

Returns to Tenure or Seniority? *

Sebastian Buhai, Miguel Portela, Coen Teulings and Aico van Vuuren[†]

This version: 27 June 2008

Abstract

This study documents two empirical regularities, using data for Denmark and Portugal. First, workers who are hired last, are the first to leave the firm (Last In, First Out; LIFO). Second, workers' wages rise with seniority (= a worker's tenure *relative* to the tenure of her colleagues). We seek to explain these regularities by developing a dynamic model of the firm with stochastic product demand and irreversible specific investments. There is wage bargaining between a worker and its firm. Separations (quits or layoffs) obey the LIFO rule and bargaining is efficient (a zero surplus at the moment of separation). The LIFO rule provides a stronger bargaining position for senior workers, leading to a return to seniority in wages. Efficiency in hiring requires the workers' bargaining power to be in line with their share in the cost of specific investment.

Keywords: irreversible investment, options, seniority, LIFO, matched employer-employee data

JEL-codes: J31, J41, J63

*We are grateful for insightful comments and suggestions to Marianne Bertrand, Paul Bingley, Pieter Gautier, Juergen Maurer, Espen Moen, Dale Mortensen, Kevin Murphy, Jean Marc Robin, Robert Topel, Valerie Smeets, Lars Villhuber and participants in seminars and conferences at LMDG in Sandbjerg, EALE in Oslo, ESEM in Budapest, University of Mannheim, Aarhus School of Business, ESPE in Verona, SOLE in Boston and CAFE in Nuremberg. All remaining mistakes are ours.

[†]Buhai (corresponding author): Aarhus School of Business, University of Aarhus and Tinbergen Institute/Erasmus Univ. Rotterdam, sebu@asb.dk; Portela: University of Minho, mangelo@eeg.uminho.pt; Teulings: CPB Netherlands Bureau for Economic Analysis, Univ. of Amsterdam and Tinbergen Institute, C.N.Teulings@cpb.nl; Van Vuuren: Free University Amsterdam and Tinbergen Institute, avuuren@feweb.vu.nl

1 Introduction

Why does Lars earn a lower wage than Jens, while they both do exactly the same job, at the same firm, and with equal skills? And why is Pedro fired and his colleague Miguel allowed to stay at the firm, when their employer has to scale down employment, where again they do the same job? Some might think that the answer to these questions is obvious: it is simply because Jens and Miguel have a longer tenure at the firm than Lars and respectively Pedro. This paper provides a simultaneous explanation and offers empirical evidence for these phenomena. Using matched worker-firm data for Denmark and Portugal, we show that a worker who is hired last, is likely to be fired first (Last In, First Out; LIFO). Analogously, we show that there is return to *seniority* in wages. In both cases, our claims are different from saying that your *tenure* at the job affects negatively your job exit hazard or affects positively your wage. Seniority is different from tenure in that it measures the worker's tenure *relative* to the tenure of her colleagues. Your seniority is your rank in the tenure hierarchy of the firm. Hence, we need all-encompassing matched worker-firm data to establish a worker's seniority because we need to know the tenure of all the firm's workers. Thus, when we claim that seniority affects your separation risk, we mean that on top of the negative duration dependence of the hazard rate, being a senior worker with many more junior colleagues has a further negative effect. Similarly, when we claim that there is a return to seniority in wages, we mean that on top of the return to tenure as usually measured, there is return to seniority. To the best of our knowledge, this paper is the first to document the existence of a return to seniority in wages.

We devise a simple economic theory for why firms and workers would agree on applying a LIFO layoff rule and why that leads to a return to seniority in wages. Our theory is based on a dynamic model of the firm with stochastic product demand and irreversible specific investments for each newly hired worker, similar to Bentolila and Bertola (1990). Dixit (1989) considers the same model, but then for an individual worker. Labor demand follows a geometric random walk in these models. Bentolila and Bertola calculate the optimal hiring and firing points, by considering, for the current employment level, the expected discounted marginal revenue of hiring an additional worker, accounting for the expected moment when it is efficient to fire that worker, taking as given all workers currently employed by the firm and disregarding any workers that might be hired in the future. In this way, the hiring and firing of each worker can be considered separately of the hiring and firing of all other workers, transforming a firm level model into a model of an individual worker, as in Dixit (1989). This turns out to be equivalent to applying a LIFO separation rule. Whereas Bentolila and Bertola (1990) and Dixit (1989) take wages as given, we allow for wage bargaining over the surplus generated by the specific investment. Here, we apply an idea developed by Kuhn (1988) and Kuhn and Robert (1989). Consider the standard monopoly union model, where the union bargains for wages above the market wage and the firm reduces its labor demand below the efficient level in response to this higher wage rate. Kuhn and Robert observe that there is an alternative way for

workers to extract rents from the firm. Their idea is to bargain for a layoff order and for a wage schedule where inframarginal workers get higher wages than marginal workers. The firm cannot fire the expensive inframarginal workers without first firing the cost effective marginal workers. When this wage schedule is properly set, the firm will pick the efficient employment level. As a consequence of this setup, equally productive workers receive different wages, based only on their position in the layoff order, just like Lars and Jens in the opening sentence of this paper. Kuhn and Robert elaborate their ideas in a static framework. Instead, we introduce them in the dynamic model of Bentolila and Bertola, leading to a return to seniority in wages. We take an eclectic approach by simply positing a log linear sharing rule of the surplus of the specific investment. However, we impose one feature that characterizes Nash bargaining, namely efficient bargaining: as long as there is a surplus, the worker and the firm will agree on a distribution of that surplus that makes continuation of the relation mutually beneficial. This guarantees that there is always efficiency on the firing side. Nevertheless, hiring decisions are efficient if and only if cost and revenues of the specific investment are shared in the same proportions. If not, hiring is below the efficient level due to a hold up problem. We shall refer to this condition as the Hosios condition, like in the search theory. We elaborate our model under the assumption that the firm must pay for the full cost of the specific investment, so that any return to seniority implies sub-efficient hiring. Under risk neutrality, contractibility of either specific investment or wages suffices to achieve efficiency, since we can always satisfy the Hosios condition by using the one to match the other. When workers are risk averse, any return to seniority is inefficient, as it assigns the worker a risky return that can better be assigned to the risk neutral firm. As an extension, we consider the effect of firing cost, accounting for its upward effect on wages.¹ By the efficient bargaining assumption, firing cost does not affect firing, but further deteriorates hiring. Finally, we consider the role of trade unions in this model. At first sight, the ideas in Kuhn (1988) and Kuhn and Robert (1989) seem to suggest the return to tenure to be higher in unionized firms, since unions are predicted to use the tenure profile as a rent extraction mechanism. This turned out to be counter-factual: unionized firms generally have a lower return to tenure, not a higher return, see for instance Teulings and Hartog (1998: 225). We observe that this fits our theory. The LIFO layoff rule allows for a decentralisation of the bargaining process –as required in the absence of a union– leading to higher wages for senior workers. Instead, the political process within a union would lead to a more egalitarian distribution of the rents among the workers, that is, to higher wages but a lower wage return to seniority.

In the empirical part, we establish a number of features of our model. We show that seniority is an important determinant of job separation. Junior workers have a larger separation probability than senior workers. This effect comes on top of the duration dependence of the hazard, that is, in addition to the fact that the

¹In Bentolila and Bertola's (1990) analysis of the effect of firing costs, wages are fixed. Accounting for the effect on wage setting turns out to be important for the conclusions.

separation probability declines with the elapsed tenure at the job. Second, we show that there is a wage return to seniority. Starting from the seminal papers by Altonji and Shakotko (1987) and Topel (1991), there is a large and still flourishing literature on the estimation of the wage return to tenure. The problem in this literature is that within a job spell, tenure is perfectly correlated with experience. Hence, the first order term of this return can only be estimated using variation between job spells, but that introduces all kind of selectivity problems, which this literature sets out to resolve. We show that this problem is absent in the estimation of the return to seniority, since seniority is not perfectly correlated with experience. Seniority increases for example because new workers enter the firm. From that perspective, changes in seniority are correlated with changes in firm size, though not perfectly. In our regressions, we use within job spell variation and we include both tenure and firm size as controls; nevertheless, we are still able to find wage returns to seniority of 1 to 2 % in Portugal, and returns half that range in Denmark.

The paper is set up as follows. Section 2 presents our theoretical framework. In Section 3, we describe the data and the relevant labour market institutions in Denmark and Portugal, and we present our estimation results. Section 4 summarizes and concludes.

2 Theoretical framework

2.1 Setup

The model of Bentalila and Bertola (1990) provides a nice starting point for our analysis. Firms face a stochastic iso-elastic demand curve for their output, in logs:

$$n_t = z_t - \eta p_t, \tag{1}$$

where $\eta > 1$ is the price elasticity of demand, N_t is demand, P_t is its price; lower cases denote the log of the corresponding upper cases, so n_t is log output. The variable z_t is a market index capturing the exogenous evolution of demand; z_t is assumed to follow a Brownian with drift, such that $\Delta z \sim N(\mu, \sigma^2)$. Labor is the only factor of production. The production function exhibits constant returns to scale. Without loss of generality, productivity is normalized to unity, so that output is equal to employment. In the model of Bentalila and Bertola (1990), hiring and firing of workers is costly. At this stage, we focus on hiring cost, denoted by I . This cost is interpreted more broadly as the specific investment that has to be made by the firm at the start of an employment relation. It is irreversible: once made, the cost cannot be recouped by ending the employment relation. For simplicity, we assume that this investment can be made instantaneously, so that no time elapses between the start and the end of the investment process. At the outside market, workers can earn a reservation wage, which is constant over time. It is most convenient to think of this reservation wage as the return to self employment. Without loss of generality, it is normalized to unity: $w^r = 0$, where w^r denotes the log reservation wage. We assume

both workers and firms to be risk neutral.

As a benchmark, we first analyze the simple case where firms pay workers their reservation wage and where there are no specific investments required for starting an employment relation, $I = 0$. In that case, labor demand can be adjusted costlessly at each point in time. Hence, the optimal strategy is to maximize instantaneous profits Π_t :

$$\Pi_t = N_t P_t - N_t w_t,$$

subject to the demand curve (1). The first term is total revenue, the second term is the wage bill. The first order condition for profit maximization implies

$$\begin{aligned} p_t &= \pi, \\ n_t &= z_t - \eta\pi, \\ \pi &\equiv \ln \frac{\eta}{\eta - 1} > 0. \end{aligned} \tag{2}$$

This ratio is greater than unity due to the monopoly power of the firm at the product market. The firm's price is constant over time, while its labor demand follows a random walk, a regularity known as Gibrat's law, which holds for larger firms, see for instance Jovanovic (1982).

Next, consider the optimal strategy with specific investments, $I > 0$. Then, labor demand cannot be adjusted costlessly. On the hiring side, an additional worker requires a specific investment, which has to be recouped from future profits. Moreover, this investment is irreversible, so that delaying hiring has an option value. On the firing side, firing per se is costless, but irreversible. If demand surges after having fired the worker, the firm is unable to benefit from that demand without incurring the cost of the specific investment again. Hence, retaining the worker has an option value, too. Bentolila and Bertola (1990) show that the optimal policy of a firm is to hire workers whenever p_t reaches a constant upper bound $p^+ > \pi$ and to fire them whenever p_t reaches a lower bound $p^- < \pi$.

