, which squares with our

---

[11] As pointed out by a referee, this causes some problems for the interpretation of the results in the case of rehiring since part of the specific investment do not have to be re-incurred in that case, compare Feldstein's (1976) discussion on temporary lay-offs. Hence our estimate of *I* is some mixture of the cost in the case of permanent separation and the much lower cost in the case of temporary lay-off.

[12] This excludes most of the holiday jobs.

[13] A complication is that the question regarding training programs in the NLSY has been changed half way the period of observation, see Parent (1999, 301). Before 1988, the question asked for training programs beyond military and government sponsored programs lasting longer than one month. This restriction to programs longer than a month was lifted afterwards. Like Parent (1999), we are therefore mixing data with and without a minimum duration requirement.

intuition. For the drift, the standard deviation of the random worker effect is large relative to the systematic component. For the distance, the standard deviation relative to the intercept is much smaller. This implies that truncation implied by the minimum condition to avoid negative distances has a limited impact on the estimation results, see Section 3.1.

Since all explanatory variables are measured in deviation of their mean, the intercept can be interpreted as an 'average' value for Δ. Loosely speaking, the estimated value of 6.6 for Δ implies that the initial surplus job productivity over the outside option is equal to 6.6 times the standard deviation of a weekly shock, or equivalently, 0.9 times the standard deviation of a yearly shock (since $\sqrt{52} = 7.2$). The peak in the hazard rate is somewhere between the job start and 29 weeks. The negative intercept of the drift provides strong evidence against a pure learning model, see the argument in Section 2.3. The size of the drift, -0.036, implies that, abstracting from the effect of random shocks, the initial surplus will dissipate in 6.6/0.036 = 183 weeks = 3.5 years. The models of Aghion and Howitt (1992) or Caballero and Hammour (1994) allow an interesting interpretation of this negative drift. New technologies are embodied in the specific investments required for new jobs. The application of the latest technology therefore requires a switch to a newly created job. The market alternative consists of with new jobs equipped with the newest technology while existing jobs have a zero drift. It is tempting to interpret the effect of living in central city as evidence in favor of this explanation. Big cities, with their large and therefore highly specialized labor markets, probably allow a faster diffusion of new technologies than small communities.

Figure 3 and 4 plot the observed (solid lines) and predicted (dashed lines) sample job exit rates for the first three jobs of each worker (two thirds of all jobs in the sample) for the first year and for the whole 14 year period. The plots indicate a good overall prediction of the random growth model. Taking into account that (conditional on the explanatory variables) only two parameters are used to fit the distribution, this is strong evidence in favor of the random growth model. As discussed in Section 2.3, the pattern of exit rates of a zero-drift random growth model can be generated by a pure learning model for extremely low values of δ only. The learning model can never match the hazard rates of a random growth model with negative drift. Our finding of a negative drift is therefore evidence against pure learning model. However, the random growth model somewhat underpredicts the peak, as can be seen most clearly from figure 3. Hence, a mix of both models might yield an even better description of the data.

Table 4 presents the values for the level of specific investment I (in weeks) and the hiring and firing threshold B_X and B_T, calculated from equation (7) and (8). We use the estimated intercept values for Δ and π, a value for ρ_r of 10 % per year, and a range of values of σ, from 0.005 to 0.075 per week, or equivalently, 4-60 % per year. For this sample of young workers at the beginning of their career, the value of specific investments ranges from the wage equivalent of a couple of days to four weeks of work, depending on the value of σ. The hiring threshold exceeds the firing threshold by a range varying from 3 % to 100 %.

The value of σ can established by the methodology outlined in Section 2.5, using the estimated dummies for the training programs as a proxy for $d\Delta/dH$. The dummies for the training programs are all positive and significant, suggesting that all categories of programs are at least partly job specific. A similar result has been reported by Parent (1999). However, the identification strategy outlined in Subsection 2.5 requires that a program is fully specific, since only then we can calculate the value of the additional specific investment that is embodied in the training program and relate that to the estimated effect on Δ. Parent (1999: 302) argues that seminars are the most fruitful candidate for this purpose. The estimations results suggest this intuition to be correct, since seminars have the lowest hours per course and nevertheless the largest effect on Δ. We focus therefore on this category in the subsequent argument.

There is one important caveat. The dummy for seminars might be a proxy for job heterogeneity. Though the inclusion of 30 occupational dummies offers a partial remedy for this problem, some unobserved job heterogeneity will persist. The standard deviation of the effect of unobserved worker characteristics on Δ shows that the impact of unobserved heterogeneity might well be substantial: the difference between the occupation with the lowest and the highest Δ is about twice the standard deviation of these unobserved characteristics. We do not have a proper instrument to account for this endogeneity bias. The estimated coefficient is therefore likely to be an upperbound of $d\Delta/dH$, since it picks up some unobserved heterogeneity. Using the median hours for a seminar, 24 hours or 0.6 week, an estimated upper bound for $d\Delta/dH$ is 6.19/0.6 = 10.3. Hence, an estimated lower bound for σ is 0.005 per week or 4 % per year. Topel and Ward (1992, Table VI) estimate the yearly standard deviation of innovations in w_t to be 13 % per year or 0.018 per week. When workers' share in the surplus β is equal to 0.30 (see e.g. Holmlund and Zetterberg (1991), Abowd and Lemieux (1993)), this yields a standard deviation for p_t of 0.018/0.30 = 0.06 per week. This estimate can be considered to be an upperbound for σ, since p_t is correlated positively with r_t when worker's human capital is not fully job specific. Hence: $0.005 < \sigma < 0.06$.

Table 5 presents some calculations of the tenure profile in wages unconditional on the separation date of the employment relation. We applied the benchmark parameter estimates for Δ, π, and ρ_r, and we set σ at its upperbound, to calculate the expected surplus $\sigma\Delta_t$ conditional on $T>t$ and the implied tenure profile $\beta\sigma(\Delta_t-\Delta)$, using equation (8). The calculations show that there is a rapid increase in the first 5 years. Later on, the profile flattens. We get close to the tenure profile that are obtained from simple OLS regressions on cross-section CPS data (numbers taken from Teulings and Hartog (1998: 37): 12 percent after 4 years, 18 percent after 8 years. It is encouraging that this simple structure can explain the tenure profile in wages.

