Returns to Tenure or Seniority? *

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

This version: 6 July 2009

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.

Keywords: irreversible investment, options, seniority, LIFO, matched worker-firm data

JEL-codes: J31, J41, J63

---

*We are grateful for insightful comments and suggestions to Jaap Abbring, Joe Altonji, Marianne Bertrand, Paul Bingley, Moshe Buchinsky, Eric French, Pieter Gautier, Larry Katz, Espen Moen, Dale Mortensen, Kevin Murphy, Derek Neal, Jean Marc Robin, Robert Shimer, Robert Topel, Till von Wachter, and participants in seminars and conferences at ELE in Bergen, University of Alicante, Harvard, UC Merced, University of Minho, IFS London, SED in Boston, 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.

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

1 Introduction

Why does Lars earn a lower wage than Jens, while they have the same human capital and work at the same firm? And why is Pedro fired and his colleague Miguel allowed to stay at the firm, when their employer has to scale down employment? Some might think that the answer to these questions is obvious: it is simply because Jens and Miguel have higher seniority at the firm than Lars and Pedro, respectively. *Seniority* is different from tenure in that it measures the *worker's tenure relative to the tenure of her colleagues.* A worker's seniority is her tenure compared to the tenure of all her co-workers at the firm. This paper provides empirical evidence and offers a theoretical explanation 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 henceforth), that is, workers are fired according to their seniority. When we claim that seniority affects your separation risk, we mean that on top of the negative effect of the elapsed tenure on the hazard rate, being a senior worker with many more junior colleagues has a further negative effect. Furthermore, we show that there is a *return to seniority in wages*, even when controlling for tenure. To the best of our knowledge, this paper is the first to document the existence of this return to seniority.

We set up an 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. The optimal hiring and firing points can be calculated by considering the expected discounted marginal revenue of hiring an additional worker while (i) accounting for the option to fire that worker at a future point in time, (ii) taking as given all workers currently employed by the firm, and (iii) disregarding any workers that might be hired in the future. In this way, the hiring and firing of each worker can be considered separately from 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 quasi-rents 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. This implies that some gains from trade between the union and the firm remain unexploited. While Kuhn and Robert elaborate these ideas in a static framework, we use the dynamic model of Bentolila and Bertola. We take an eclectic approach on the functional form of the return to seniority, by simply positing a log-linear sharing rule of the quasi-rents. However, we impose efficient bargaining: as long as there are positive quasi-rents, the worker and the firm will reach agreement on their distribution. Hence, firing decisions are efficient. Hiring decisions are efficient if and only if costs 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 (1990) condition, like in 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 because workers capture part of the quasi-rents. As an extension, we consider the effect of firing costs, accounting for its upward effect on wages.[1] By the efficient bargaining assumption, firing costs do not affect firing, but further deteriorate 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. However, the sparse empirical evidence on the issue does not confirm this hypothesis. Following up on a further suggestion by Kuhn (1988), we observe that the state of facts might fit our theory. The LIFO layoff rule allows for a decentralisation of the bargaining process to the individual level –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 need all-encompassing matched worker-firm data to establish workers' seniority because

[1] In 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 crucial for the conclusions.

we need to know the tenure of all the workers in the firm. We show that junior workers have *ceteris paribus* a higher separation probability than senior workers. 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 main 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 kinds 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. Furthermore, Neal (1995) and others have shown that the return to the tenure at the firm, as measured by the previous studies, is partly a proxy for a return to tenure in the industry or occupation. We show that this finding does not affect either our estimation of the return to seniority. We find wage returns to seniority of 1 to 2 % in Portugal, and returns half that range in Denmark. We argue that our results cannot be explained by tournament models such as Lazear and Rosen (1981) and Malcomson (1984), learning models à la Jovanovic (1979), or job search models like Postel-Vinay and Robin (2002) and Burdett and Coles (2003).