The situation is sketched in Figure 1. The hiring bound p^+ exceeds π due to the necessity for the firm to recoup the cost of specific investments and due to the option value of postponing hiring, while the firing bound p^- is below π due to the option value of postponing firing. The downward sloped curves are the logarithm of marginal revenue $mr(\cdot)$ as a function of log employment n_t and the market index z_t , that is

$$\ln [d(N \cdot P) / dN] \equiv mr(n, z) = \frac{1}{\eta} (z - n) - \pi.$$

The current value of the market index is denoted by z_0 in Figure 1. In the case $I = 0$, the firm sets equal log marginal revenue $mr(n_0, z_0)$ to log marginal cost $w_r = 0$. Hence: $n_0 = z_0 - \eta\pi$. There is a wedge π between p_0 and $mr(n_0, z_0)$ due to the monopoly power of the firm. Any change in z_t will immediately affect

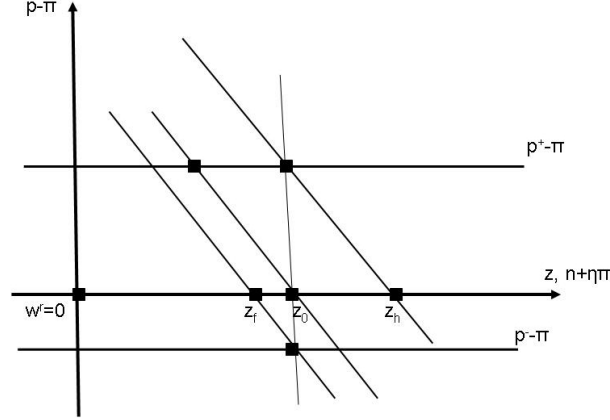


Figure 1: Firing-hiring boundaries with stochastic market index

log employment n_t , but will leave the log price constant at π . In the case $I > 0$, employment is insulated from shocks to the market index z_t within certain bounds. Only when z_t rises above z_h , the firm hires additional workers to prevent p rising above p^+ , and only when the market index falls below z_f , the firm fires workers to avoid p falling below p^- . Hence, p_t follows a random walk between p^- and p^+ , while n_t is constant at n_0 in this interval. However, when p_t drifts outside these boundaries, the firm uses n_t as an instrument to control p_t . Then, p_t is held constant, and n_t starts drifting, either up (if $p = p^+$), or down (if $p = p^-$). We provide expressions for p^+ and p^- later on.

2.2 The LIFO rule and rent sharing

For the moment, let us suppose that firms are obliged to apply a LIFO (Last-In-First-Out) separation rule. Later on, we shall offer a rationale why a firm and its workers would agree to using such a rule. We assume that workers never quit or retire. We can index each worker by the log employment level of the firm at the date that the worker is hired: a worker hired at time h gets rank q , $q = n_h = z_h - \eta p^+$. Her seniority index at time t is defined as $n_t - q$. Hence, the most senior worker has $q = 0$, and her seniority index is $n_t - q = n_t$, while the least senior worker at time t has $q = n_t$, and her seniority index is $n_t - q = 0$. The LIFO layoff rule says that when a firm wants to fire a worker, it has to fire the worker with the lowest seniority $n_t - q$. Under this separation rule, the distribution of completed job tenures is characterized by the following proposition.

Proposition 1 *Consider a firm that satisfies the previous assumptions. The distribution of the completed tenure of workers hired by this firm is equal to the first passage time distribution, that is, the distribution of the time it takes a random walk z , $\Delta z \sim N(\mu, \sigma^2)$, with initial value z_h , to pass the barrier $z_h - \eta(p^+ - p^-)$ for the first time. This distribution does not depend on the initial value z_h , and hence, it is identical for all workers hired by the firm, irrespective of the number of workers hired previously.*

Proof: The LIFO layoff rule implies that a worker hired at time h , with $q = n_h$, will be fired at the first moment $f > h$ that employment is back at the level n_h and $p_f = p^-$, since this worker can only be fired when all workers hired after, at $t > h$, have been fired, since their rank is $q > n_h$, whereas all workers that have been hired before, at $t < h$, cannot be fired before the worker hired at time h is fired. Hence, $z_h - z_f = \eta(p^+ - p^-)$ and $z_h - z_t = \eta(p^+ - p^-)$, $\forall t \in (f, h)$. The distribution of the first moment where $z_h - z_f = \eta(p^+ - p^-)$ is independent of the starting value z_h of the random walk at time h , due to the Markov property of a random walk. ■

Bentolila and Bertola's (1990) model of firm level employment supplemented with a LIFO layoff rule corresponds therefore one-to-one with a simple model of individual job tenures. Buhai and Teulings (2006) analyse the characteristics of the distribution of completed tenures implied by this model: its hazard rate starts at zero, then rises quickly to a peak, and then falls slowly, to zero for $\mu > 0$, and to some positive number for $\mu < 0$. Buhai and Teulings estimate this model on tenure data for the United States, and show that it fits the data well.

Why would a firm commit to using a LIFO layoff rule? In the simple world discussed above, where the firm pays the worker her reservation wage, there is no rationale for such a rule. Since the worker receives her reservation wage, she is indifferent between working at the firm or being laid off. Hence, there is no point in fixing an order of layoff. However, if we relax the assumption that the firm pays its workers their reservation wage and we attribute incumbent workers some bargaining power, the quasi rents of the specific investment might enable these workers to capture wages above the reservation wage. In that case, a layoff order carries practical relevance, as it protects the 'rights' of senior workers. Kuhn (1988) and Kuhn and Robert (1989) offer a neat further legitimation for using such a rule. Consider the standard monopoly union model, where the union bargains for union wage rate above the workers' reservation wage. The firm reduces its labor demand below the efficient level in response to this higher wage rate. This leaves some gains from trade unexploited, since there are workers who would be willing to work at the firm for the reservation wage and the firm would be willing to hire them, but these trades do not occur because the firm is not allowed to pay wages below union rate. Kuhn and Robert observe that there is an alternative way for workers to extract rents from the firm without leaving gains from trade unexploited. The idea is to bargain simultaneously for a LIFO layoff rule and a wage schedule that grants higher wages to inframarginal workers (those who

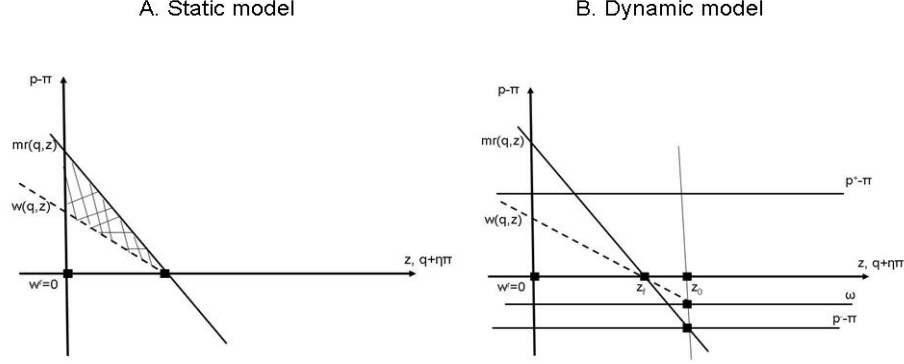


Figure 2: Static vs. dynamic framework

were hired first, and, by the LIFO rule, are therefore the last to be fired) than to marginal workers. Hence, log wages depend positively on the seniority index $n_t - q$. We assume this relation to be linear:

$$w(q, z_t) = \beta \cdot mr(q, z_t) + \omega = \frac{\beta}{\eta} (z_t - q) - \beta\pi + \omega, \quad (3)$$

with $0 \leq \beta \leq 1$.

The situation is depicted in Figure 2, Panel A; the continuous line is marginal revenue $mr(q, z_0)$, the dotted line is the wage schedule $w(q, z_0)$. The shaded area is the surplus of marginal revenue above marginal cost for that level of log employment, q . As long as the log wage schedule is between the marginal revenue curve $mr(q, z_0)$ and the log reservation wage $w^r = \omega = 0$, the firm has no incentive to deviate from the first best level of log employment, n_0 , since the surface of the shaded area is maximized by setting $n = n_0$. In that sense, equation (3) implies efficient bargaining: all gains from trade are exploited. Although inframarginal workers earn a log wage $w(q, z)$, $q < n_0$ above the log marginal revenue $mr(n_0, z_0)$, the firm has no incentive to fire them, because it is obliged to fire the less expensive marginal workers first. This wage setting scheme is a form of price discrimination on the side of the union. The parameter β can be interpreted as the bargaining power of workers.² The case $\beta = 1$

²Strictly speaking, this interpretation lacks a foundation in a formal bargaining model. In the case of a single worker firm, where we could apply the theory of two player bargaining, as in Buhai

corresponds to first degree price discrimination, where the union has full bargaining power. In case $\beta = 0$, workers have no bargaining power and they get just their reservation wage. The return to seniority, β/η , is increasing in the bargaining power of the workers, β , and in the monopoly power of the firm, η^{-1} . Since $0 \leq \beta \leq 1$ and $\eta > 1$, this return is between zero and unity.

Figure 2, Panel A, corresponds to the original model of Kuhn and Robert (1989), who specify their theory in static framework. In the dynamic framework à la Bentolila and Bertola (1990) that we apply here, there are some complications, see Figure 2, Panel B. Since investment I is irreversible, workers have an option value for the firm. Hence, the firm keeps workers when marginal revenue has fallen just slightly below marginal cost, that is, when $w^r > mr(n_0, z_0) > p^- - \pi$. Similarly, the marginal worker, $q = n_0$, prefers staying employed at the firm even when her wage rate is just slightly below the reservation wage, $w^r < w(n_0, z_0) < \omega$, since by quitting the worker loses the option value of benefitting from the market index z_t rising above z_0 at later stage, and hence log wages rising above w^r . Hence, for some range of $z \in [z_f, z_0]$, marginal revenue is below marginal cost, $mr(n_0, z) < w^r$, and wages are below reservation wages, $w(n_0, z) < w^r$. Hence, whereas w^r is the outside wage for a worker, ω is the reservation wage for a worker currently employed by the firm. Only when $w(n_0, z_0)$ is below ω , $\omega < w^r$, the marginal worker (for whom $q = n_0$) wants to quit the firm, because the prospect of future wage increases no longer outweighs the cost of the current wage being below the outside wage.

Since Kuhn and Robert (1989) specify their theory in a static framework, the layoff ordering can be based on any variable, height, IQ, experience, or what else springs to mind. When using the dynamic framework of Bentolila and Bertola (1990), ordering layoffs by a LIFO rule has a natural economic rationale. The senior worker's future wage claims are sensitive to the firm hiring new workers, since after the specific investments have been made, these new workers are perfect substitutes for senior workers. The firm could in principle hire new workers for a low wage, and fire the senior workers instead. Hence, senior workers have an incentive to block the hiring of new workers. Suppose that the specific investment of new workers is largely made up from acquiring the tacit knowledge of the firm's production process and the transfer of this knowledge is a monopoly of senior workers, or suppose that senior workers can harass newcomers, as suggested by Lindbeck and Snower (1990). In that case, hiring new workers requires the consent of senior workers. So, the firm has an incentive to commit on not using new hires as substitutes for senior workers. A LIFO separation rule serves as a commitment device, by providing senior workers protection against being laid off before newly hired workers.

and Teulings (2006), the log linear sharing rule would be almost equivalent to Nash bargaining, which would yield a linear sharing rule.