4 Relating the Evolution of Match Productivity to Observed Shocks

4.1 Methodology

The previous analysis yields two empirical regularities contributing to the credibility of the random growth model: the humped-shaped pattern in separation rates and (compared to the predictions of the learning model) the high separation rates beyond the hump. However, both pieces of evidence offer only indirect support. It is always possible to specify some strange pattern of duration dependence to make any model fit the observed pattern of exit rates. It would be more convincing if we were able to show that shocks to the productivity of individual employment relations are correlated with observed shocks. We can think of many variables that can be expected to be related to the evolution of p_t or r_t, like for example, changes in log output prices or log employment either at the industry or the firm level. We can then test directly the main feature of the random growth model, that separations are driven by the accumulated random shocks to productivity running from the start of the job till the moment of separation, and not by the latest shock, as is suggested by search models. In this section, we work out a methodology for this type of analysis.

The main problem is that we do not observe the excess log productivity of a worker in his present job above her outside option, $b_t = \sigma\Delta_t$. The only information that is available is that Δ_t must be equal to zero at the moment of separation, $t = T$. Given the initial value Δ at the start of the job, we can infer that the random shocks must have accumulated to $-\Delta$ in the time period spanned by the job tenure. We have to rely on this cumulative information. However, the problem is that the realizations of Δ_t at the moment of separation are obviously a highly selective sample of all conceivable of trajectories of the underlying Brownian motion.

In order to deal with this selectivity problem, we consider the case of a single observed shock variable. Suppose that the Brownians r_t-r_s and p_t-p_s can be decomposed in an observed Brownian component a_{st} and two unobserved Brownian components e^r_{st} and e^p_{st} in the following way:

$r_t-r_s = \gamma_r a_{st} + e^r_{st},$
$p_t-p_s = \gamma_p a_{st} + e^p_{st},$

where $a_{st} \sim N[0,\sigma_a^2(t-s)]$ and where the unobserved components e^r_{st} and e^p_{st} are uncorrelated to a_{st}. The assumption of the drift of a_{st} being equal to zero is made for the sake of convenience and does not imply a loss of generality, since the observed variable a_{st} can be easily detrended. From the above assumptions we have:

$b_t-b_s = \sigma\Delta_t = \gamma\, a_{st} + e_{st},$

where $\gamma = \gamma_p-\gamma_r$ and $e_{st} = e^p_{st}-e^r_{st}$. Since the distribution of Δ_t and a_{st} have been specified previously, the distribution of e_{st} is fully determined:

$e_{st} \sim N[\pi\sigma(t-s), (\sigma^2-\gamma^2\sigma_a^2)(t-s)].$

Hence, it must hold that: $\gamma^2\sigma_a^2 < \sigma^2$. The parameters γ_r and γ_p can be expected to have the same sign. If a particular shock affects the productivity of a particular worker in the present job, the market option (being the max of all alternatives) will probably change in the same direction. However, this brings us no clue regarding the sign of γ, as this sign depends on which of the two is most sensitive to the observable shocks, r_t or p_t.

We want to relate the observed values of a_{st} to realized patterns of p_t and r_t. However, as long as we do not observe productivity directly, we have no information regarding the magnitude of b_t-b_s for a particular employment relation at a particular point in time. However, we do know that b_T-b_X has to be equal to $-\sigma\Delta$ at the moment of separation. This provides a basis for empirical inference. Clearly, running a regression where Δ is the left-hand variable and a_{XT} the explanatory variable fails since our model implies that b_T-b_X is equal to $-\Delta$ only at the moment of separation, and greater than $-\Delta$ before. This is a highly selective sample from all possible trajectories b_T-b_X, violating the standard assumptions of regression analysis.

Our solution to the selectivity problem in the observations of Δ_t benefits from the fact that the distribution of the observed component a_{XT} is normal. Hence, in an aselect sample, $\sigma\Delta_t$ is the sum of two independent normal variates, γa_{st} and e_{st}. The distribution of one of these variates, conditional on separation (that is: on their sum being equal to $\sigma\Delta$), is again a normal variate with parameters:

$[a_{XT}|\gamma a_{XT} + e_{XT}=\sigma\Delta] \sim N[-\gamma^2\sigma_a^2\,\sigma^{-1}(\Delta-\pi\,T), (\sigma^2-\gamma^2\sigma_a^2)\,\sigma_a^2\,\sigma^{-2}T]$

where we normalize X to zero, so that T measures job tenure. Hence, the stochast $e_{XT}^* \equiv [a_{XT} + \gamma^2\sigma_a^2\,\sigma^{-1}(\Delta-\pi\,T)]/\sqrt{T}$ has zero mean and constant variance and is orthogonal to Δ and T. We can therefore specify the following regression model:

$$a_{XT}/\sqrt{T} = \delta_0 + \delta_1\,T + \delta_2\,\sqrt{T} + \delta_3\,\pi\,\sqrt{T} + \delta_4\,\Delta/\sqrt{T} + e_{XT}^* \quad (10)$$

From the distribution of a_{XT}, we have the following equalities:

$\delta_0 = \delta_1 = \delta_2 = 0,$
$\delta_3 = -\delta_4 = \gamma^2\sigma_a^2/\sigma,$
$V[e_{XT}^*] = (\sigma^2-\gamma^2\sigma_a^2)\,\sigma_a^2\,/\sigma^2$

The estimation results from Section 3 can be used for calculation of Δ and π. Because γ is the only unknown parameter (since a_{st} is observed, σ_a^2 can be inferred from the data), we have five over-identifying restrictions, which can be applied for testing.

4.2 Recovering the Brownian from Unemployment Data

The next issue is what variable can serve for a_{XT}. The productivity of individual matches is affected by a multitude of shocks. Most of these shocks are idiosyncratic (either industry or firm or even individual specific), see Davis and Haltiwanger (1992). However, part of them will be aggregate shocks. The methodology set out above can be applied to both types of shocks. However, here we use data on aggregate shocks only. Our plan is to use monthly unemployment data at the state level. Obviously, unemployment is not a Brownian. Random shocks push the unemployment rate up and down every now and then, but Smith's invisible hand pushes it back to its natural rate in a properly functioning market economy. This suggests the evolution of unemployment to be an auto-regressive process. So, although the unemployment rate itself is not a Brownian, the underlying process may very well be.

