ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Teulings, Coen N.; van der Ende, Martin A.

Working Paper

A Structural Model of Tenure and Specific Investments

Tinbergen Institute Discussion Paper, No. 00-009/3

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Teulings, Coen N.; van der Ende, Martin A. (2000) : A Structural Model of Tenure and Specific Investments, Tinbergen Institute Discussion Paper, No. 00-009/3, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/85701

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

A Structural Model of Tenure and Specific Investments

Coen N. Teulings* and Martin A. van der Ende**

November, 2002

Abstract

Two stylized facts on job tenures are the increasing tenure profiles in wages and the hump shaped tenure profile in job separation rates. Both can be generated by a model with specific investments and random shocks to productivity after the job start. Assumptions on risk neutrality and efficient bargaining on wages allow the job separation decision to be analyzed separately from wage setting. A simple bargaining rule for surplus sharing between worker and firm generates the tenure profile in wages. Estimation results using job tenure data from the NLSY support the humped shaped pattern in job separation rates and favor this model above the learning model. When assuming a Last-In-First-Out (LIFO) separation rule, this model of individual employment 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.

Key words: option value, job tenure, tenure profiles.
J.E.L. J63: Turnover, Vacancies
Field positioning: labor economics

* Erasmus University Rotterdam and Tinbergen Institute, P O Box 1738, NL-3000 DR Rotterdam, The Netherlands (e-mail: teulings@few.eur.nl).
** Nederlands Economisch Instituut, P O Box 4175, NL-3006 AD Rotterdam, The Netherlands (e-mail: endem@nei.nl).

The authors thank Nicholas Bloom, Pieter Gautier, Lennart Janssens, Maarten Lindeboom, Audra Bowlus, Joop Hartog, Jan van Ours, Guiseppe Bertola and participants of seminars at UCL, NorthWestern University, University of Uppsala, and CREST in Paris for helpful comments.

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 some specific investments. When the match productivity evolves unfavorably, the investments loose 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 models on the one hand and learning models on the other hand. In search models (e.g. Jovanovic (1979b), Burdett and Mortensen (1998)), 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. When the value of a newly arrived offer exceeds the value of the present job, there is separation. In search models, both types op shocks are usually modelled as transitory shocks. The arrival of a job offer today does not affect the probability of the arrival of an offer tomorrow, neither does the value of one job offer affect the value of the next. 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 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 mean value of the output is the quality of the match. The worker and the firm gradually learn about match quality by observing realizations of this random output. Hence, shocks have a permanent effect on the probability of future separations. However, the more shocks are accumulated, the less is the share of each shock. In this paper, we contrast the (Bayesian) implication that the impact of new information decreases as the worker and firm are better informed, with the random growth model. The learning model generates a hump shaped separation 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. Gradually, beliefs become more accurate and bad matches are eliminated. Finally, beliefs are almost exact, and all bad matches have been eliminated previously, so that the separation rate drops to zero.

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 permanent, cumulative effect on future separations.[1] However, unlike the learning model, the effect of new shocks does not decline in the course of time.

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 loose 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 at the moment of job start is equal to the cost of investment.

This model yields a hump shaped separation rate. At first, the separation rate is low. Players would not have made the specific investment if the resulting initial 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 while the remaining matches will be predominantly those which have accumulated positive shocks.

Lancaster, Imbens, and Dolton (1987) and Miller (1984) use the hump shaped separation rates as evidence for the learning model. However, their evidence suggests that separation rates do not decline as quickly as predicted by the estimated learning model. With a negative drift, the separation rates from the random growth decline more slowly, and converge to a positive value as empirically observed.

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

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 extend the empirical analysis by using 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.

We explore the implications for the tenure profile in wages. 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 that develop unfavorably, are eliminated from the stock of ongoing matches. The remaining stock is therefore a selective sample of the random walks where productivity has developed favorably. This selection mechanism generates a tenure profile in wages, which is consistent with what is reported by for example Topel (1991).

An important advantage of the random growth model of individual employment relations is that it can be 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.

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. In Section 4, we analyze the relation of our model with that of Bentolila and Bertola (1990). Section 5 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 which we normalize to unity per unit of time. 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 I have to be paid. All specific investments are made at the start of the employment relationship. This assumption is obviously restrictive but it is a reasonable first approximation. 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. Both the worker and the firm are perfectly informed about the present value of P_t. However, its future evolution is subject to uncertainty. P_t follows 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 between the worker and the firm. We 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 the moment of separation 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.