2.3 The worker's problem

The value of the parameter ω in equation (3) can be derived using the theory of option values, see Dixit and Pindyck (1994, 136-140)³. Efficient bargaining implies that at the moment of separation, both the worker and the firm are indifferent between continuation of the employment relation. In this section, we analyse the worker's side of the problem, which yields an expression for ω . The next section looks at the firm's side of the problem, which yields expressions for the hiring and firing thresholds, p^- and p^+ . First, we observe that $w(q, z_t)$ can be written as function of a single argument, $z_t - q$, see equation (3), as we do in what follows. Let $V(z_t - q)$ be the asset value of holding a job at a firm. By Ito's lemma $V(z_t - q)$ satisfies the Bellman equation

$$\rho V(z_t - q) = \exp[w(z_t - q)] + \mu V'(z_t - q) + \frac{1}{2}\sigma^2 V''(z_t - q),$$

where ρ is the interest rate. For a bounded solution to exist, we must assume that $\rho > \mu + \frac{1}{2}\sigma^2$. The left hand side is the return on the asset. The first term on the right hand side, $\exp[w(z_t - q)]$, is the current wage income, the second and the third term are the (expected) wealth effects of the drift and the volatility in the market index z_t . The solution to this second order differential equation reads

$$\begin{aligned} V(z_t - q) &= \frac{1}{r(\beta/\eta)} \exp\left[\omega + \frac{\beta}{\eta}(z_t - q - \eta p^-)\right] + A^- \exp[\lambda^-(z_t - q)] \\ &\quad + A^+ \exp[\lambda^+(z_t - q)] \end{aligned} \quad (4)$$

$$r(x) \equiv \rho - \mu x - \frac{1}{2}(\sigma x)^2,$$

$$\lambda^{+,-} \equiv \frac{-\mu \pm \sqrt{\mu^2 + 2\rho\sigma^2}}{\sigma^2}.$$

where we substitute $w(z_t - q)$ from equation(3) and where A^- and A^+ are constants of integration that remain to be determined. $r(x)$ is a modified discount rate, accounting for the drift and the variability of z_t . We have:

$$\begin{aligned} r(0) &= \rho, r(1) = \rho - \mu - \frac{1}{2}\sigma^2 > 0, \\ \lambda^- &< 0, \lambda^+ > 1. \end{aligned}$$

where the final inequality follows from $\rho > \mu + \frac{1}{2}\sigma^2$. The first term of the right hand side of equation (4) is the net discounted value of expected wage payments, disregarding the worker's option to quit the firm when wages fall too far below the

³Dixit and Pindyck (1994) consider an optimal investment problem instead of an optimal stopping problem, but this is *mutatis mutandis* the same. Their notation is somewhat different in that their z denotes a geometric Brownian instead of standard Brownian. The equations are exactly the same, but we feel our notation to be simpler.

reservation wage. The final two terms, $A^{+,-} \exp [\lambda^{+,-} z_t]$, are the option value of separation. Only one of the roots $\lambda^{+,-}$ is relevant, due to a transversality condition. For large values of z_t , the firm is doing well and hence, keeping the job is attractive to the worker for the foreseeable future. The option value of separation must converge to zero, which is the case for the negative root λ^- , since $\lim_{z \rightarrow \infty} A^- \exp [\lambda^- z] = 0$. Hence, the constant of integration A^+ must be equal to zero.

Efficient bargaining implies that it is optimal for a worker with rank q to separate when $z_t = q + \eta p^-$. Two conditions need to hold for that value of z_t to be optimal:

$$\begin{aligned} V(z_t - q) &= \frac{1}{r(\beta/\eta)} e^\omega + A^- e^{\lambda^- z} = \frac{1}{\rho}, \\ V'(z_t - q) &= \frac{\beta}{\eta r(\beta/\eta)} e^\omega + \lambda^- A^- e^{\lambda^- z} = 0. \end{aligned}$$

The first condition is the value matching condition, which states that the asset value of holding the job should be equal to the asset value after separation, that is, the net discounted value of the reservation wage, ρ^{-1} . The second condition is the smooth pasting condition, which states that for small variations in z_t the worker remains indifferent between holding the job and separation. Since separation is irreversible, the worker should not regret separating after a small perturbation of z_t . This requires the first derivative of $V(z_t - q)$ with respect to z_t to be zero. Elimination of A^- yields an expression for ω :

$$\omega = \ln r(\beta/\eta) - \ln \rho - \ln \left(1 - \frac{\beta}{\eta \lambda^-} \right). \quad (5)$$

Proposition 2 (i) $\omega \leq 0$; (ii) $\frac{\partial \omega}{\partial \beta} < 0$; (iii) $\frac{\partial \omega}{\partial \mu} < 0$; (iv) $\frac{\partial \omega}{\partial \sigma} < 0$.

Proof: The proof is straightforward calculus. ■

ω is below the log reservation wage $w^r = 0$ since separation is an irreversible decision. If the demand for the firm's product, z_t , goes up after the separation decision, the worker is no longer able to benefit from the wage increase. Hence, workers are prepared to incur some loss before they decide to separate. The higher the worker's bargaining power β , the lower is ω , since expected future revenues are higher so that workers are prepared to accept greater losses before separation. Similarly, ω is declining in the drift μ since a higher drift raises expected future revenues, and ω is declining in the variability of demand σ^2 , since a higher variability raises the option value of hoping for a future increase in the surplus.

2.4 The firm's problem

We now turn to the firm's optimal strategy. Like $w(q - z_t)$, $mr(q, z_t)$ can be written as a function of $q - z_t$ only, as we do in what follows. We observe that hiring

a worker with rank q^* affects neither log marginal productivity⁴ $mr(q - z_t)$ nor log wage $w(q - z_t)$, for workers with rank $q < q^*$ (the workers who have been hired before worker q^*). Similarly, whether or not the firm hires any workers $q > q^*$ after hiring worker q^* affects neither $mr(q^* - z_t)$ nor $w(q^*)$. Furthermore, the option of firing worker q^* at a future date is unaffected by the hiring of workers $q > q^*$. Hence, in its cost calculation, the firm can attribute each worker her marginal revenue $\exp[mr(q - z_t)]$ and her wage $\exp[w(q - z_t)]$, taking the employment of workers hired previously as given, and then consider when it is optimal to hire and subsequently fire this worker. In this way, we can consider the decision to hire and fire the N_t -th worker ($N_t \equiv \exp n_t$) separately of the hiring and firing of workers hired before this worker, and of workers hired afterwards. Then, the model is a straightforward extension of Dixit and Pindyck (1994: 216), the only difference being that wages are constant in Dixit and Pindyck, while they vary with the state of demand z_t in this model. Let $F(n_t - z_t)$ be the asset value of the firm for the N_t -th worker. The Bellman equation for $F(n_t - z_t)$ satisfies

$$\rho F(n_t - z_t) = \exp[mr(z_t - n_t)] - \exp[w(z_t - n_t)] + \mu F'(n_t - z_t) + \frac{1}{2} \sigma^2 F''(n_t - z_t).$$

The first term is the marginal revenue of that worker, the second term is the wage for that worker. The final two terms capture the option value of separation. The relevant solution to this differential equation reads

$$F(n_t - z_t) = \frac{1}{r(\eta^{-1})} \exp\left[\frac{1}{\eta}(z_t - n_t) - \pi\right] - \frac{1}{r(\beta/\eta)} \exp[w(z_t - n_t)] + B^- \exp[\lambda^-(z_t - n_t)]. \quad (6)$$

The final term is the option value of separation, with B^- being the constant of integration. As in the case of the worker, the positive root λ^+ is irrelevant, due to a transversality condition: the option value converges to zero for large values of z_t . Suppose the firm employs less than N_t workers. Then, the option value of hiring the N_t -th worker at some future date is

$$G(n_t - z_t) = B^+ \exp[\lambda^+(z_t - n_t)], \quad (7)$$

where B^+ is the constant of integration. There are no current costs or revenues, hence only the option value term matters. Since this option value converges to zero for low values of z_t , only the positive root λ^+ applies here. The value matching and smooth

⁴Note that the function $mr(q, z_t)$ measures the marginal revenue of adding the q -th worker, or equivalently, it is the firm's marginal revenue *as if* employment were only q . When actual employment is larger, $n_t > q$, the actual marginal revenue of the firm is smaller, see Figure 2, which is drawn for $t = 0$.

pasting conditions read

$$\begin{aligned}
F(z_t - \eta p^-, z_t) &= G(z_t - \eta p^-, z_t), \\
F'(z_t - \eta p^-, z_t) &= G'(z_t - \eta p^-, z_t), \\
F(z_t - \eta p^+, z_t) &= G(z_t - \eta p^+, z_t) + I, \\
F'(z_t - \eta p^+, z_t) &= G'(z_t - \eta p^+, z_t).
\end{aligned} \tag{8}$$

The first pair refers to the firing decision, the second to the hiring decision. The first condition states that at the moment of firing, when by definition $n_t = z_t - \eta p^-$, the asset value of keeping the worker is equal to the option value of a vacancy. The second equation is the smooth pasting condition, which states that this condition also applies for slight variations of z_t , so that the firm wouldn't regret a decision to fire after a slight variation in z_t . The third equation is the value matching condition for the moment of hiring, when $n_t = z_t - \eta p^+$: the asset value of hiring the worker should be equal to the cost of investment plus the option value of filling the vacancy at a later point in time. The final equation is the smooth pasting conditions for the moment of hiring. This system of four equations determines four unknowns, the constants of integration, B^- and B^+ , and the hiring and firing boundaries, p^- and p^+ .

Proposition 3 *The system of equations (8) has a unique solution for p^+, p^-, B^+, B^- for which (i) $p^+ - \pi > 0 > \omega > p^- - \pi, B^+ > 0, B^- > 0$; (ii) $\frac{\partial B^+}{\partial \beta} < 0$; (iii) $\frac{\partial p^+}{\partial \beta} > 0$; (iv) $\frac{\partial p^-}{\partial \beta} < 0$.*

Proof: see Appendix. ■

A unique, economically meaningful solution exists, see Proposition 3-(i). In the static version of the model by Kuhn and Robert (1989), the distinction between hiring and firing is meaningless, so that there is only a single issue of efficiency, namely the level of employment. In the dynamic version of the model, the efficiency of hiring and firing are two separate issues. Since the asset value for outside workers is equal to the net discounted value of their reservation wage ρ^{-1} , the option value of a firm to hire an outside worker, on its current level of log employment n_t and the value of its market index z_t , is the residual claim in this economy. Any inefficiency shows up as negative effect on this option value. This option value is proportional to B^+ , see equation (7). Hence, Proposition 3-(ii) implies that $\beta = 0$ is most efficient, since then B^+ is at its maximum. This result is due to a hold up problem. For efficiency, the Hosios (1990) condition should apply: the revenues of specific investment should be shared between players according to their share in the cost of this investment. Since firms are assumed to bear the full cost I , any share of its revenues being assigned to workers leads to inefficiency. Hence, $\beta = 0$ is most efficient. If $\beta > 0$, firms will postpone the hiring of new workers till the surplus of log marginal productivity $mr(q - z_t)$ above the log reservation wage w^r is such that, even though the firm gets only part of that surplus, its net discounted value is sufficient to cover the cost of investment. This

explains why p^+ is positively related to β , see Proposition 3-(iii). We now turn to the firing decision. The elimination of B^- from the first two equations of (8) yields

$$p^- - \pi = \ln r (\eta^{-1}) - \ln \rho - \ln \left(1 - \frac{1}{\eta \lambda^-} \right) + \ln \left(1 - \rho \frac{\lambda^+ - \lambda^-}{\lambda^-} B^+ \exp [\eta \lambda^+ p^-] \right). \quad (9)$$

The firing bound p^- does not depend on β , except for its effect on B^+ . This is an application of the Coase theorem: under efficient bargaining, the distribution of the surplus of the employment relation does not matter for the actual level of employment. The only exception is the option value of hiring another worker for this vacancy at a later stage. This option value comes in because at the same time that the firm fires the N_t -th worker, it acquires the option to rehire at a later stage, provided that it pays the cost I again. This option makes firing more attractive, so it raises p^- . The larger is β , the lower is this option value of future rehiring, and the more attractive it is to fire a worker, see Proposition 3-(iv). Proposition 3-(iii), (iv) implies that the higher is the workers' bargaining power β , the less volatile is the employment, since employment is insulated from shocks to market index z_t over a larger interval of $\eta p^- < z_t - n_t < \eta p^+$, and the larger is therefore the expected tenure of a newly hired worker. These implications square well with the findings in Bertrand and Mullainathan (2003), who show that when firms are insulated from takeovers, the wages of the incumbent employees are higher, suggesting a higher value of β . This goes hand in hand with lower rates of creation of new plants, which in the context of our model is similar to a higher hiring bound, p^+ . Bertrand and Mullainathan also report a lower rate of destruction of old plants, or in the context of our model, a lower firing bound, p^- .