We have data available for 50 states and Washington DC, running from January 1978 through January 1996 (215 months). We use these data to reconstruct this underlying process by estimating a VAR(2) model, allowing for lagged interstate effects. Per state, we therefore estimate 2 lags x 51 states + 1 intercept = 103 coefficients. Unemployment is stationary in all but two states. The restriction of all interstate effects being equal to zero is rejected. A VAR(1) model is rejected against the VAR(2) model. Data limitations do not allow the estimation of a full VAR(3) model, but without interstate effects, the third lag is insignificant in all but 2 states. Observations for a_{st} for each state are calculated by adding up minus the residuals of this model from s to t; a_{st} is therefore constructed to be a Brownian with zero drift. These data will be applied for the estimation of equation (10).

4.3 Estimation Results for the Effect of Aggregate Shocks

The estimation results for the standard version of equation (10) suggest that aggregate shocks do not affect the trajectory of b_t. None of the coefficients is significant. However, the sensitivity of employment to aggregate shocks is likely to vary strongly between industries. For example, in construction, employment goes up and down with every recovery and slowdown, while in retail trade employment might even be countercyclical. Hence, workers quit from retail trade and stay in construction during the upswing while they are laid off from construction and stay in retail trade in the downswing, and vice versa. This might explain why we do not find any effect on separations when we enter aggregate shocks with the same coefficient for all jobs. We investigate this issue by adding cross effects of $a_{X\text{-}1,X}$ with π and Δ:

$$a_{XT}/\surd T = \delta_0 + \delta_1\, T + \delta_2 \surd T + \delta_3\, \pi \surd T + \delta_4\, \Delta/\surd T + \\ \delta_{21}\, a_{X\text{-}1,X} \surd T + \delta_{31}\, a_{X\text{-}1,X}\, \pi \surd T + \delta_{41}\, a_{X\text{-}1,X}\, \Delta/\surd T + e_{XT}^{*} \qquad (11)$$

$a_{X\text{-}1,X}$ measures the state of the business cycle at the moment that the employment relation starts. By using $a_{X\text{-}1,X}$ instead of $a_{X,X+1}$, spurious correlation with the endogenous variable is avoided.

The estimation results presented in Table 6 strongly support the notion that jobs that started at a trough will likely end when the economy recovers and jobs that started at a peak will likely end when the economy gets into a recession. A downward shock to the economy raises b_t for jobs that started at a trough. This seems counter-intuitive. However, $b_t = p_t\text{-}r_t$. The downward shock is expected to have a negative impact on both the productivity in the present match and on the outside option. When the negative effect on the latter exceeds that on the former ($\gamma_r>\gamma_p$), the relative productivity b_t moves up. In that case, workers stay in their present job, not because its productivity is so high, but because the alternatives are even worse. Aggregate shocks account, on average, for 9 percent of the variance. This is an amazingly large number. In a world so diversified and specialized as a modern economy, where the impact of each technological innovation differs strongly between job types or industries, we had not expected so large a contribution of a single indicator. We also tested specifications where we used somewhat different indicators for the state of the labor market at the moment of job start, e.g. $a_{X\text{-}2,X\text{-}1}$ and $a_{X\text{-}3,X}$. These specifications all yielded similar results.

The test statistic for the F-test $\delta_0 = \delta_1 = \delta_2 = \delta_{21}=0$, $\delta_3= -\delta_4$, $\delta_{31}= -\ \delta_{41}$ is 3.35, which exceeds the critical value at 1 percent (2.82). When we drop the final restriction, the others are acceptable. This final restriction refers to the proportionality of the effect of the initial surplus Δ and the drift π. For longer lasting jobs, their ratio as estimated from Section 3 gets distorted, which might be due to a non-linearity in the drift. The restriction becomes almost acceptable when we drop jobs with long tenures from the sample. For the 5,616 jobs with tenures less than 250 weeks, the F-statistic is 2.85. This lower F-statistic is not due to the lower precision in the estimation by the reduction in the number of observations, since when we estimate the model only for the observations above 250 weeks of tenure the F-statistic goes up to 3.68. When we account for the effect of unobserved characteristics on Δ, see Van der Ende (1997), the restrictions are acceptable. The random growth model therefore fits the data well, as long as we consider jobs that last less than 5 years. This evidence offers support for the notion that separations of matches are triggered by the accumulated shocks over their total duration, not by the last shock, as is predicted by search models.

Can this evidence be explained by learning models of the type of Jovanovic (1979a)? In our view, this is hard to believe. Remember from Section 2.3 that in the learning model the expected value of the output of a match is match specific, but constant over time. Present realizations of output provide worker and firm information on this match specific expected value, but they do not affect the expected value itself. Hence, the only way in which aggregate shocks can lead to separations in a learning model is when workers and firms erroneously take the aggregate shock to be match specific.

5 From Individual Matches to a Model of Firm Level Employment

5.1 The General Model

Hitherto, the literature on individual matches has not been well connected to that on firm level employment. An important advantage of the random growth model is that it can be easily embedded in Bentolila and Bertola's (1990) model of the evolution of firm level employment when there is uncertainty about future labor demand. This requires only one additional assumption, which moreover has an interesting economic interpretation. In the model of Bentolila and Bertola (1990), a profit-maximizing firm faces a demand curve with constant elasticity η>1:

$$q_t = -\eta(p_t^{\,*} + z_t)$$

where q_t and p_t^* are log output and log price respectively, and where log market index z_t is a Brownian with drift. For simplicity, the log outside option of workers r_t is normalized to zero, so that $p_t = b_t$. Furthermore, productivity per worker is normalized to unity, so that output is equal to employment. Consider the case where firms make all specific investments and reap their full surplus (β = 0). In the model of Bentolila and Bertola, firms pay hiring and firing costs proportional to the number of workers they hire and fire respectively. Since our model does not have firing cost (though their introduction would be rather simple), we set firing cost to zero. The per worker hiring cost are identified as the specific investment I.