2.2 The Tenure Distribution

Using lower cases for logarithms, the law of motion for the market value of log output between arbitrary dates s and t shocks is:

$$[p_t - p_s] \sim N[(t - s)\,\mu, (t - s)\,\sigma^2] \tag{1}$$

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

$$\begin{aligned} V(P_t) &= \mathrm{E}_t\left[e^{-\rho(X-t)}\left(J(P_X) - I\right)\right] \\ J(P_t) &= \mathrm{E}_t\left[e^{-\rho(T-t)}\,V(P_T)\right] + \mathrm{E}_t\left[\int_t^T e^{-\rho(s-t)}(P_s - 1)ds\right] \end{aligned} \tag{2}$$

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, $t \leq X,T$ and where $\mathrm{E}_t[.]$ is the expectation operator for the filter P_s, $s > t$. The value of a vacancy is fully determined by the option of filling the vacancy by making the specific investment 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, minus what she would earn on the outside market. 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):

$J(P_X) = V(P_X) + I,$
$V(P_T) = J(P_T).$

The stopping rules for hiring and separation depend only on P_t because the law of motion formula (1) implies the strong Markov property for P_t and because the time horizon is infinite, see McDonald and Siegel (1986, page 712-713). Its logarithm P_t is a Brownian with drift μ and variance rate σ^2. The firm hires a worker at the moment X when P_t hits an upper bound P_X and the worker separates from the firm at the moment T when this ratio hits a lower bound P_T.

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 P_t is at the separation level P_T. Since the employment relation has started at time t such that P_t is at the hiring level P_X, the duration until separation is determined by the time that is required for the random walk to travel down the distance P_X-P_T. Define: $\Delta_t \equiv (P_t - P_T)/\sigma$, $\Delta \equiv (P_X - P_T)/\sigma$, and $\pi \equiv \mu/\sigma$. 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. Then the job duration is the required time for Δ_t to travel down the distance Δ. Hence, the distribution of job tenures is fully determined by two parameters, Δ and π. Contrary

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

to π, Δ is not a structural parameter. It depends on the optimal hiring and separation thresholds p_X and p_T, which will be determined in Section 2.5.

The distribution of $\Delta_t - \Delta$ conditional on the hiring time X is $N[(t-X)\pi,(t-X)\sigma^2]$. 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 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. These trajectories should be subtracted when calculating the density of all trajectories that never crossed the line $\Delta_s = 0$. Hence, the joint 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 is equal to unity 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 $0 < t_0 < {}^2/_3\Delta^2$;

[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 $\lambda'' < 0$.

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 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 productivity in remaining jobs shifts upward. When the drift is positive, the drift term dominates the random shocks in the long run. When the drift is negative, there is 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 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)) also yield a hump shaped pattern in separation rates. However, the random growth model generates 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 P_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 P_t about match productivity by Bayes' rule:

$$P_t = \frac{1}{1+\delta t}\left[\sum_{s=1}^{t} \delta x_s + P_0\right]$$

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 P_t is below a certain threshold, P_T. In the full model, the separation threshold is time dependent since the variance of the beliefs P_t decreases over time. The smaller the variance of P_t, the lower the option value of continuing the employment relation. Hence, the separation threshold increases over time. However, this effect is 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, $\Delta \equiv (P_0 - P_T)/\sigma_0$ 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 $\Delta_t \equiv (P_t - P_0)/\sigma_0$ have traveled down the distance Δ.

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 $\text{Var}[\Delta_t] = t$. Hence, in the learning model, the accumulation of shocks leading to the gradual evolution of Δ_t proceeds in a time scale 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: $\text{Var}[P_t - P_0] = \text{Var}[x_0 - P_0] = \sigma_0^2$, since beliefs converge to actual match productivity x_0. When we set: $\Delta_{\text{learning}} = \sqrt{\delta}\,\Delta_{\text{random growth}}$, 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 $\Delta_{\text{random growth}} = 4$, while the fat line represents the separation rate for a learning model with $\Delta_{\text{learning}}/\sqrt{\delta} = 3.8$ 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 P_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

[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.