2.5 Unemployment and the welfare cost of hold up

To close the model, we have to explain who gets hired by a firm and who does not. The log wage of a worker who is just hired is higher than the wage of a worker who is at the borderline of being laid off, that is

$$w(z_t - \eta p^+, z_t) > w(z_t - \eta p^-, z_t) = \omega.$$

Since the asset value of a worker who is on the borderline of being laid off is equal to the net present value of her reservation wage, $1/\rho$, the asset value of a worker who is just hired must be higher than $1/\rho$. Hence, new jobs at the firm are rationed. A convenient way to model this rationing process is to introduce unemployment. A worker who is just laid off has two options. Either she decides to collect her outside wage by becoming self employed, or she can decide to queue for a new job at a firm. During this waiting period she cannot produce as a self employed. For simplicity, we assume that leisure has no value.⁵ New jobs at firms arrive at a rate λ per unit of the

⁵Allowing for a value of leisure would not change the predictions of model. It would make unemployment less costly per unit of time, but this effect would be exactly offset by the rise in unemployment.

labor force and are distributed randomly among the unemployed. Hence, the asset value of unemployment, V^U , satisfies

$$\rho V^U = \frac{\lambda}{u} [V(\eta p^+) - V^U],$$

where u is the unemployment rate. λ/u is the arrival rate of a new job for unemployed. The lower unemployment, the higher this arrival rate, since there are less people among whom new jobs have to be distributed. $V(\eta p^+) - V^U$ is the asset gain of getting a job offer. The level of unemployment follows from the no-arbitrage condition between self employment and unemployment

$$u = \lambda \left[V(\eta p^+) - \frac{1}{\rho} \right], \quad (10)$$

where we use $V^U = 1/\rho$, the asset value of self employment⁶. The higher the asset gain of getting a job at a firm, the higher is unemployment. Hence, there are two types of inefficiency.⁷ First, not all gains from trade between the worker and the firm are exploited. Firms would hire more workers if $\beta = 0$, since p^+ is an increasing function of β . Firms' perception of the marginal cost of hiring a worker in net present value terms exceeds the social cost by the same amount as the asset gain for an unemployed of getting a job offer, $V(\eta p^+) - 1/\rho$. This gives rise to a Harberger triangle. Next to this Harberger triangle, there are the cost of rationing that dissipate workers' surplus. The no-arbitrage condition (10) implies that the workers as a group spoil their whole share in the quasi rents in wasteful unemployment.

As discussed before, the inefficiency is due to a violation of the Hosios condition that cost and revenues of specific investments should be shared in the same proportion between workers and firms. At a deeper level, the failure to satisfy the Hosios condition is due to the combination of the non-verifiability of specific investment and the inability of workers to commit on not using their bargaining power after the specific investment has been made. If wages were contractible, workers could commit on not demanding any return to seniority, such that the firm bears the full cost and gets the full revenues of the specific investment, thereby satisfying the Hosios condition. Alternatively, if specific investments were verifiable, the inefficiency would be resolved by shifting some share of the burden of investment to the worker, such that workers bear an equal share of the cost of the specific investment as they get from its revenues, again satisfying the Hosios condition. It is useful to consider this case where workers share in the cost of specific investment, more closely. The asset

⁶We assume: $u < 1$. If $u > 1$, the outside option of self employment would become irrelevant, and the reservation wage ω and job arrival rate λ would become endogenous. ω would rise till so many workers are fired, and so few workers are hired till λ is such that the no arbitrage condition holds for $u = 1$.

⁷Throughout the paper, we do not pay attention to a third type, the inefficiency caused by the monopoly power of the firm vis-a-vis consumers.

value of a worker at the moment of hiring, $V(\eta p^+)$, is independent of the worker's rank q . Hence, although at a particular point in time senior workers receive a higher return on their specific investment than juniors, each worker has the same net present value of expected rents at the moment she is hired, that is, independent of her rank q . Seniors getting higher rents than juniors at a particular point in time reflects the fact that they are able to realize the upside of the risky returns on their share in the cost of specific investment I . Hence, the LIFO separation rule can be interpreted as a protection of their property right of senior workers on their share in the quasi rents, against the temptation of the firm to fire the expensive senior workers, thereby depriving them from the upside of their risky returns. A LIFO separation rule is then a device for implementing an efficient contract.

In the basic model of this paper, workers are overcompensated. However, when workers bear the full cost of the specific investment, the hold up problem is reversed. Then, the non-verifiability of workers investment and the inability of firms to commit on not using their bargaining power $1-\beta$ leads to inefficiency. Workers are only willing to enter the firm when the net present value of quasi rents of their investment are so high that their share in this present value suffices to cover the cost of investment. These arguments imply that as long as we do not know what share of the cost of specific investment is born by workers, empirical evidence showing that there is a return to seniority in wages is inconclusive on the issue of whether or not employment is below its inefficient level. However, there is an alternative statistic enabling the observer to establish which side, either the worker or the firm, is overcompensated in the ex post bargaining over the quasi-rents of the specific investment: when workers queue for jobs, so that there is unemployment, firms are held up, as in the basic model; when firms chase after workers, so that there are vacancies, workers are held up.

2.6 Explanation of the firm size wage effect?

The firm size effect on wages has been extensively documented, see Brown and Medoff (1989). Can our model offer an explanation for the firm size wage effect? When we look at the issue from the point of view of an individual worker, the evolution of her seniority $n_t - q$ is driven by the evolution of log firm size n_t . At first sight, this suggests that our theory could explain the firm size wage effect. This turns out not to be true. The average log wage in a firm at the firing bound p^- satisfies

$$\frac{1}{N_t} \int_{-\infty}^{n_t} w(z_t - q) Q dq = \frac{1}{N_t} \int_{-\infty}^{n_t} [\omega + \beta/\eta (z_t - q + 1 - \eta p^-)] Q dq = \omega + \beta/\eta,$$

where in the first expression the factor Q comes in as the Jacobian $dQ/dq = Q = e^q$, and where in the final equality we use the fact that at the firing bound, $n_t = z_t - \eta p^-$. Hence, the average log wage does not depend on firm size. The intuition is that the positive effect on the average log wage of the wage increase for the incumbents is exactly offset by the negative effect of the below average log wage for new hires. Thus,

although the model predicts firm size to be a driver of wage changes for incumbents, it does not explain why wages for the firm as a whole depend on firm size. The average log wage does depend on the parameter β/η . Other things equal, the model predicts the return to seniority, β/η , to be increasing in the average log wage.

2.7 Extensions

Some extensions to this model are worth discussing. First, relaxing the assumption of risk neutrality on the side of the worker introduces a trade off. As discussed in the previous section, verifiability of the specific investment I is sufficient to implement first best in the standard case with risk neutral workers. With risk averse workers, this conclusion no longer applies. First best requires that workers get paid their reservation wage all the time, and hence that firms bear the full cost of investment. The inability of workers to commit on not using their bargaining power makes first best unattainable in that case. The case of risk averse workers and risk neutral firms is particularly relevant because one can expect capital markets to be more complete for firms than for workers. It is easier for the firm to diversify firm specific risks on the capital market than it is for its workers.

A second extension is the introduction of firing cost, imposed either by law or by trade unions. We think of a firing cost as a wealth transfer L from the firm to the worker at the moment of firing. By the assumption of efficient wage bargaining, this wealth transfer has an impact on the wage bargaining process. Firing cost raise value of the outside option of the worker by the wealth transfer L . The value matching condition reads $V(\eta p^-) = \rho^{-1} + L$. Hence, the expression for ω reads, compare equation (5)

$$\omega = \ln r(\beta/\eta) - \ln \rho - \ln \left(1 - \frac{\beta}{\eta \lambda^-}\right) + \ln(1 + \rho L). \quad (11)$$

Firing cost raises ω . Hence, there are two counteracting effects on the firing bound p^- : the direct effect of firing cost makes layoffs less attractive to the firm, while the indirect effect via higher wages makes layoffs more attractive, since workers are more costly due to the higher level of ω . The first order condition for optimal firing now reads, compare equation (8)

$$F_z(z_t - \eta p^-, z_t) = G(z_t - \eta p^-, z_t) + L,$$

where we use the value of ω from equation (11) to account for the effect of firing cost on wages. Some calculation, see the Appendix, shows that the value for p^- remains the same as in equation (9). The direct and the indirect effect cancel therefore exactly, except for the indirect effect via B^+ , the option value of rehiring. Again, this is an implication of the efficient bargaining assumption and the Coase theorem. On the hiring side, firing cost has two effects with the same sign: first, it raises wages via its effect on ω and, second, there is the prospect of having to pay firing cost in case

of future layoff. To the extent that workers have excessive bargaining power, as in the basic model, this increase comes to the detriment of efficiency. The paradox here is that firing cost aggravates the unemployment problem that it is meant to resolve. Since the hiring threshold p^+ goes up due to the introduction of firing cost, the expected value of future rehiring is lower, so the firing threshold is lower. Hence, firing cost raises the distance $p^+ - p^-$ and therefore the expected tenure of newly hired workers.

Now that we have discussed these extensions, risk aversion and firing cost, it makes sense to consider the nature of workers' bargaining power. Most economists associate this power with the trade unions.⁸ Only unions provide workers bargaining power. Without unions, firms are supposed to have complete bargaining power. A notable exception are Lindbeck and Snower (1990), who point out that the insiders' ability to harass new hires gives them bargaining power vis-à-vis their employer. Without the insiders' consent, firms are effectively unable to train new hires. The interesting aspect of Kuhn (1988) and Kuhn and Robert (1989) is that their rank related compensation scheme allows a decentralization of the bargaining process. As soon as the layoff order has been set, each worker can negotiate for herself. When a marginal worker negotiates a wage increase raising her wage above marginal cost, she endangers her own employment, not that of the inframarginal workers.⁹ Hence, a LIFO scheme enables workers to exploit their individual bargaining power without workers having to solve their collective action problem. Hence, using a LIFO separation rule is a rationale of the firm and its workers on a commitment problem. By using this rule, workers can be sure that training new hires does not threaten their employment at the firm. When workers are united in a trade union, more elaborate strategies are available, that yield a higher expected payoff, in particular when workers are risk averse. By trading a higher firing cost L , in exchange for a lower seniority premium, such that the firm's asset value $F(n_t - z_t)$ remains the same, the expected utility for new hires can be improved by shifting the firm specific risk z_t to the firm. The transfer L allows this insurance to be extended even beyond the time spell covered by the employment relation with the firm, by providing insurance for the fall in wages after being laid off by the firm. Moreover, the political decision making process within the union, where senior and junior workers have to compromise on the distribution of the rents, is likely to generate support for an egalitarian outcome, as implied by the median voter model.¹⁰ All these arguments suggest that the LIFO model is probably a more appropriate description of a non-unionized than of a unionized environment.

⁸This problem has been suggested to us by Kevin Murphy.

⁹The reverse is not necessarily true. An inframarginal worker can bargain a wage above her productivity, if workers with lower seniority capture less than their full productivity. In that case, the firm has an incentive not to fire the inframarginal worker because it first has to fire the marginal worker.

¹⁰Equation (3) implies a Pareto distribution of wages within the firm, which is heavily skewed to the right, so that the median voter has a strong incentive to limit the return to seniority.

These arguments can also explain, why tenure profiles are flatter in unionized firms, in contrast to what Kuhn (1988) and Kuhn and Robert (1989) seem to predict, see e.g. Teulings and Hartog (1998: 225).

3 Empirical framework

The model discussed in the previous section has three testable implications:

1. Gibrat’s law: log firm size follows a random walk, in particular for large firms.
2. The Last-in-First-Out separation rule: the workers hired last, leave the firm first.
3. A return to seniority in wages: a worker’s wages depends on her seniority in the firm, that is her tenure relative to that of her colleagues.