Bentolila and Bertola show that under these assumptions the firm hires workers when p_t^* reaches an upper bound $p^+ > 0$ and fires workers when p_t^* reaches a lower bound $p^- < 0$. Our claim is that the model of Bentolila and Bertola yields the same pattern of job durations of individual workers as the model in Section 2, if we supplement their model with a particular rule for the order in which workers are laid off. Firms have to fire the workers first that are hired last (Last-In-First-Out). A simple way to model this is to attribute to every worker a seniority index q, which is conveniently defined as the log employment level q_t at the moment that the worker is hired. If the firm wants to fire workers it is obliged to fire the workers with the highest seniority index, which are by construction the workers which are hired last. Kuhn (1988) and Kuhn and Roberts (1989) offer a rationalization for this type of agreement, which will be discussed in Section 5.2.

The situation is depicted in the graph form in Figure 5. The horizontal lines at p^+ and p^- represent the hiring and firing thresholds. Suppose that at time s, employment is equal to q^* and output prices are equal to p^+. Hence, the firm is at its hiring threshold. A further upward shift in product demand due to an increase of z_s will lead to an increase in employment. A downward shift will have no immediate effect on employment because the output price ends up between the hiring and firing threshold. Whatever, the future evolutio of z_t, worker q^* will be fired at the first moment t, $t>s$, that $z_t = z_s + p^- - p^+$, that is, the first moment that the random walk has traveled down the distance $p^+ - p^-$. This is exactly the same process as described in Section 2. The distance $p^+ - p^-$ is equivalent to $\sigma\Delta$ in Section 2.

The analogy can even be extended to the nature of the optimization process itself. Consider the log marginal value of product of the q-th worker conditional on the state of demand, z_t, denoted $mr(q,z_t)$. This log marginal value can be calculated by considering the case that the firm does not hire any workers workers beyond the seniority index q. Then, log total revenue for all workers up to q would be equal to $^{\eta-1}/_{\eta}\ q + z_t$ and hence the log marginal revenue of worker q equals:[14]

$$mr(q,z_t) = -^{1}/_{\eta}\, q + z_t + \ln(^{\eta-1}/_{\eta}).$$

Since z_t follows a random walk, the log marginal revenue also follows a random walk. When this log marginal revenue is set equal to the variable b_t as in applied in Section 2, the optimal hiring and firing thresholds of the firm are exactly equal to those derived in Section 2. In fact, every worker is attributed her marginal product of labor as if no further workers were hired. The additional revenues that are collected by hiring extra workers are attributed to these extra workers. In this way, the hiring and firing decision of a worker indexed q can be decoupled from the hiring and firing decisions of workers with a higher or a lower seniority rank. As long as the rank-order of hiring and firing is preserved, the only relevant information for the hiring and firing decision of a particular worker is the marginal productivity given her seniority index (that is: ignoring the output of all workers with a higher seniority index).

When hiring and firing cost in the model of Bentolila and Bertola (1990) converges to zero the interval between the upper bound p^+ and lower bound p^- converges to zero, too. Then, p_t^* is a constant and hence $q_t = -\eta\ z_t$. Log employment follows a random walk, which is known as Gibrat's law. This law offers a quite accurate description for the evolution of firm size, provided that it is above a certain minimum threshold, see Jovanovic (1982) and the references cited there. The model for individual job durations set out in Section 2 is therefore closely related to this empirical law. This equivalence suggests a simple test of this model of the relation between individual employment relations and the evolution of firm level employment. Consider the empirical testing procedure set out in Section 4.1. Instead of aggregate shocks one can use the evolution of log employment of the firm as explanatory variable. The separation process of individual matches would be closely related to the evolution of the firm's employment level.

5.2 Hold Up and LIFO

The LIFO lay-off rule provides a device for decoupling the hiring and separation decisions of different workers within the same company. This decoupling gets practical importance, as soon as workers pay for some share β of the specific investments I, $0 < \beta < 1$. Workers are compensated for these investments by awarding them a share β of the surplus of log productivity above their outside option, $mr(q,z_t)$:

$$w(q,z_t)= \beta\, mr(q,z_t) \tag{12}$$

where we use the linear Taylor expansion that is also applied in equation (4). Since $mr(q,z_t)$ plays the same role as b_t in the model of Section 2.5, this set up relates the wage setting in this

[14] $mr(q,z_t)$ is below the log market price at output level q by a constant term $\ln(^{\eta-1}/_{\eta})$, since it takes account of the negative effect of hiring an additional worker on the price obtained for the output of the intramarginal workers.

firm level model to the discussion on tenure profiles in Section 2.4. The *expected* discounted value of the tenure profile is equal to the worker's share in specific investment. However, the *actual* return depends on the evolution of firm's demand curve, that is, on future realizations of z_t.

The implication of this set up is that workers who differ by their degree of seniority q, but who are otherwise homogeneous, receive different wages. The problem with this set up is that senior worker's wages are vulnerable to the firm hiring new workers, since these are perfect substitutes for incumbents. The firm could negotiate a lower wage today to these new hires by promising them parts of the returns on specific investments that would otherwise go to incumbents. This threat of the firm introduces a hold-up problem. Workers invest less because they know that they will not be able to appropriate their full expected share in future surpluses. One possible strategy for incumbents is to oppose any further hiring, because this endangers their claims on the surplus. This is the extreme insider-outsider theory. The drawback of this strategy is that gains of trade remain unexploited. A more efficient solution therefore is to protect the claims of incumbents by a LIFO lay-off rule, which prevents the firm to replace expensive incumbents by cheap new hires. A LIFO lay-off rule can therefore be viewed as a device to deal with the hold-up problem in firms with otherwise homogeneous workers who bear part of the cost of specific investments and therefore share in subsequent surpluses.[15]

This model of wage setting has direct empirical implications. The firm sets employment q_t as such that $mr(q_t,z_t)=\sigma\Delta$, or:

$$q_t = \eta[z_t + \ln(^{\eta-1}/_{\eta}) - \sigma\Delta].$$

Hence, $\exp[q-q_t] = \exp[\eta\sigma\Delta-\eta mr(q_t,z_t)]$ is a rank index of the seniority of a worker q within the firm's seniority hierarchy. The index takes value 0 for the most senior worker and value 1 for the least senior worker. It then follows immediately that wages are a function of this seniority rank index, compare equation (12). So, if this model is correct, the tenure variable showing up in a cross section log earnings regression is in fact a proxy for a tenure rank index. With proper data on this seniority index $\exp[q-q_t]$, the prevalence of this index above tenure as an explanatory variable in a wage regression could be tested.