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 $\lambda(\infty) = \frac{1}{2}\pi^2 > 0$. 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 drop the working assumption that the firm captures all surpluses of the specific investment. Instead, we extend the model with a simple sharing rule for the distribution of surpluses between the worker and the firm. Let w_t be the log wage of the worker. By the normalization of the worker's outside option to unity, it vanishes in a log linear representation. Our sharing rule simply distributes instantaneous surpluses proportional to the worker and the firm[6]:

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

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

[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.

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):

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

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 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 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 includes randomness in

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

the outside option, while the outside option is normalized to a non-random constant in our calculations. 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 should be lower just before separation. This prediction gets support in Topel and Ward (1992, Table VI, model (v)) not in Topel (1991, Table 4). This difference matters in particular when a 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 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 on the gains of trade 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 Separation Threshold and the Level of Specific Investments

By applying Ito's lemma, the value functions in equation (2) are defined by two Bellman equations, see Dixit (1989):

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

The first term on right hand side of both equations takes care of the drift in P_t. The second term accounts for the non-vanishing second order effect of shocks to P_t. The final terms in the second equation measure the current output of a filled job, net of the outside option of the worker. The hiring threshold P_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:

$$\begin{aligned} I &= \frac{1}{\sigma\rho}\frac{\sigma\alpha_1\left(1-D^{\sigma-\alpha_1}\right)\left(D^{\alpha_1}-D^{\alpha_1-\alpha_2}\right)+\sigma\alpha_2\left(D^{\sigma-\alpha_2}-1\right)\left(D^{\alpha_1}-1\right)+2\rho\left(D^{\sigma}-1\right)\left(D^{\alpha_1-\alpha_2}-1\right)}{\alpha_1\left(1-D^{\sigma-\alpha_1}\right)D^{\alpha_1-\alpha_2}+\alpha_2\left(D^{\sigma-\alpha_2}-1\right)+\frac{2\rho}{\sigma}\left(D^{\alpha_1-\alpha_2}-1\right)} \\ B_X &= \frac{1}{\sigma}\frac{2(\rho-\pi\sigma-\frac{1}{2}\sigma^2)\left(D^{\alpha_1-\alpha_2}-1\right)D^{1/\sigma}}{\alpha_1\left(1-D^{\sigma-\alpha_1}\right)D^{\alpha_1-\alpha_2}+\alpha_2\left(D^{\sigma-\alpha_2}-1\right)-\frac{2\rho_y}{\sigma}\left(D^{\alpha_1-\alpha_2}-1\right)} \end{aligned} \tag{7}$$

[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 $D \equiv e^{\Delta}$, $\alpha_1 \equiv - \pi + \sqrt{(\pi^2+2\rho)}$, and $\alpha_2 \equiv - \pi - \sqrt{(\pi^2+2\rho)}$. The value of a job is finite if and only if $\rho - \pi \sigma - \frac{1}{2} \sigma^2 > 0$ and $\rho > 0$, which we assume to be the case. Equation (7) implies that P_X is positive and P_T is negative. Immediately after the hiring decision, the productivity in the job P_X exceeds the outside option by at least the interest payments on the specific investment, ρ. If not, the firm would be able to increase its profits by postponing the investment. At the moment of separation, P_T is below unity. 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 model has four parameters: π, σ, ρ, and I. Consider the identification of these parameters, assuming exogenous information on ρ 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. The advantage of the assumptions on risk neutrality, no hold-up, and efficient separation, is that we can separate the wage process completely from the investment and separation decisions. At the same time, wage data do not add to the identification. All additional information in these data would be needed for the identification of an additional parameter, the bargaining power of the worker, determining the sharing of the surplus between the worker and the firm, see Section 2.4. Hence, we have to look for an alternative identification strategy.

The parameter σ is 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, 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 ($P_X > 1$), since intramarginal specific investments have been made, which raise the productivity of the worker above her outside market option. Workers and/or firms 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: $dI/dH = P_X$. Variation in the actual amount of on job specific training can be related to the observed tenure distribution, which allows us to estimate $d\Delta/dH$. I is a (non-linear) function of Δ, denoted $I(\Delta)$, see equation (7). Hence:

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

Expressions for P_X and $I'(\Delta)$ can be obtained from equation (7). Given the availability of estimates of $d\Delta/dH$, 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 $d\Delta/dH$, 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:

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

where $d_{ij} \equiv 1$ if a job is uncensored and $d_{ij} \equiv 0$ otherwise and where j_i 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

[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: $\Delta = \exp(x\beta_{\Delta} + u_{\delta i})$. However, the disadvantage of this specification is that the additive structure $\Delta_t = x\beta_{\Delta} + t\, x\beta_{\pi}$ is lost.