The paper is set up as follows. Section 2 presents our theoretical framework. In Section 3, we describe the data for Denmark and Portugal, and present our empirical specifications and estimation results. Section 4 concludes.

2 Theoretical framework

2.1 Assumptions

Consider a world where firms produce a differentiated product. Each firm faces a stochastic iso-elastic demand curve for its output:

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

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 demand (= 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, labor productivity is normalized to unity, so that output is equal to employment. The market index is non-verifiable, such that the worker and the firm cannot contract on it. The hiring order of workers is verifiable, such that the contract may be contingent on this. Both workers and firms are risk neutral. Workers live forever and maximize the discounted value of their expected income. Firms maximize the discounted value of their expected profits. The discount rate is denoted by ρ. Finite discounted values require this discount rate to be larger than the growth rate of expected demand:

$$\rho > \mu + \frac{1}{2}\sigma^2. \tag{2}$$

At the start of an employment relation, the firm has to make a specific investment I.[2] This investment 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. The specific investment is also unverifiable. Hence, the firm and the worker cannot write a contract in which future payments are contingent on whether or not this investments is actually made. Workers have some bargaining power; workers will capture some share of the quasi-rents generated by this specific investment. Hiring new workers and making the specific investment requires the consent of the incumbent workers. The motivation for this assumption is that a main part of the specific investment is the transfer of the firm's tacit knowledge, which is a monopoly of senior workers. An alternative motivation is that senior workers can harass newcomers, as suggested by Lindbeck and Snower (1990). At the outside market, workers can earn an outside wage, which is constant over time. It is most convenient to think of this outside 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-outside wage.

[2] In Bentolila and Bertola (1990), there is both hiring and firing costs per worker. For the moment, we set the firing costs equal to zero. As an extension, we consider the implications of firing costs.

2.2 A benchmark case: $w = 0,\ I = 0$

As a benchmark, we analyze the simple case in which firms pay workers the outside wage, $w = w^r = 0$, where w denotes the log wage paid by the firm, 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:

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

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 yields the standard Amoroso-Robinson rule:

$$\begin{aligned} p_t &= \pi, \qquad (3) \\ n_t &= z_t - \eta\pi, \\ \pi &\equiv \ln \frac{\eta}{\eta - 1} > 0. \end{aligned}$$

The firm's price is constant over time, since the firm applies a fixed markup over marginal cost: log marginal cost is $w^r = 0$, so that $\pi > 0$. The log of the firm's labor demand follows a random walk, a regularity known as Gibrat's law, which holds particularly for larger firms, see for instance Jovanovic (1982).

2.3 The case with specific investment: $w = 0, I > 0$

When $I > 0$ 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. If demand falls after hiring the worker, the firm cannot get rid of the worker without losing the specific investment. Hence, the firm does not hire additional workers at the first moment that p_t rises above π. Instead, the firm postpones hiring till p_t is substantially above π. The reverse holds at 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 specific investment again. Hence, retaining the worker has an option value. The firm postpones firing till p_t is substantially below π. Analogous to Bentolila and Bertola (1990), we shall show that the optimal policy of a firm is to hire workers whenever p_t reaches a constant hiring bound p^+ and to fire workers whenever p_t reaches a constant firing bound p^-. The hiring bound is above the price level for the case $I = 0$, while the firing bound is below it: $p^+ > \pi > p^-$.

The situation is sketched in Figure 1. The downward sloping curves are inverse labor demand curves, specifying log marginal revenue $mr(\cdot)$ as a function of $z_t - n_t$:

$$\ln\left[\frac{d(N_t \cdot P_t)}{dN_t}\right] = \frac{1}{\eta}(z_t - n_t) - \pi \equiv mr(z_t - n_t). \tag{4}$$

Let the current level of the market index be z_0. In the case $I = 0$, the firm sets log