6 Concluding Remark

In this paper, we present some pieces evidence for the relevance of the random growth model of the job tenure distribution. The hump shaped pattern of separations rates predicted by the model fits the pattern observed empirically quite closely. In particular, where the learning model predicts separation rates to converge to zero quickly, the random growth model generates the fat tail in separations observed empirically. The evidence on the impact of shocks suggest that separations are driven by the accumulated history of shocks running from the start of the job up till the moment of separation, contrary to what is predicted by search model, where only the last shock matters. The results in the paper provide therefore for evidence in favour of the random growth model. However, it seems unlikely that either random growth, or learning or search can explain exclusively the empirical evidence. A realistic model should probably include all three

[15] A similar argument can be found in Kuhn (1988), who considers a world where first workers set wages and then the firm sets employment.

ingredients. The more relevant question would then be their relative importance in explaining the data.

The evidence on the importance of accumulated shocks above presented in Section 4 favors the random growth model above the search model. However, the observed aggregate shocks explain only 10 % of the variance of innovations in job productivity. Nothing guarantees that there is no transitory component in the remaining 90 %. The merit of the search model is that it offers a natural explanation for the negative correlation between separation rates and the experience at the date of job-start. Search selects good draws from the offer distribution. The more experienced the worker, the longer this selection process has been going on and the less likely it is that a new offer yields a further improvement. However, search is not the only explanation for the fall in separation rates in the course of the worker's career. When the optimal level of specific investments is complementary to the worker's experience, this leads to a similar experience related pattern in separation rates. In our model, this implies a positive correlation between the initial surplus and experience at the date of job-start, as we observe empirically, see Section 3. Alternatively, one can assume that experienced workers are more efficient in the accumulation of job-specific experience, leading to an upward drift in job productivity, or equivalently, that they are less efficient in the accumulation of general experience, leading to a downward drift in the outside option. In the random growth model, both mechanisms lead to a positive correlation between the drift and experience, again as is observed empirically. Hence, a more complete evaluation of the importance of search vis-à-vis the random growth model requires further implications to be incorporated in the model.

Appendix: Derivation of the Investment Equation

The derivation follows Dixit and Pindyck (1993), Section 7.1. It starts from Bellman equations (3) in the text:

$$\rho_r V = (\sigma\pi + \tfrac{1}{2}\sigma^2)V'B + \tfrac{1}{2}\sigma^2 V'' B^2$$
$$\rho_r J = (\sigma\pi + \tfrac{1}{2}\sigma^2)J'B + \tfrac{1}{2}\sigma^2 J'' B^2 + B - 1$$

where:
$V \equiv V(B_t)$,
$J \equiv J(B_t)$,
and where *V’* and *J’* are the first and *V'* and *J'* the second derivatives of *V* and *J*. We leave out the subscript *t* of *B* for the sake of convenience. These are two second order differential equations. A particular solution to the second equation is:

$$\rho_r J = B/\rho_p - 1/\rho_r$$

The characteristic roots of the homogeneous part of both differential equations read:

$$\beta_{1,2} = -\pi/\sigma \pm \sqrt{(\pi/\sigma)^2 + 2\rho_r/\sigma^2}$$

For our application it is often more convenient to work with the parameter $\alpha = \sigma\beta$. Hence:

$$\alpha_{1,2} = -\pi \pm \sqrt{\pi^2 + 2\rho_r}$$

The solution to the Bellman equations can now be written as:

$$V = A_1 B^{\beta_1}$$
$$J = A_2 B^{\beta_2} + B/\rho_p - 1/\rho_r$$

The first term in both equations is the option values. The option value of a vacancy is increasing in *B* for a vacancy (the option increases in value when the prospect of filling becomes closer) and converges to zero for low values of *B*. The option value of a job is decreasing in *B* (separation becomes more realistic when productivity goes down) and converges to zero for high values of *B*. Hence, we apply the positive and negative root of β in both equations respectively. The value equivalence conditions ($V = J$ for separation; $V = J - I$ for hiring) and the smooth pasting conditions ($V' = J'$, both for separation and hiring) read:

$$A_1 B_T^{\beta_1} = A_1 B_T^{\beta_2} + B_T \big/ \rho_p - 1 \big/ \rho_r$$

$$\beta_1 A_1 B_T^{\beta_1} = \beta_2 A_2 B_T^{\beta_2} + B_T \big/ \rho_p$$

$$A_1 B_R^{\beta_1} = A_2 B_R^{\beta_2} + B_R \big/ \rho_p - 1 \big/ \rho_r - I$$

$$\beta_1 A_1 B_R^{\beta_1} = \beta_2 A_2 B_R^{\beta_2} + B_R \big/ \rho_p$$

Usually this system is solved for A_1, A_2, B_T, and B_X. In our case, we want to have a solution for I, given our estimate of Δ. Since $e^{\Delta} = [B_X / B_T]^{\sigma} \equiv D$, we use this relation to substitute for B_T and solve the system for the remaining unknowns. This procedure yields the system of equation (5) in subsection 2.3.

References

Abowd, J.A. and T. Lemieux (1993), "The effect of product market competition on collective bargaining agreements: the case of foreign competition in Canada", *Quarterly Journal of Economics* 108, pp. 983-1014.

Abraham, K.G. and H.S. Farber (1987), "Job duration, seniority, and earnings", *The American Economic Review* 7, pp. 278-297.

Aghion, P. and P. Howitt (1992), “A model of growth through creative destruction”, *Econometrica* 60, pp. 323-351.

Altonji, J.G. and R.A. Shakotko (1987), "Do wages rise with job seniority", *Review of Economic Studies* 59, pp. 437-59

Becker, G.S. (1962), "Investment in human capital: A technological analysis", *Journal of Political Economy* 70, pp. S9-S49.

Bentolila, S. and G. Bertola (1990), "Firing costs and labour demand: How bad is euroclerosos?", *Review of Economic Studies* 57, pp. 417-434.

Blanchard, O.J. and P.A. Diamond (1994), "Ranking, unemployment duration, and wages", *Review of Economic Studies* 61, pp. 417-434.