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 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 are 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 except grade which is measured at the start of the career. 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 is 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.

[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.

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 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 is dissipated 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 negative drift could be interpreted as a depreciation rate indicating how fast a job is replaced by a new job with a new technology. 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 only be generated by a pure learning model for extremely low values of δ. 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 equations (7) and (8). We use the estimated intercept values for Δ and π, a value for ρ 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 be 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. Parent (1999) reports a similar result. 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 $d\Delta/dH$. 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 or 43% per year.

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, 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 From Individual Matches to a Model of Firm Level Employment

4.1 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 $\eta>1$:

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

where q_t and p_t^* are log output and log price respectively, and where the log market index z_t is a Brownian with drift. Without loss of generality, 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 ($\beta = 0$). 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 simple), we set firing cost to zero. The hiring costs per worker 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 4.2.

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 evolution 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 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 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 p_t, 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 the relation between individual employment relations and the evolution of firm level employment predicted by this model. The evolution of log employment of the firm serve as explanatory variable for the separation process. The separation rates for individual matches should be negatively related to the evolution of the firm's employment level.

4.2 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, $mr(q,z_t)$:

[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.

$w(q,z_t) = \beta\, mr(q,z_t)$ (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 p_t, this set up relates the wage setting in this 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 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 highering new workers, since these are perfects 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.

[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.

5 Concluding Remark

In this paper, we present 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 closely. In particular, where the learning model predicts separation rates to converge to zero, 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 provide therefore evidence in favor of the random growth model.

Appendix: Derivation of the Investment Equation

The derivation follows Dixit and Pindyck (1993), Section 7.1. It starts from Bellman equations (6) in the text:

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

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:

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

where for shorthand, $\rho_p = \rho - \pi\sigma - 1/2\sigma^2$. The characteristic roots of the homogeneous part of both differential equations read:

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

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

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

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

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

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:

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

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 (7) 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**, 983-1014.

ABRAHAM, K.G. and H.S. Farber (1987), "Job Duration, Seniority, and Earnings", *The American Economic Review*, **7**, 278-297.

AGHION, P. and P. Howitt (1992), "A Model of Growth through Creative Destruction", *Econometrica*, **60**, 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**, S9-S49.

BENTOLILA, S. and G. Bertola (1990), "Firing Costs and Labour Demand: How bad is Euroclerosos?", *Review of Economic Studies*, **57**, 417-434.

BLANCHARD, O.J. and P.A. Diamond (1994), "Ranking, Unemployment Duration, and Wages", *Review of Economic Studies*, **61**, 417-434.

BOWLUS, A.J. (1995), "Matching Workers and Jobs: Cyclical Fluctuations in Match Quality", *Journal of Labor Economics*, **13**, 335-350.

BURDETT, K. and D. Mortensen (1998), "Equilibrium Wage Differentials and Employer Size", *International Economic Review*, **39**, 257-274.

CABALLERO, R.J. and M.L. Hammour (1994), "The Cleansing Effect of Recessions", *American Econmic Review*, **84**, 1350-1368.

DAVIS, S.J. and J. Haltiwanger (1992), "Gross Job Creation, and Employment Reallocation", *Quarterly Journal of Economics*, 820-863.

DIXIT, A.K. (1989), "Entry and Exit Decisions under Uncertainty", *Journal of Political Economy*, **97**, 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**, 554-593.

FELDSTEIN, M. (1976), "Temporary Lay-offs in the Theory of Unemployment", *Journal of Political Economy*, **84**, 937-957.

GROUT, P.A. (1984), "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach", *Econometrica*, 449-460.

HOLMLUND, B. and J. Zetterberg (1991), "Insider Effects in Wage Determination: Evidence from Five Countries", *European Economic Review*, **35**, 1009-35.

JACOBSON, L.S., R.J. LaLonde, and D.G. Sullivan (1993), "Earnings Losses of Displaced Workers", *The American Economic Review*, **83**, 685-709.

JOVANOVIC, B. (1979a), "Job Matching and the Theory of Turnover", *Journal of Political Economy*, **87**, 972-990.