We shall analyze these implications empirically. The first implication is not specific for our model. Moreover, it has been tested many times before. We show that it holds for the datasets at hand, just for the sake of completeness. The challenging aspect of this paper is testing the second and third implication. For that purpose, we need longitudinal matched employer-employee data. Only by knowing the tenure distribution of the entire workforce of the firm, at all times, we can calculate the seniority of a worker. Though using this type of data has become more fashionable in recent years, they are still not widely available. We have been able to get access to such data on Denmark and Portugal. We give a brief description of both data sets from these countries in the next subsection. Subsequently, we discuss the test of the three implications of our model, each in a separate subsection.

3.1 Data description

For Denmark, we use the *Integrated Database for labor Market Research* (IDA) for 1980-2001, from the Danish Bureau of Statistics, which has been used previously e.g. in Mortensen (2003). IDA tracks every single individual between 15 and 74 years old. The labor market status of each person is recorded once a year, at November 30. The dataset contains a plant identifier, which allows the construction of the total workforce of a plant, and hence of the firm as a whole. There is information on earnings, occupation, education, and age, and on the plant’s location, firm size, and industry. Industry is defined as the industry employing the largest share of the firm’s workforce. Firm size is defined as the number of individuals holding primary jobs in that firm and earning a positive wage. The tenure of workers hired since 1980 can be calculated straightforwardly from the IDA. For workers hired between 1964 and 1980, the tenure can be calculated from a second dataset on the contribution histories to a mandatory pension program, the *ATP*. The tenure in job spells started before 1964 is left censored (less than 3% of the observations). We calculate potential experience as age-schooling-6.

For Portugal, we use the *Quadros de Pessoal* for 1991-2000 provided by the Ministry of Employment, which has been used before e.g. in Cabral and Mata (2003). It is based on a compulsory survey of firms, establishments and all their workers; the compulsory participation enhances the quality of the data. The information available is similar to that for Denmark except that workers' tenure is directly reported; the industry of the firm is that industry with the highest share of sales or, when the allocation by sales is not possible, the industry with the highest employment share. We use all full-time employees in their main job, aged between 16 and 66, and working for a firm located in Portugal's mainland. The hourly gross earnings were computed as the monthly base-wage plus seniority-indexed components plus other regularly paid components, divided by normal hours of work per month.

Table 1: Descriptive Statistics Denmark and Portugal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Pooled sample											
DK 1980-2001	22364083	2771627	6870869	301015	164.22 (59.94)	5.41 (5.58)	20.25 (12.19)	12.05 (3.05)	37.79 (11.95)	0.35	0.63 (0.72)	4.68 (2.44)
PT 1991-2000	11420191	3211990	4268149	330270	3.68 (2.52)	8.43 (8.61)	20.26 (11.39)	6.81 (3.65)	36.73 (11.18)	0.40	0.87 (0.85)	4.04 (2.14)
	Cross-section year 2000											
DK 2000	1087922	1087922	1087922	79188	23.53 (9.02)	5.26 (5.78)	21.10 (11.68)	12.69 (2.74)	38.98 (11.75)	0.36	0.61 (0.73)	4.86 (2.48)
PT 2000	1403686	1403686	1403686	184957	4.06 (2.75)	7.93 (8.54)	20.47 (11.42)	7.59 (3.83)	37.14 (11.08)	0.42	0.75 (0.81)	3.64 (2.18)

Columns' labels: (1) Observations, (2) Workers, (3) Spells, (4) Firms, (5) Average Real Hourly Wage (base year=2000, measure unit: Euro), (6) Average Tenure, (7) Average Potential Experience, (8) Average Education, (9) Average Age, (10) Proportion of Women, (11) Average Relative Log Rank, (12) Average Log Firm Size. (Standard deviations in paratheses.)

For both countries, we use data for all private sector jobs, except agriculture, fishing and mining. We eliminate outliers by deleting all wage observations lower than the legal minimum wage and drop the top 1% of the wage distribution, for each year. Summary statistics for both countries are presented in Table 1, both for the pooled data and for 2000 separately¹¹. There are several obvious differences between the two countries. The mean level of education is more than 5 years higher in Denmark than in Portugal, while the mean tenure is almost 3 years longer in Portugal than in Denmark. The number of firms is far higher in Portugal than in Denmark, and the average firm size in Portugal is only 30% of that in Denmark. Finally, Danes earn on average almost six times more than Portuguese.

3.2 Testing Gibrat’s law

The assumptions underlying our theoretical model imply log employment size of firm j , n_{jt} , to follow a random walk, apart from the dampening effect of the hiring and firing boundary on short run fluctuations in n_{jt} . This regularity is known as Gibrat’s law. It holds in particular for large firms, see Jovanovic (1982). Though slight deviations from Gibrat’s law do not affect the main economic implications of the theoretical model, it is useful to have at least some idea how close this assumption is in the data. There is a massive literature on testing Gibrat’s law, see e.g. De Wit (2005) or Sutton (1997). Here, we use two tests.

The first approach is laid out in Abowd and Card (1989) and Topel and Ward (1992) for log wages; we adapt this methodology for log firmsizes. First, we estimate

$$\Delta n_{jt} = \delta_0 + \delta_1 Z_{jt} + \varepsilon_{jt}, \quad (12)$$

where Δ is the first difference operator and where Z_{jt} is vector of controls: age category of the firm, time effects and industry indicators. Second, we construct the autocovariance matrix of the residuals ε_{jt} of this regression. If n_{jt} follows a random walk, ε_{jt} should be uncorrelated across time t . The resulting covariograms for (12) are reported in Table 2, both for the whole sample of firms and for the subsample of larger firms (at least 20 employees each year over the sample period of that firm). The evidence from Table 2. suggests that Gibrat’s law holds closely, in particular when we exclude small firms. Most lagged correlations are really small relative to the variance of shocks that is reported in the first line. Gibrat’s law holds slightly better for Portugal than for Denmark.

The second approach follows Bond et al. (2005), who show that for micro panels with large cross-sectional and small time dimension, OLS in levels is consistent and typically more efficient than more complex GMM and ML estimators. Consider a

¹¹Summary statistics for each separate industry (both as broader and 2-digit industry categories) are available upon request.

Table 2: 1st Gibrat's Law Test: Residual Autocovariances

Lag	Denmark		Portugal	
	(1)	(2)	(1)	(2)
0	0.1587 (0.0005)	0.0424 (0.0112)	0.1162 (0.0005)	0.0255 (0.0007)
1	-0.0030 (0.0002)	-0.00003 (0.0005)	0.0002 (0.0002)	-0.0001 (0.0003)
2	-0.0094 (0.0002)	-0.0008 (0.0003)	-0.0024 (0.0002)	0.0012 (0.0004)
3	-0.0020 (0.0002)	-0.0002 (0.0002)	-0.0013 (0.0002)	0.0006 (0.0003)
4	-0.0016 (0.0002)	-0.00004 (0.0002)	-0.0008 (0.0003)	0.0006 (0.0002)
5	-0.0008 (0.0002)	-0.0006 (0.0005)	-0.0010 (0.0003)	0.0001 (0.0003)
6	-0.0008 (0.0002)	-0.0002 (0.0005)	-0.0013 (0.0004)	0.0002 (0.0004)
N obs generating reg	1505926	79425	878919	66369

Specification (1) uses all the firms; specification (2) uses all firms that have at least 20 employees in each year of their life spans. All generating regressions use the first differenced log firm size as dependent variable and control for age of the firm, time and industry effects. (Robust standard errors in parentheses)

simple dynamic AR(1) panel data model:

$$n_{jt} = \beta n_{j,t-1} + u_{jt}, \quad (13)$$

where $u_{jt} \equiv (1 - \beta)\gamma_j + v_{jt}$ and the initial firm size $n_{j1} = \alpha_0 + \alpha_1\gamma_j + \varepsilon_j\gamma_j$, with v_{jt} and γ_j error terms such that $E(\gamma_j) = E(v_{jt}) = 0$ and $E(v_{jt}v_{js}) = 0$ for $t \neq s$. Under the null of $\beta = 1$ the OLS estimator of β in (13) is consistent. We refer to this estimator of β as the OLS estimator. Under the alternative $\beta < 1$, the OLS estimator is biased upwards, the more so when $\text{Var}(\gamma_j)/\text{Var}(v_{jt})$ is large. In the latter case, one could use the transformed statistic in Breitung and Meyer (1994), which estimates β from:

$$n_{jt} - n_{j1} = \beta(n_{j,t-1} - n_{j1}) + \varepsilon_{jt} \quad (14)$$

where $\varepsilon_{jt} = v_{jt} - (1 - \beta)(n_{j1} - \gamma_j)$. The OLS estimator of (14) is consistent again under the null and again upwards biased under the alternative $\beta < 1$, but this time the bias does not depend on $\text{Var}(\gamma_j)/\text{Var}(v_{jt})$. The results for both methods are shown in Table 3, both for the full sample and the sample excluding small firms. The results are very similar to those in Table 2: Gibrat's law holds reasonably well, in particular for large firms, and in particular in Portugal. The value for the mean squared error of the regressions (MSE) is a good estimate for the parameter σ of our model. It is similar for both countries, being quite large (0.40) for the whole sample, and about half of that (0.20) for the subsample of larger firms.

Table 3: 2nd Gibrat's Law Test: Unit Root Type Regressions

Coef	Denmark				Portugal			
	all firms		large firms		all firms		large firms	
	OLS1	BM1	OLS2	BM2	OLS1	BM1	OLS2	BM2
β	.9361 (.0003)	.9208 (.0006)	.9755 (.0012)	.9806 (.0030)	.9594 (.0004)	.9537 (.0009)	.9791 (.0011)	1.043 (.0030)
N obs	1505926		79425		878934		66340	
R ²	0.87	0.70	0.95	0.82	0.91	0.66	0.96	0.84
MSE	0.42	0.43	0.21	0.21	0.36	0.36	0.17	0.17

The dependent variable is logfirm size in OLS columns and (logfirm size-initial logfirm size) in BM columns. Columns indexed 1 correspond to estimates using the sample of all firms, while columns 2 correspond to the sample of firms with at least 20 employees in each year of their life spans. Both regressions control for age of the firm, time and industry effects. (Robust standard errors in parentheses).

3.3 Testing the LIFO separation rule

Next, we turn to the second implication of our model, the LIFO separation rule. Since we apply an efficient bargaining model, the distinction between quits and layoffs is

arbitrary, compare McLaughlin (1991). As long as there is a positive surplus of the worker's marginal revenue to the firm above the worker's reservation wage, the worker and the firm will strike a deal. As soon as this surplus has vanished, it is in their mutual interest to separate. Whether the separation is initiated by the worker or by the firm is irrelevant. Hence, the model predicts the LIFO separation rule to apply to separations as a whole, not just to layoffs separately. We use duration analysis to test this for this implication. Let the function $j(i, t)$ denote the firm j in which worker i is employed in period t . We drop the arguments of this function whenever the identification of the individual and the period of observation are clear. The seniority level q_{ijt} is defined as the log of the number of workers employed at firm $j(i, t)$ at time t , for at least as long as or longer than worker i ; this number includes therefore worker i herself. Hence, q_{ijt} is equal to n_{jt} at the moment t when worker i is hired (assuming that i is the only one hired at time t). Furthermore, for the most senior worker $q_{ijt} = 0$ because there is only one worker who is employed at the firm as least as long as herself. Then, the seniority index r_{ijt} is defined as the log of the ratio of the number of people employed at least as long as worker i to the size of firm j at time t , in logs

$$r_{ijt} \equiv n_{jt} - q_{ijt}. \quad (15)$$

The seniority index r_{ijt} is a reasonable proxy for the variable $z_t - q$, since z_t is equal to n_t , up to a constant, ηp , and except for the insulation of n_t from shocks in z_t when $p^- < p_t < p^+$, recall the setup of our theoretical model. Were the LIFO separation rule to apply literally, the seniority index r_{ijt} would be the only determinant of separation. However, there are two reasons why this is not likely to be the case. First, the workforce of the firm is not completely homogeneous, so that a firm may wish to diminish its workforce in one skill class but not necessarily for other skill classes employed within that firm. This may disrupt a strict application of the LIFO separation rule. Second, workers separate not only due to shocks of the demand for the firm's product, but also due to worker specific shocks, e.g. when a worker's partner gets a new job in another city, which might cause the worker to quit from his or her current job. A particularly important worker specific factor that does not fit in the LIFO model is retirement. Hence, our ambition is more limited than what would follow from a strict interpretation of the LIFO separation rule. We just want to show that r_{ijt} has a strong impact on the job separation rate.