Bowlus, A.J. (1995), "Matching workers and jobs: Cyclical fluctuations in match quality", *Journal of Labor Economics* 13, pp. 335-350.

Burdett, K. and D. Mortensen (1998), “Equilibrium wage differentials and employer size”, *International Economic Review* 39, pp. 257-274.

Caballero, R.J. and M.L. Hammour (1994), “The cleansing effect of recessions”, *American Econmic Review* 84, pp. 1350-1368.

Davis, S.J. and J. Haltiwanger (1992), "Gross job creation, and employment reallocation", *Quarterly Journal of Economics*, pp. 820-863.

Dixit, A.K. (1989), "Entry and exit decisions under uncertainty", *Journal of Political Economy* 97, pp. 620-638.

Dixit, A.K., and R.S. Pindyck (1993), "Investment under uncertainty", Princeton University Press, Princeton, U.S.A.

Dustman, C. and C. Meghir (2001), "Wages, experience and seniority", working paper, IFS/UCL, London.

Farber, H.S. (1994), "The analysis of interfirm worker mobility", *Journal of Labor Economics* 12, pp. 554-593.

Feldstein, M. (1976), "Temporary lay-offs in the theory of unemployment", *Journal of Political Economy* 84, pp. 937-957.

Grout, P.A. (1984), "Investment and wages in the absence of binding contracts: A Nash bargaining approach", *Econometrica*, pp. 449-460.

Holmlund, B. and J. Zetterberg (1991), "Insider effects in wage determination: evidence from five countries", *European Economic Review* 35, pp. 1009-35.

Jacobson, L.S., R.J. LaLonde, and D.G. Sullivan (1993), "Earnings losses of displaced workers", *The American Economic Review* 83, pp. 685-709.

Jovanovic, B. (1979a), "Job matching and the theory of turnover", *Journal of Political Economy* 87, pp. 972-990.

Jovanovic, B. (1979b), "Firm-specific capital and turnover", *Journal of Political Economy* 87, pp. 1246-1260.

Jovanovic, B. (1982), "Selection and the evolution of industry", *Econometrica* 50, pp. 649-670.

Jovanovic, B. and R. Moffitt (1990), "An estimate of a sectoral model of labor mobility", *Journal of Political Economy* 98, pp. 827-852.

Keane, M.P., and K.I. Wolpin (1997), "The career decisions of young men", *Journal of Political Economy* 105, pp. 473-522.

Kuhn, P. (1988), "A nonuniform pricing model of union wages and unemployment", *Journal of Political Economy* 96, pp. 473-508.

Kuhn, P. and J. Robert (1989), "Seniority and distribution in a two-worker trade union", *Quarterly Journal of Economics*, pp. 485-505.

Kuhn, P. and A. Sweetman (1996), "Vulnerable seniors: Unions, tenure and the cost of job loss", Working Paper McMaster University.

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

Lancaster, T., G. Imbens, and P. Dolton (1987), "Job separations and job matching", in: R.D.H. Heijmans and H. Neudecker (Eds.), *The Practice of Econometrics*, pp. 31-43, Martinus Nijhoff Publishers.

McDonald, R. and D. Siegel (1986), "The value of waiting to invest", *Quarterly Journal of Economics* 101, pp. 707-727.

Miller, R.A. (1984), "Job matching and occupatinal choice", *Journal of Political Economy* 92, pp. 1086-1120.

Parent, D. (1999), "Wages and mobility: the impact of employer-provided training", *Journal of Labor Economics* 17-2, pp. 298-317.

Teulings, C.N. and J. Hartog (1998), “Corporatism and competition: labour contracts, institutions and wage structures in international comparison”, Cambridge University Press.

Topel, R.H. (1991), "Specific capital, mobility, and wages: Wages rise with job seniority", *Journal of Political Economy* 99, pp. 145-176.

Topel, R.H. and M.P. Ward (1992), “Job mobility and the careers of young men”, *Quarterly Journal of Economics* 107, pp. 439-479.

Van der Ende, M.A. (1997), "On the values of jobs and of specific training", Ph.D. thesis, University of Amsterdam, The Netherlands.

Table 1 Means of selected variables	
Variable	Mean
% jobs observed to end	73.6
tenure (in weeks)	100.3
grade/10	1.28
local unemployment rate (in %)	8.7
experience at job start/100 (in weeks)	2.10
unemployment spell prior to job start/100 (in weeks)	.15
spouse present	.34
central city	.12
goverment job	.07
union	.14

Table 2 Job training in 8,339 jobs					
type	Observati-ons	Hours Known	Median Hours	Mean Hours	Survives Job Exit
on-job-training	159	115	107	400	12
vocational/technical	112	64	144	782	23
business college	23	18	35	149	5
corres. course	405	328	44	171	14
seminars	299	262	24	63	4
other	91	75	80	313	15
total	1,089	862	40	226	73

Table 3 ML estimates of (9)

Parameter	π	*t*-value	Δ	*t*-value
intercept	-.036	14.23	6.642	55.28
grade/10	.039	5.54	-1.118	3.56
unemployment (%)	.008	1.85	- 3.33	2.03
experience/100	.010	7.69	.018	0.48
prior unempl./100	-.015	3.81	-.491	3.93
spouse present	.017	4.66	.229	1.63
central city	-.015	2.83	.136	0.69
government job	.032	5.17	-.851	3.62
union	.031	7.01	-.067	0.39
dummies:				
on job training	-	-	5.014	7.69
voc/technical	-	-	4.726	8.63
business college	-	-	3.917	2.72
corresponding course	-	-	4.925	13.37
seminar	-	-	6.194	13.63
other programs	-	-	2.985	4.54
26 occupation dummies	no	-	yes	-
sdev. worker chars.	.045	16.10	1.897	34.13