JOVANOVIC, B. (1979b), "Firm-specific Capital and Turnover", *Journal of Political Economy*, **87**, 1246-1260.

JOVANOVIC, B. (1982), "Selection and the Evolution of Industry", *Econometrica*, **50**, 649-670.

JOVANOVIC, B. and R. Moffitt (1990), "An Estimate of a Sectoral Model of Labor Mobility", *Journal of Political Economy*, **98**, 827-852.

KEANE, M.P., and K.I. Wolpin (1997), "The Career Decisions of Young Men", *Journal of Political Economy*, **105**, 473-522.

KUHN, P. (1988), "A Nonuniform Pricing Model of Union Wages and Unemployment", *Journal of Political Economy*, **96**, 473-508.

KUHN, P. and J. Robert (1989), "Seniority and Distribution in a Two-worker Trade Union", *Quarterly Journal of Economics*, 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**, 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*, 31-43, Martinus Nijhoff Publishers.

McDONALD, R. and D. Siegel (1986), "The Value of Waiting to invest", *Quarterly Journal of Economics*, **101**, 707-727.

MILLER, R.A. (1984), "Job Matching and Occupatinal Choice", *Journal of Political Economy*, **92**, 1086-1120.

PARENT, D. (1999), "Wages and Mobility: the Impact of Employer-provided Training", *Journal of Labor Economics*, **17-2**, 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**, 145-176.

TOPEL, R.H. and M.P. Ward (1992), “Job Mobility and the Careers of Young Men”, *Quarterly Journal of Economics*, **107**, 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	Observations	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	1.549	2.83	105
nonsecondary 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/dΔ	$\partial\Delta/\partial H$
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 $\sigma(\Delta_t-\Delta)$
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

Figure 1: The reflection principle

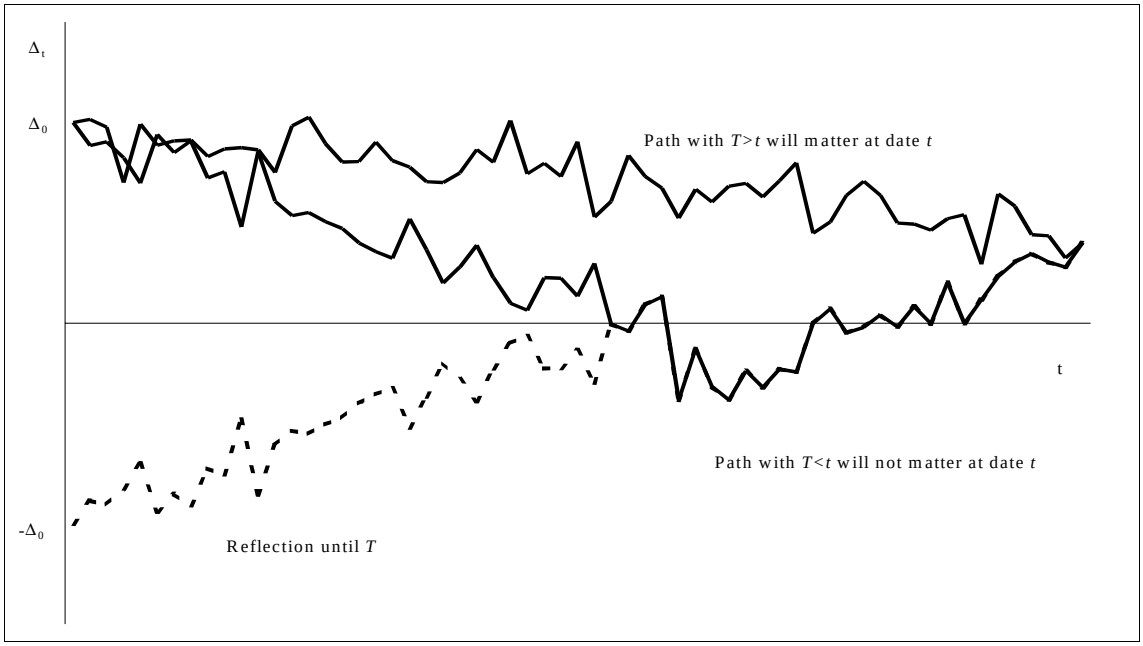

Figure 2: A comparison of the learning and random growth model ($\Delta_{\text{random growth}} = 4$, $\Delta_{\text{learning}} / \sqrt{\delta} = 3.8$, $\delta = 0.01$)