We model the transition process by a mixed proportional hazard rates model with discrete time periods. This implies that the conditional probability of leaving the firm (i.e. the hazard rate) after T_{ijt} years of tenure can be written as:

$$\theta(r_{ijt}, Z_{ijt}, T_{ijt}, v_i) = \frac{\exp\left(\beta r_{ijt} + \gamma Z_{ijt} + \psi_{T_{ijt}} + v_i\right)}{1 + \exp\left(\beta r_{ijt} + \gamma Z_{ijt} + \psi_{T_{ijt}} + v_i\right)} \quad (16)$$

where Z_{ijt} is a vector of observed characteristics of the individual and the job, and where v_i represents the unobserved worker heterogeneity. We include a full set of

dummies ψ_T for every tenure category, which is equivalent to a fully flexible specification of the baseline hazard. Identification of the parameter β of the seniority index r_{ijt} separate of the parameters of the baseline hazard ψ_T requires variation in r_{ijt} that is independent of the tenure T_{ijt} . Such independent variation is available since the seniority index also depends on the hiring and firing of other workers and thus is a non-deterministic function of tenure. A LIFO separation rule implies that β should be negative. For our estimation method we use a two mass-point distribution for the unobserved heterogeneity. We use up to 10 spells of an individual, which helps to estimate the unobserved heterogeneity distribution. We use a discrete time model, since workers are observed only once per year. Hence, we cannot observe the exact moment at which the worker enters or leaves the firm.¹² In addition, short spells are underrepresented since a worker has to stay at least till the next period of observation. With the data at hand, we cannot correct for these problems. Using a discrete time analysis is the best we can do.

At some point in time, older workers leave the firm for retirement. This process is independent of the LIFO separation rule. Therefore, we exclude workers above the age of 55 from the analysis. Spells started before the age of 55 and finished afterwards are therefore right censored. Women are also more likely to leave the firm for non-participation. Hence, we separate our results for men and women. We delete spells that are left censored since we cannot compute the seniority of an individual for the periods before she enters our observation sample. Since this seniority affects the probability that the individual survives till the start of the sample period, we cannot easily correct for left censoring. Deleting the left-censored spells implies that we have a maximum of 22 years of tenure in Denmark and 10 for Portugal. The vector Z_{ijt} includes education, potential experience and indicators for region, industry and occupation.

Table 4 lists the main results. We find a negative and significant impact of seniority for both women and men, with small differences between these categories, in both Denmark and Portugal, in accordance with the LIFO separation rule. Though the actual coefficients are not reported here, we also find negative duration dependence and evidence of unobserved heterogeneity.¹³ Apparently, seniority does not pick up all the variation in separation rates over the course of a job spell. There are two explanations for this phenomenon. First, as noted before, our seniority index might not exactly correspond to the actual layoff ordering, since the firm's workforce is likely to be heterogeneous with separate LIFO ordering to apply to subsets of the workforce. This is equivalent to measurement error in our seniority index r_{ijt} , leading to an attenuation bias in the estimate of β and unobserved variation in the seniority index being picked up by correlated variables. Second, not all separations are driven by the fluctuations in the demand for the firm's product, and hence, the log seniority

¹²For Portugal, tenures is reported in months. We use this information in the estimation. For the rest, the modelling is identical to that for Denmark.

¹³The full estimation results are available upon request.

index. For example, some separations might be driven by the worker and the firm learning about the quality of the match, see Jovanovic (1979). These separations do not fit the LIFO pattern.

Table 4: Main results LIFO test

	Denmark		Portugal	
	Males	Females	Males	Females
Logrank	-0.0577 (0.0019)	-0.0357 (0.0025)	-0.0549 (0.0054)	-0.0669 (0.0065)
Education	-0.1169 (0.0003)	-0.1267 (0.0005)	-0.1204 (0.0009)	-0.1446 (0.0012)
Experience	-0.0771 (0.0001)	-0.0732 (0.0001)	-0.0490 (0.0003)	-0.0656 (0.0004)
N obs	10788368	5990891	2118405	1488687

The estimation also controls for occupation, region and industry indicators. (Standard errors in parantheses)

3.4 Testing dependency of wages on seniority

3.4.1 Empirical specification and returns to seniority within the firm

The third implication of the model is the return to seniority in wages. This can be tested by extending the standard specification of the log earnings equation with the seniority index, r_{ijt} . Consider the following specification of log wages w_{ijt}

$$w_{ijt} = \alpha + \chi X_{ijt} + \gamma T_{ijt} + \delta r_{ijt} + \zeta n_{jt} + \varepsilon_{ijt}, \quad (17)$$

where X_{ijt} is experience. We omit higher order terms in experience and tenure and other controls (including time effects) from equation (17) for the sake of convenience, but include them in the estimation. The unobservable term can be decomposed into four orthogonal components, a match, a firm, a worker, and an idiosyncratic effect¹⁴

$$\varepsilon_{ijt} = \varphi_{ij} + \psi_j + \mu_i + \nu_{ijt}. \quad (18)$$

The idiosyncratic effect ν_{ijt} can also include measurement error. There are all kinds of reasons for φ_{ij} , ψ_j , and μ_i to be correlated to T_{ijt} , see Topel (1991) or Altonji and Williams (2005): good worker-firm relationships tend to survive as the worker and the firm learn about the quality of their match and bad matches are broken up, leading to a positive correlation between $\varphi_{ij} + \psi_j + \mu_i$ and T_{ijt} . Search theories imply

¹⁴This formulation is similar to Topel (1991: 150), except that we add a firm effect and that we delete the subscript t from the match effect φ_{ij} , as Topel does in his application.

that workers sample new jobs from a job offer distribution. The longer this selection process is going on, the higher the expected value of $\varphi_{ij} + \psi_j$ since bad jobs do not survive, leading to a positive correlation between $\varphi_{ij} + \psi_j$ and T_{ijt} . There are two obvious solutions to this problem, either within-job first differencing (FD) or adding fixed effects for every job spell (FE). First differencing yields

$$\Delta w_{ijt} = \chi + \gamma + \delta \Delta r_{ijt} + \zeta \Delta n_{jt} + \Delta \nu_{ijt}. \quad (19)$$

Adding fixed effects per job spell is equivalent to estimating (17) by taking deviations from the mean over time, within a job spell:

$$\tilde{w}_{ijt} = (\chi + \gamma) \tilde{T}_{ijt} + \delta \tilde{r}_{ijt} + \zeta \tilde{n}_{jt} + \tilde{\nu}_{ijt}, \quad (20)$$

where the upper tilde denotes deviations from the mean per job spell, e.g. $\tilde{w}_{ijt} = w_{ijt} - \bar{w}_{ijt}$, with $\bar{w}_{ijt}$ the mean over time of w_{ijt} . We exclude $\tilde{X}_{ijt}$ from (20) because it is perfectly collinear with $\tilde{T}_{ijt}$. In both specifications above, it is immediately clear that the first order effects of tenure and experience are not separately identified. This problem has troubled all attempts to estimate the return to tenure, see e.g. Altonji and Shakotko (1987) and the large stream of subsequent papers. The perfect collinearity of experience and tenure within a job spell rules out estimating the return to tenure on with spell variation in wages. Hence, researchers had to revert to between job spell variation in wages. However, since job mobility is not exogenous, using this type of information introduces all kind of selectivity issues, which the literature has tried to resolve. Happily, this problem does not affect the estimation of δ , since r_{ijt} is not perfectly correlated to T_{ijt} . This means that we can use only within job spell variation in wages to estimate δ , and hence we do not have to bother about the selectivity problems that plague data on job mobility.

The choice between the FE and FD estimators above depends on the error structure of v_{ijt} . The closer is v_{ijt} to a unit root, the more efficient is the FD method; the closer v_{ijt} is to being serially uncorrelated, the more efficient estimation method is the FE estimator. Previous empirical studies have typically found a high degree of autocorrelation in v_{ijt} , even close to a unit root, see for instance Abowd and Card (1979) and Topel and Ward (1992). From that perspective, equation (19) is likely to be most efficient. However, this equation assumes that the effect of r_{ijt} and n_{jt} on w_{ijt} is immediate. Any lagged impact will not be captured after first differencing. From that perspective, equation (20) is preferred, since there lagged effects of r_{ijt} and n_{jt} will be captured. Hence, one would expect higher estimates for δ and ζ from using equation (20) than from (19).¹⁵ In the strict version of our model, where separation is completely governed by the LIFO separation rule, r_{ijt} and n_{jt} are perfectly correlated within a job spell, since more senior workers will never leave the firm before worker i , so that the only variation in r_{ijt} comes from variation in n_{jt} . The same argument

¹⁵We report robust standard errors, so that correlation between the residuals over time does not affect their validity.

applies to $\tilde{r}_{ijt}$ and $\tilde{n}_{jt}$. Hence, δ and ζ are not separately identified in that world neither in equation (19) nor in (20). Happily, LIFO does not apply in a strict sense. The most compelling reason for a violation of the LIFO separation rule is workers' retirement, but also other individual specific shocks discussed earlier in this section. These separations allow separate identification of δ and ζ with FE and FD estimators.

Table 5: Residual Autocovariances for Within-Job LogWage Innovations

Lag	Denmark	Portugal
0	0.0231 (0.00002)	0.0273 (0.00007)
1	-0.0043 (0.00001)	-0.0082 (0.00006)
2	-0.0006 (8.7e-06)	-0.0008 (0.00003)
3	-0.0003 (9.0e-06)	-0.0004 (0.00003)
4	-0.0003 (9.5e-06)	9.2e-06 (0.00003)
5	-0.00008 (0.00001)	-0.00008 (0.00004)
6	-0.0001 (0.00001)	-0.0006 (0.00005)
N obs generating reg	14907897	5758655

The generating regressions are the FD wage regressions with logrank includes, see the FD2 columns in the next table. (Robust standard errors in parentheses)

First, we check the characteristics of the dynamic process of v_{ijt} . Table 5 reports the variance-covariance of Δv_{ijt} , analogous to what we did for log firm sizes in Table 2. For both countries, the covariance of ε_{ijt} with its first lag is substantial, the covariance with higher lags is negligible. Hence, the process is well approximated by an MA(1) process, made up of a mixture permanent and transitory shocks. Abowd and Card (1979) and Topel and Ward (1992) find similar results for the United States. The standard deviation of the permanent shocks can be calculated as 0.12 for Denmark and 0.10 for Portugal.¹⁶ These numbers are of the same order of magnitude as found for the United States.

¹⁶Let q_{ijt} and u_{ijt} be the transitory and permanent shock respectively. Then:

$$\Delta v_{ijt} = u_{ijt} + q_{ijt} - q_{ij,t-1}.$$

Hence: $\text{Var}(\Delta v_{ijt}) = \text{Var}(u_{ijt}) + 2\text{Var}(q_{ijt})$ and $\text{Cov}(\Delta v_{ijt}, \Delta v_{ij,t-1}) = -\text{Var}(q_{ijt})$, so that: $\text{Var}(u_{ijt}) = \text{Var}(\Delta v_{ijt}) + 2\text{Cov}(\Delta v_{ijt}, \Delta v_{ij,t-1})$.