Table 3 (cont'd) with occupational effects. Skilled craft workers are the reference group			
Occupation	Distance	|*t*-value|	Nr. Jobs
early jobs			
waiter	0.014	0.03	72
unskilled craft	-0.258	-0.99	300
food, household	0.139	0.61	451
nurse aid, maid	0.400	1.61	359
communication worker	0.320	0.55	60
agriculture worker	-0.297	-1.25	334
middle class jobs			5,174
sales worker	0.584	2.47	513
medium skilled craft			1,495
guard, packer	0.388	0.97	142
medical assistant	1.880	1.70	27
low skilled craft	0.564	0.98	1,093
restaurant manager			79
secretary	0.708	0.76	30
teacher nonsecondary	1.549	2.83	105
clerk	1.291	4.89	450
skilled craft			1,335
technician	1.151	2.67	162
specialized jobs			
specialized craft	0.871	1.64	106
protective services	1.908	2.14	53
manager and related	1.249	3.41	241
sales representative	2.226	3.33	92
medical specialist, pilot	1.453	1.19	25
medical skilled	3.566	2.64	28
manager n.e.c.	1.980	5.63	370
computer operator	1.933	2.23	51
public servant, scientist	2.393	5.94	228
leisure service	1.392	1.15	24
specialist, lawyer	2.509	5.70	192

Table 4 Standard deviation of shocks, investment, (both in weeks), and separation tresholds					
π=-0.036, Δ = 6.66					
σ	I	B_T	B	$dI/d\Delta$	$\partial\Delta/\partial H$
0.005	0.235	0.985	1.018	0.106	9.649
0.010	0.471	0.970	1.036	0.211	4.902
0.015	0.707	0.956	1.055	0.318	3.320
0.020	0.943	0.941	1.074	0.425	2.529
0.025	1.181	0.927	1.093	0.532	2.056
0.030	1.418	0.912	1.112	0.639	1.740
0.035	1.656	0.898	1.132	0.747	1.515
0.040	1.895	0.884	1.152	0.855	1.347
0.045	2.134	0.871	1.172	0.964	1.216
0.050	2.373	0.857	1.192	1.072	1.111
0.055	2.612	0.843	1.212	1.181	1.026
0.060	2.852	0.830	1.233	1.291	0.955
0.065	3.092	0.817	1.254	1.400	0.896
0.070	3.333	0.803	1.275	1.510	0.845
0.075	3.573	0.790	1.297	1.620	0.801

Table 5	Tenure profile for σ = 0.060, π = -0.036, Δ = 6.66	
years	$\sigma\Delta_t$	0.30 $\sigma(\Delta_t\text{-}\Delta)$
0	0.400	0.000
1 month	0.388	- 0.002
3 months	0.403	0.002
1	1.022	0.188
2	1.322	0.278
3	1.523	0.338
5	1.795	0.420
10	2.175	0.534
15	2.389	0.598
20	2.531	0.641

Table 6	Extended estimates of (10) for uncensored tenures; Δ_T= 0. Dependent variable: a_{1T} (monthly aggregated unemployment residuals).	
Variable	OLS	|*t*-value|
intercept	.0000	.0
T (in weeks)	-.0000	.1
$\sqrt{T}$	-.0001	.2
$a_{01}\sqrt{T}$	.0023	2.2
$\pi\sqrt{T}$	.0008	.4
$a_{01}\pi\sqrt{T}$	-.0180	1.0
$\Delta/\sqrt{T}$	-.0005	.4
$a_{01}\Delta/\sqrt{T}$	-.0560	17.9
σ_a^2	.0012	
R^2	.091	

Figure 1: The reflection principle

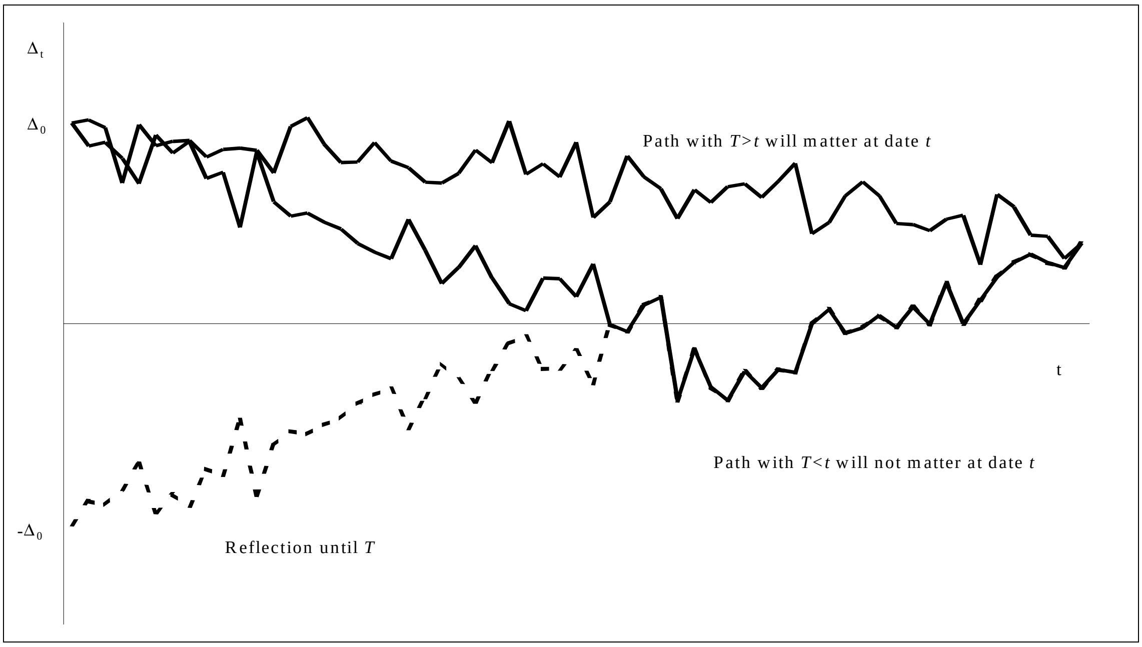

Figure 2: A comparison of the learning and random growth model ($\Delta_{\text{random growth}} = 4$, $\Delta_{\text{learning}} / \sqrt{\delta} = 3.8$, $\delta = 0.01$).