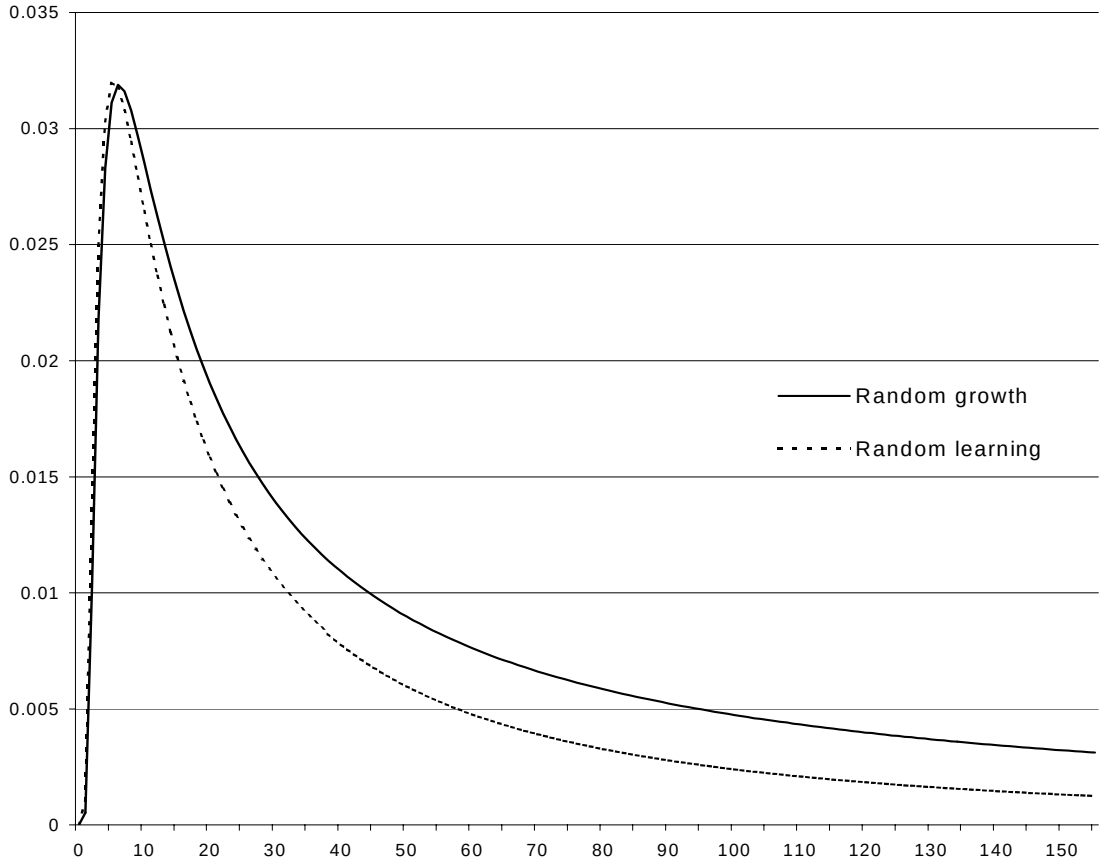

Figure 3: Weekly job exit percentage for the first year

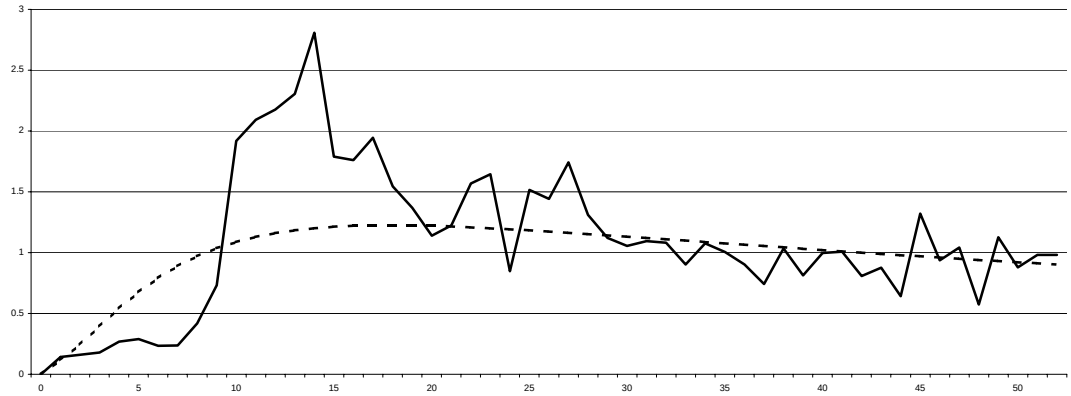

Figure 4: Monthly job exit percentage