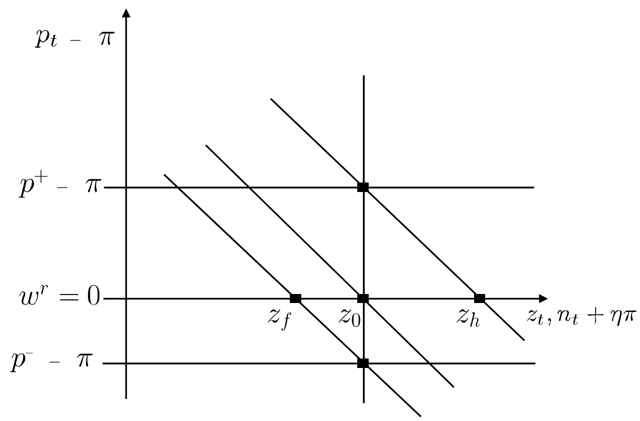

Figure 1: Firing-hiring boundaries with stochastic market index

employment n_0 such that log marginal revenue $mr\,(z_0 - n_0)$ is equal to log marginal cost $w^r = 0$; hence $n_0 = z_0 - \eta\pi$, see equation (3). Any change in z_t will immediately affect log employment n_t, but will leave the log price constant at π. In the case $I > 0$, the firm hires additional workers only when z_t rises above z_h, because at that point the price level hits the hiring bound, $p_t = p^+$. Similarly, the firm fires workers only when the market index falls below z_f, since then the price hits the firing bound, $p_t = p^-$. 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 keep p_t in the interval $[p^-, p^+]$. Then, p_t is held constant, and n_t starts drifting, either up (if $p_t = p^+$), or down (if $p_t = p^-$). We provide expressions for p^+ and p^- later on.

2.4 The case with a LIFO rule and rent sharing

Specific investment creates quasi-rents. Since workers have some bargaining power, they will be able to capture part of these rents. Hence, workers have an interest in the continuation of their relationship with the firm. Since hiring requires the consent of the incumbent workforce and since incumbents have an interest in the continuation of their relationship with the firm, they want to make sure that new hires will not be used as replacement for themselves when the firm has to downsize at some future time. Since the layoff order is contractible, they can refuse to train new workers unless the firm agrees not to fire any of the current incumbents before having fired all workers that will be hired. The firm has no incentive to oppose this demand of the incumbents; it is indifferent between firing either the one or the other worker when necessary, since after having fired either of them, the bargaining position vis-à-vis the remaining workforce is identical. Because similar bargaining has taken place at all previous points in time when the firm has hired additional workers, this outcome generalizes to the firm applying a LIFO (Last-In-First-Out) separation rule: fire the latest hires first.

Claim 1 *Consider a firm that satisfies the previous assumptions. If this firm wants to fire some of its workers, it applies a LIFO separation rule, firing first the workers whom it has hired last.*

When the firm pays its workers their outside wage, see Sections 2.2 and 2.3, there is no rationale for a LIFO layoff rule. Since the worker receives her outside wage, she is indifferent between working at the firm and being laid off. Hence, there is no point in fixing an order of layoff. However, when the firm pays its workers some share of the quasi-rents of the specific investment, a layoff order carries practical relevance, as it protects the 'property rights' of senior workers on their part of the quasi-rents.

Under this LIFO layoff rule, the distribution of completed job tenures is characterized by the following proposition.

Proposition 1 *Consider a firm that satisfies the assumptions of Section 2.1, and that hires workers when $p_t = p^+$ and fires workers when $p_t = p^-$. 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_t, with initial value z_0, to pass the barrier $z_0 - \eta\left(p^+ - p^-\right)$ for the first time. This distribution 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 particular worker who is hired at time h, when log employment is n_h, will be fired at the first moment $f > h$ that log employment is equal to n_h and $p_f = p^-$, since this worker can only be fired when all workers hired after her, by implication at $h < t < f$, have been fired, due to the LIFO rule, whereas all workers that have been hired before her, at $t < h$, cannot be fired before this worker, also due to the LIFO rule. Hence, $z_h - z_f = \eta\left(p^+ - p^-\right)$ and $z_h - z_t > \eta\left(p^+ - p^-\right), \forall t \in (f, h)$. The distribution of the first moment in time that $z_h - z_f = \eta\left(p^+ - p^-\right)$ 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 if fits the data well.