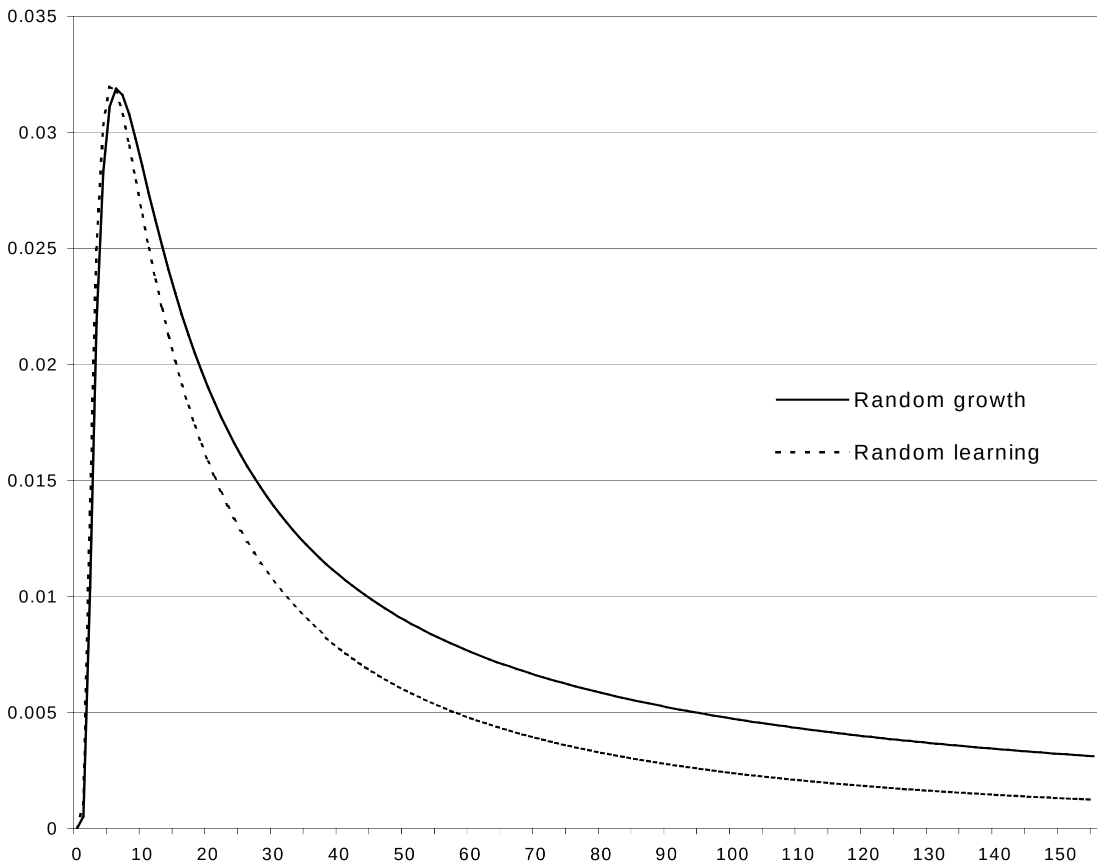

Figure 3: Weekly job exit percentage for the first year

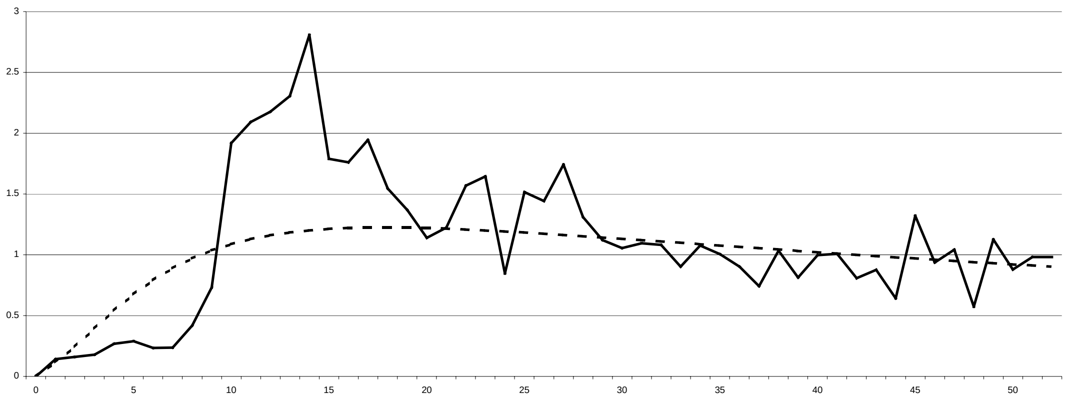

Figure 4: Monthly job exit percentage for the first fourteen years

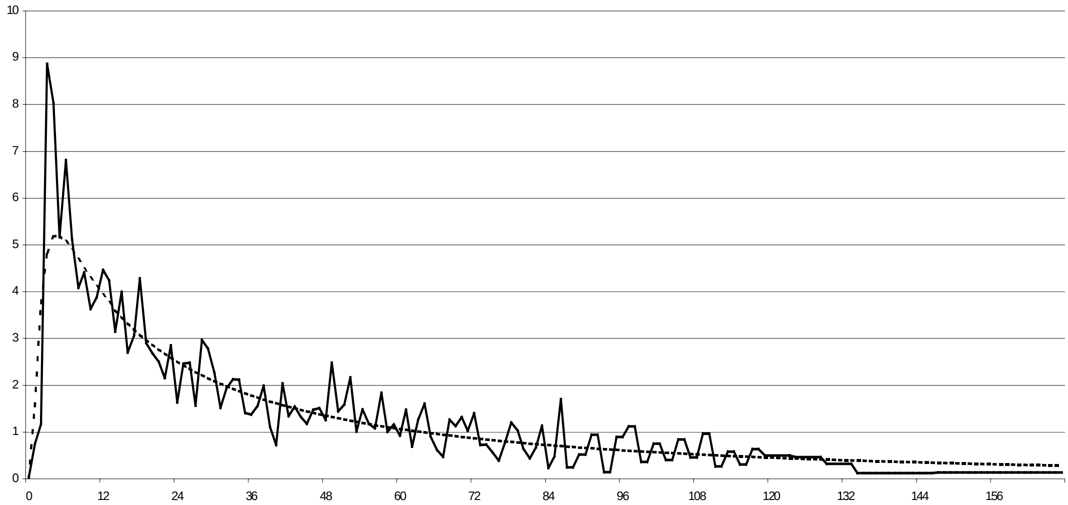

Figure 5: The relation between the individual and the firm level