Table 6: FE and FD Wage Regressions for the Entire Private Sector in Denmark and Portugal

	Denmark			Portugal		
	FD1	FD2	FE1	FE2	FD1	FE1
logrank		.003*** (.0003)		.008*** (.0003)	.016*** (.0005)	.022*** (.0005)
logsize	.013*** (.0002)	.011*** (.0003)	.026*** (.0003)	.021*** (.0003)	.015*** (.0005)	.040*** (.0004)
tenure+exper	.047*** (.0003)	.045*** (.0003)	.010*** (.0001)	.007*** (.0002)	.065*** (.0005)	.059*** (.0003)
tenure ²	.191*** (.002)	.199*** (.002)	-.052*** (.002)	-.036*** (.002)	-.069*** (.003)	-.083*** (.002)
tenure ³	-.101*** (.001)	-.105*** (.001)	.014*** (.099)	.008*** (9.88e-07)	.021*** (.001)	.024*** (.0007)
tenure ⁴	.002*** (.0002)	.002*** (.0002)	-.0009*** (.0002)	-.0002 (.0002)	-.002*** (.0002)	-.002*** (.00009)
exper ²	-.224*** (.002)	-.223*** (.002)	.099*** (.0006)	.100*** (.0006)	-.204*** (.004)	-.149*** (.002)
exper ³	.039*** (.0007)	.039*** (.0007)	-.039*** (.0002)	-.039*** (.0002)	.043*** (.001)	.030*** (.0007)
exper ⁴	-.003*** (.00007)	-.003*** (.00007)	.004*** (.00003)	.004*** (.0003)	-.003*** (.0001)	-.002*** (.00007)
N obs	14907897		22364083		5758655	10743244
Workers	2116307		277162		1752000	3092329
Spells	3745050		6870869		1965560	4053649
Firms	221106		301015		206361	322502

The dependent variable is the (Δ) time-detrended log real hourly wage for the (FD) FE columns; the covariates have Δ in front for FD columns. Columns 1 report results for the same regressions as corresponding columns 2, but without logrank included as covariate. The higher order polynomials in tenure and experience are divided by the corresponding powers of 10. All regressions also control for region, industry and occupation indicators. Significance levels: * : 10% ** : 5% *** : 1%. (Robust standard errors in parentheses).

This evidence suggests that in terms of efficiency of the estimation method we might prefer FD, while in terms of allowing for a lagged effect of Δr_{ijt} on Δw_{ijt} we might prefer FE. Hence, we report both the FD and FE estimator. Our regressions control for up to quartic terms in tenure and experience, log firm size and industry, occupation, and region dummies. In Table 6 we report the results¹⁷. We present the estimation results for two specifications, one excluding log seniority r_{ijt} and another including it. We can draw the following conclusions. First, all coefficients for log seniority are positive and statistically significant. Second, the coefficients are larger for FE than for FD, as was expected because FE allows for a lagged effect of r_{ijt} on w_{ijt} , while FD does not. Third, comparing the estimation results with and without seniority, including seniority reduces the coefficients for tenure and log firm size by 5-30 %. The coefficients for experience are hardly affected by including seniority. The effect of tenure and log firm size on wages is at least partly a proxy for the effect of seniority. We can expect seniority to be measured with the greater measurement error than tenure and firm size. Apart from straightforward reporting errors, the main source of measurement error in tenure is who exactly is the relevant employer. Some job changes might either be classified as between firms, justifying the tenure clock being set back to zero, or as within the firm, which does not affect the tenure clock. However, this source of measurement error only affects changes at the borderline of the definition of a firm. This is likely to be only a small fraction of the firm's workforce. However, misclassification of the tenure of even a single worker affects the measurement of the seniority of all other workers in the firm. In general, any measurement error in tenure or firm size automatically feeds into seniority, while on top of that, seniority is also affected by measurement errors because separate seniority statistics are likely to apply for subgroups of the workforce. We can therefore expect that actual effect of seniority on wages is larger than estimated here, and that part of the effect of tenure and log firm size is still a proxy for measurement error in the seniority variable. Finally, the effect of seniority is twice as high in Portugal as in Denmark.

3.4.2 Returns to seniority within gender and education subgroups

We repeat the analysis separately for males and females, and for low- and high-educated workers. The results are reported in Table 7. The results for male and female categories do not differ much. The only apparent exception is for Denmark, when using the FE estimator, where the estimated coefficient for males is twice as high as for females, though they are the same when using the FD estimator. At the same time, and linked to the previous observation, the estimates by gender categories really do not differ much from the estimate when using the whole samples, for either country.

¹⁷Results for the same analysis performed for each of our broad industry categories—enumerated in subsection 3.4.3— are qualitatively identical (with some between-industry heterogeneity in the magnitude of the estimated seniority coefficients) to the results at the national level in Table 6. They are available upon request from the authors.

Our interpretation is therefore that seniority positions within gender categories are no more relevant for wage determination than the seniority position within the firm as a whole, and hence splitting by gender is not likely to attenuate the measurement error in seniority index. However, the estimation results for education groups show that the effect of seniority is much larger for higher educated workers than for low educated workers. The impact of seniority on wages is lower for the low-educated workers, compared to corresponding estimates from Table 6, and the FE estimate is even significantly negative for Denmark, though small in absolute value. The impact of seniority on wages within the high educated group is much larger, both in Denmark and in Portugal. These results are consistent with the fact that high educated workers have steeper wage-tenure profiles than their low-educated peers. At the same time, they give support to the fact that the relevant seniority hierarchy within the firm is already more realistically captured when accounting for education levels.

Table 7: FE and FD Regressions by Gender and Education Rank Groups

		Denmark				Portugal			
		Gender Categories							
		Females		Males		Females		Males	
		FD	FE	FD	FE	FD	FE	FD	FE
logrank		.005*** (0.0004)	.005*** (0.0005)	.005*** (0.0004)	.010*** (0.0004)	.015*** (0.0007)	.019*** (0.0006)	.014*** (0.0007)	.019*** (0.0006)
logsize		.002*** (0.0005)	.014*** (0.0005)	.014*** (0.0004)	.025*** (0.0004)	.015*** (0.0007)	.028*** (0.0006)	.019*** (0.0006)	.031*** (0.0006)
ten+exp		.032*** (0.0004)	.009*** (0.0002)	.052*** (0.0004)	.007*** (0.0002)	.053*** (0.0007)	.042*** (0.0005)	.080*** (0.0007)	.073*** (0.0005)
N obs		5049388	7745676	9858509	14618407	2300767	4353808	3457888	6389436
		Education Categories							
		HighEduc		LowEduc		HighEduc		LowEduc	
		FD	FE	FD	FE	FD	FE	FD	FE
logrank		.010*** (.0003)	.020*** (.0004)	.002*** (.0004)	-.002*** (.0004)	.029*** (.002)	.032*** (.002)	.013*** (.0005)	.016*** (.0005)
logsize		.007*** (.0004)	.016*** (.0004)	.014*** (.0005)	.024*** (.0005)	.026*** (.002)	.026*** (.002)	.016*** (.0005)	.031*** (.0004)
(ten+exp)		.040*** (.0004)	.006*** (.0002)	.031*** (.0007)	.006*** (.0002)	.116*** (.002)	.099*** (.001)	.056*** (.0005)	.049*** (.0003)
N obs		9567345	14054988	5268672	8309095	259793	536920	5492034	10206324

The dependent variable is the (Δ) time-detrended log real hourly wage for the (FD) FE columns; the covariates have Δ in front for FD columns. Logrank has been computed separately for each category. "LowEduc" stands for category of people with at most 12 years of education. All regressions include also up to 4th order polynomials in tenure and experience and indicators for region, occupation and industry. Significance levels: * : 10% ** : 5% *** : 1%. (Robust standard errors in parentheses).

3.4.3 Returns to seniority and firm monopoly power

Our theoretical model also predicts that the return to seniority, β/η , is partly driven by the degree of monopoly power, η^{-1} . We test this hypothesis by analyzing whether the variation in the return to seniority across industries can be explained by the degree of monopoly power in each industry. We take the log of the number of firms in each industry as proxy for the degree of monopoly power. We regress the estimated coefficient for seniority for each industry on this log number of firms and a constant term, using both simple OLS and Weighted Least Squares (WLS) specifications. We use two measures as "number of firms" in an industry: the sum of all firms that were at any time active in that industry during the sample period, and respectively, the median number of firms over the sample period. We use two industries classifications, a broad classification with 12 industries for Denmark and 11 for Portugal¹⁸ and a more refined classification where we use all 2-digit Standard Industry Classification (SIC) industry sub-categories available, increasing the number of observations in the regressions to 40 for Denmark and 49 for Portugal. For our prediction to be verified, we expect negative estimates of the coefficients of log number of firms.

The estimation results for the regressions of returns to seniority on the log number of firms by industry are presented in Table 8. Most of the estimated coefficients of interest are not significantly different from 0 (though most slightly negative in magnitude), both when using the WLS and the OLS methods and regardless of using as dependent variables the FD or the FE coefficients previously estimated in this paper, and as independent variables the sum or the median of the number firms in an industry. There are very few cases where the results are statistically significant: when using the broad industry categories for Portugal we get significant coefficients of the expected sign with the FE method, but significant coefficients of the opposite sign in Denmark; when using the OLS for 2-digit industries in Portugal we get significantly positive coefficients for the FD method and again significantly positive when using the WLS for the FD, sum of firms, and FE, median of firms. In conclusion, we regard this test as inconclusive. The explanatory variables used as proxy for the monopoly power of an industry are not strong enough to isolate the effect of the degree of monopoly power on the return to seniority.

¹⁸The broad industry categories we use are: 1. Manufacturing; 2. Electricity, gas and water supply; 3. Construction; 4. Wholesale and retail trade, repairs; 5. Hotels and restaurants; 6. Transport, post and communications; 7. Financial intermediation; 8. Real estate, renting and business activities; 9. (private firms operating in) Public administration and defense, compulsory social security; 10. Education; 11. Health and social work; 12. Other community, social and personal service activities. *Note:* For Portugal we miss category 9 (no firms are privately owned in that sector).

Table 8: Monopoly Power Test

	Denmark			Portugal		
	Sum	Median	Sum	Median	Sum	Median
	FD	FE	FD	FE	FD	FE
OLS, Broad Industry Categories						
log Nfirms	-0.00009 (0.0014)	0.0058** (0.0024)	0.0003 (0.0014)	0.0060** (0.0022)	0.0046 (0.0028)	-0.0069* (0.0032)
constant	0.0107 (0.0135)	-0.0435* (0.0227)	0.0072 (0.0111)	-0.0378* (0.0178)	-0.0275 (0.0267)	0.0892** (0.0303)
					0.0046 (0.0027)	-0.0068* (0.0030)
					-0.0234 (0.0230)	0.0817** (0.0262)
WLS, Broad Industry Categories						
log Nfirms	-0.0008 (0.0014)	0.0053** (0.0024)	-0.0004 (0.0014)	0.0002 (0.0013)	-0.0004 (0.0031)	-0.0057 (0.0035)
constant	0.0174 (0.0134)	-0.0395 (0.0227)	0.0128 (0.0117)	0.0075 (0.0109)	0.0229 (0.0302)	0.0778** (0.0340)
					0.0212 (0.0269)	-0.0102 (0.0251)
OLS, 2-Digit Industry Categories						
log Nfirms	-0.0010 (0.0011)	-0.0010 (0.0016)	-0.0010 (0.0012)	-0.0007 (0.0016)	0.0136** (0.0058)	-0.0003 (0.0030)
constant	0.0213** (0.0093)	0.0224* (0.0125)	0.0198** (0.0080)	0.0193* (0.0108)	-0.0852* (0.0440)	0.0314 (0.0228)
					-0.0559 (0.0367)	0.0358* (0.0187)
WLS, 2-Digit Industry Categories						
log Nfirms	-0.0012 (0.0013)	0.0013 (0.0015)	-0.0013 (0.0013)	-0.0010 (0.0013)	0.0094* (0.0048)	-0.0004 (0.0025)
constant	0.0227** (0.0110)	-0.0022 (0.0125)	0.0213** (0.0085)	0.0194** (0.0091)	-0.0515 (0.0365)	0.0306 (0.0195)
					-0.0294 (0.0307)	-0.0421 (0.0315)

The dependent variable is the estimated coefficient of seniority rank from FD and respectively FE regressions, by industry. "Nfirms" is measured as the sum of all firms (corresponding to columns labeled "Sum") or as the median of firms (corresponding to columns "Median"), over the sample period, respectively. There are 12 'broad industries' in Denmark and 11 in Portugal, while the number of 2-digit SIC sub-categories is 40 in Denmark and 49 in Portugal. Significance levels: * : 10% ** : 5% *** : 1%. (Standard errors in parentheses).