for the first fourteen years

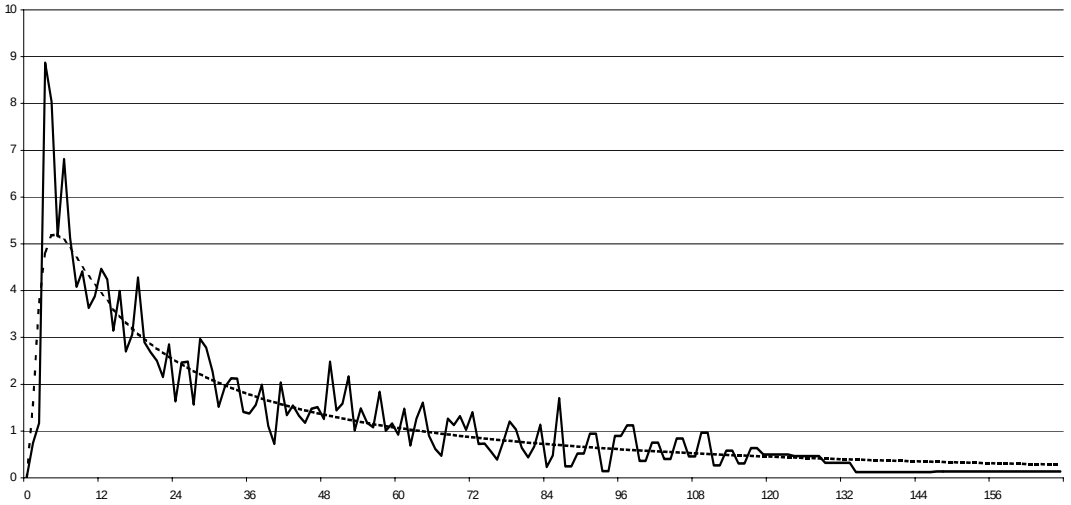
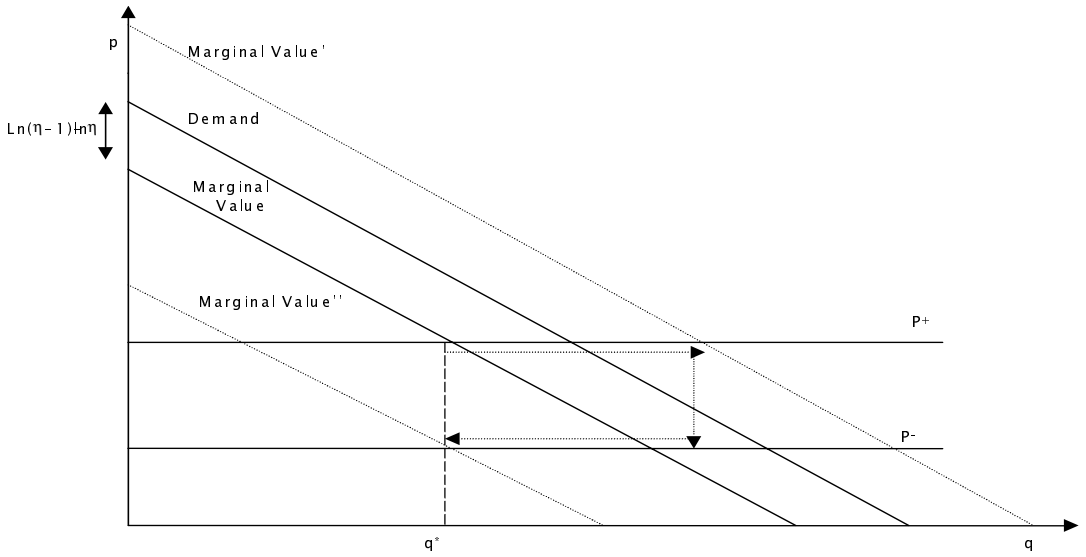
Figure 5: The relation between the individual and the firm level