4 Summary and conclusions

We have shown beyond reasonable doubt that for both Denmark and Portugal there exists a return to seniority in wages, with an elasticity in the order of magnitude of 0.01-0.02: a 1 % increase in the seniority hierarchy of the firm raises your wage by 0.01-0.02 %. This return is almost twice as high for higher educated workers, and it is twice as high in Portugal as it is in Denmark. Some 5-30 % of what has been known as the return to tenure and to firm size is in fact a return to seniority. This implies that standard explanations of the return to tenure, like Jovanovic's (1979) learning, or the search models, and subsequent versions of these models, cannot provide a full explanation of what is going on, if only because these explanations focus solely on the features of the worker herself (in case of learning, her ability; in case of search, her job offer history), while the return to seniority links the fate of the worker to that of the firm as a whole. A return to seniority implies that a worker is to some extent shareholder in her own firm. Hence, it makes the link between labor economics and finance.

Our theoretical model provides a special interpretation of the return to seniority, as being due to a hold up problem, where firms pay the full cost of the specific investment, while workers capture part of the return. This setup leads to inefficiently low hiring. All these conclusions are conditional on the assumption that the firm bears the full cost of specific investments, an assumption that has not been tested empirically in this paper. How to do that remains an open question. An indirect answer can be obtained by analysing who is queueing for whom: when workers queue for jobs, so that there is unemployment, firms are held up by their incumbent workforce; when it is the other way around, and there are vacancies, workers are held up by their employer. As long as workers are risk neutral and either investment or wages are contractible, efficient hiring can be obtained by using the sharing rule of the costs for the one, to mirror the sharing rule for the other, thereby satisfying the Hosios (1990) condition. When workers are risk averse, efficiency can only be obtained when both investment and wages are contractible, such that the costs of investment are fully attributed to the firm and there is no seniority profile. Any other allocation assigns part of the risky return to the risk averse player. In that sense, our estimation results point to incompleteness in the insurance market. Nevertheless, our analysis does not imply that LIFO layoff rules are bad per se. They can offer a useful protection to the property rights of incumbent workers on their share of the specific investment, thereby helping the firm to solve a commitment problem. Without a resolution of this commitment problem, incumbents would have all reasons not to cooperate in the transfer of tacit knowledge to newly hired workers.

We have established the existence of a return to seniority for Denmark and Portugal. Whether such a return exists in other countries, in particular in the United States, remains an open question. We bet it does; the large return to tenure in the United States as compared to Denmark and Portugal, see e.g. Teulings and Hartog (1998:36-37) suggests so. One might argue that returns to seniority are largely driven

by legal institutions, and that these institutions are entirely different and more market oriented in the United States. We think however that the economic mechanisms for having a LIFO layoff rule exist everywhere, and that the legal institutions might very well just be a formalisation of rules of conduct and implicit contracts that would have emerged anyway.

Our model suggests that hold up problems reduce turnover, and thereby specific investment (because turnover requires new specific investment to be made). This conclusion is contingent on the way specific investment is modelled here, namely as a fixed amount to be invested in one shot, at the beginning of the job. When the amount of investment can vary both in size and in timing, this conclusion might change. Then, a longer expected job duration might invoke more specific investment, which in turn would lengthen the expected job duration since the productivity at the job is raised relative to the productivity at the outside market. In such a world, a firm responds along two margins of adjustment, when the demand for its product goes up. First, it hires additional workers, and second, it expands the specific investment in its incumbent workforce. This model would provide further legitimation for a LIFO rule, not as legal constraint, but as an efficient economic institution. Again, we postpone this for future research.

5 References

- Abowd, J.M. and D. Card (1989), "On the covariance of earnings and hours changes", *Econometrica* 57(2), 411-45
- Altonji, J.G. and R.A. Shakotko (1987), "Do wages rise with seniority?", *Review of Economic Studies*, 54, 437-459.
- Altonji, J. G. and N. Williams (2005), "Do wages rise with job seniority? A reassessment", *Industrial and Labor Relations Review*, 58 (3), 370- 397.
- Bentolila, S. and G. Bertola (1990), "Firing costs and labour demands: how bad is eurosclerosis?", *Review of Economic Studies*, 57 (3), 381-402
- Bertrand, M. and S. Mullainathan (2003), "Enjoying the quiet life? Corporate governance and managerial preferences", *Journal of Political Economy*, III(5), 1043-1075
- Bond, S., C. Nauges and F. Windmeijer (2005), "Unit Roots: Identification and Testing in Micro Panels", *Cemmap working paper CWP07/05*
- Breitung, J. and W. Meyer (1994), "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?", *Applied Economics*, 26, 353-61
- Brown, C. and J. Medoff (1989), "The employer size wage effect", *Journal of Political Economy*, 97, 1027-1059.
- Buhai, I.S. and C.N. Teulings (2006), "Tenure Profiles and Efficient Separation in a Stochastic Productivity Model", *TI DP 05-099/3*, revised & resubmitted to the

- Review of Economic Studies, May 2008
- Cabral, L.M.B. and J. Mata (2003), "On the evolution of the firm size distribution: facts and theory", *American Economic Review* 93, 1075–1090.
- Dixit, A.K. and R.S. Pindyck (1994), "Investment under uncertainty", Princeton University Press.
- Dixit, A.K. (1989), "Entry and exit decisions under uncertainty", *Journal of Political Economy*, 97, 620-638.
- Hosios, A.J. (1990), "On the efficiency of matching and related models of search and unemployment", *Review of Economic Studies*, 57, 279-98.
- Jovanovic, B. (1979), "Job matching and the theory of turnover", *Journal of Political Economy*, 87(5), 972–990.
- Jovanovic, B. (1982), "Selection and the evolution of industry", *Econometrica*, 50(3), 649-670.
- Kuhn, P. (1988), "A nonuniform pricing model of union wages and employment", *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.
- Lindbeck, A. and D. Snower (1990), "Do Cooperation and Harassment Explain Involuntary Unemployment?", *American Economic Review*, June, 80(3), 167-188.
- McLaughlin, K. J.(1991), "A Theory of Quits and Layoffs with Efficient Turnover," *Journal of Political Economy*, vol. 99(1), pages 1-29, February
- Mortensen, D.T. (2003), *Wage dispersion: why are similar workers paid differently?*, MIT Press, Cambridge
- Sutton, J. (1997), "Gibrat's legacy", *Journal of Economic Literature*, 35(1), 40-59
- Teulings, C.N. and J. Hartog (1998), "Corporatism or 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, 145-175.
- Topel, R.H. and M.P. Ward (1992), "Job mobility and the careers of young men", *Quarterly Journal of Economics*, 107(2), 145-176
- Wit, G. de (2005), "Firm size distributions. An overview of steady-state distributions resulting from firm dynamics models", *International Journal of Industrial Organization*, 23, 423-450

A Derivation

Substitution of equation (3), (5), (6), and (7) in equation (8) yields

$$\begin{bmatrix} 0 \\ 0 \\ I \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -\beta & \mu^- & -\mu^+ \\ E & -F & G & -H \\ E - \beta F & \mu^- G & -\mu^+ H \end{bmatrix} \begin{bmatrix} R \\ \psi \\ C^- \\ C^+ \end{bmatrix},$$

where

$$\begin{aligned} \eta\lambda^- &\equiv \mu^- < 0, \eta\lambda^+ \equiv \mu^+ > 1, \psi \equiv \frac{\mu^-}{\rho(\mu^- - \beta)} > 0, \\ C^- &\equiv B^- \exp[\mu^- p^-], C^+ \equiv B^+ \exp[\mu^+ p^-], \\ R &\equiv r(\eta^{-1})^{-1} \exp[p^- - \pi] > 0, \Delta \equiv p^+ - p^-, \\ E &\equiv \exp[\Delta], F \equiv \exp[\beta\Delta], G \equiv \exp[\mu^- \Delta], H \equiv \exp[\mu^+ \Delta]. \end{aligned}$$

Elimination of C^- from the first two equations of this system yields equation (9). Matrix inversion yields

$$\begin{bmatrix} R \\ \psi \\ C^- \\ C^+ \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & -1 \\ 1 & -\beta & \mu^- & -\mu^+ \\ E & -F & G & -H \\ E - \beta F & \mu^- G & -\mu^+ H \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ I \\ 0 \end{bmatrix},$$

We are interested in solution with $\Delta > 0$. The second equation of this system can be written as

$$\begin{aligned} I \cdot R(\Delta) &= \psi \cdot S(\Delta, \beta), \\ R(\Delta) &\equiv -(\mu^+ - \mu^-) E + (\mu^+ - 1) \mu^- G + (1 - \mu^-) \mu^+ H, \\ S(\Delta, \beta) &\equiv (\mu^+ - 1)(\beta - \mu^-)(FG + EH) + (1 - \beta)(\mu^+ - \mu^-)(EF + GH) \\ &\quad - (1 - \mu^-)(\mu^+ - \beta)(FH + EG), \\ R(0) &= 0, R_\Delta(0) = (1 - \mu^-)(\mu^+ - 1)(\mu^+ - \mu^-) > 0, R(\Delta) \geq 0, \\ S(0, \beta) &= 0, S_\Delta(0, \beta) = 0, S_{\Delta\Delta}(0, \beta) > 0, S(\Delta, \beta) \geq 0. \end{aligned} \tag{21}$$

Hence

$$\begin{aligned} \lim_{\Delta \rightarrow 0} \frac{S(\Delta, \beta)}{R(\Delta)} &= \frac{S_{\Delta\Delta}(0, \beta)}{2R_\Delta(0)} \Delta = 0, \\ \frac{\partial}{\partial \Delta} \frac{S(\Delta, \beta)}{R(\Delta)} &> 0, \\ \lim_{\Delta \rightarrow \infty} \frac{S(\Delta, \beta)}{R(\Delta)} &= \infty. \end{aligned}$$

where the second line follows from the evaluation of higher order derivatives. Hence, there is a unique positive solution for Δ . The other three equations can be written as

$$\begin{aligned} R &= \frac{-(\mu^+ - \mu^-) \beta F + (\mu^+ - \beta) \mu^- G + (\beta - \mu^-) \mu^+ H}{R(\Delta)} \psi, \\ C^- &= \frac{-(\mu^+ - \beta) E + (\mu^+ - 1) \beta F + (1 - \beta) \mu^+ H}{R(\Delta)} \psi, \\ C^+ &= \frac{(\beta - \mu^-) E - (1 - \mu^-) \beta F + (1 - \beta) \mu^- G}{R(\Delta)} \psi. \end{aligned} \tag{22}$$

By a similar argument, one can prove that the numerators of these expression are always positive. Hence, a positive solution for R, C^- , and C^+ exists for every $\Delta > 0$. This implies a positive solution for B^- and B^+ . Totally differentiating equation (21) with respect to Δ and β yields an expression for $d\Delta/d\beta$. Totally differentiating equation (22) with respect to Δ and β and using the expression for $d\Delta/d\beta$ proves $\frac{\partial B^+}{\partial \beta} < 0$, $\frac{\partial p^+}{\partial \beta} > 0$, and $\frac{\partial p^-}{\partial \beta} < 0$. ■