Having established that a firm and its workers agree to apply a LIFO layoff rule, the next question is whether all workers receive the same wage or whether the amount of rents each worker can capture differs by her degree of seniority. Kuhn (1988) and Kuhn and Robert (1989) offer a neat further legitimation for applying a wage schedule that differentiates wages between otherwise equal workers. Consider the standard monopoly union model, where the union bargains for a 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 outside 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 the 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 simultaneously set a layoff order and bargain on a wage schedule that differentiates wages according to the worker's position in the order of layoff. Since Kuhn and Robert (1989) specify their theory in a static framework, the layoff ordering can be based on anything: height, IQ, experience, or what else springs to mind. In the dynamic framework that we apply here, this order must necessarily be the LIFO rule, see Claim 1.

For the implementation of the LIFO layoff rule, define the *rank* q_t of a particular worker to be the log number of workers in the firm with tenure greater or equal than this worker (hence including herself), at time t. Suppose this particular worker has been hired at time $h, h < t$. Due to the assumption that workers live forever and due to the LIFO rule, all workers who have been hired before her are still employed by the firm, because they can only be fired after she is fired. Hence, her rank q_t is equal to log employment at the moment when she was hired: $q_t = n_h$. Since n_h is constant, so is q_t, such that we can drop the suffix t. Hence, a worker hired at time h gets rank q, $q = n_h = z_h - \eta p^+$.

Consider a wage schedule that is in between the reservation wage and the firm's inverted labor demand curve $mr\,(z_t - q)$; then the firm will never find it attractive to fire any worker, since it makes a profit on the marginal worker at each employment level. Thus, even though inframarginal workers earn wages above the marginal revenue of the last worker hired, the firm has no incentive to fire these workers since it may only fire these expensive inframarginal workers after having fired the cheap marginal workers. Hence, the employment is set at the efficient level where marginal revenue is equal to the reservation wage.[3] The LIFO rule provides senior workers protection against layoff, which they use to demand a higher wage. Given the lack of a formal theory of multiplayer bargaining, we take an eclectic approach in wage setting. We assume the relation between the log-marginal revenue $mr\left(z_t - q\right)$ and the log-reservation wage on the one hand, and log wages on the other hand, to be linear:

$$w(z_t - q) = \beta \cdot mr\left(z_t - q\right) + \beta\left(\pi - p^{-}\right) + \omega = \frac{\beta}{\eta}\left(z_t - q\right) - \beta p^{-} + \omega, \tag{5}$$

with $0 \leq \beta \leq 1$. The parameter β can be interpreted as the bargaining power of workers.[4] This wage setting scheme is a form of price discrimination on the side of the union. The case $\beta = 1$ corresponds to first degree price discrimination, where the union has full bargaining power. In the case $\beta = 0$, workers have no bargaining power and they get just their reservation wage. The parameter 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 parameter is between zero and unity. At the moment of firing f, $z_f - q = \eta p^{-}$, and hence, by equation (4), $w(z_f - q) = \omega$. The parameter ω can therefore be interpreted as the log reservation wage of an incumbent worker. We present an expression for ω later on.

[3] We ignore here the inefficiency that is due to the monopoly power of the firm at its product market, which causes its marginal revenue to be below the output price.

[4] In the case of a single worker-firm pair, where we could apply the theory of two player bargaining, as in Buhai and Teulings (2006), the log-linear sharing rule would be almost equivalent to Nash bargaining, which would yield a linear (instead of log-linear) sharing rule of the surplus between the asset values of the worker and the firm (instead of the surplus between current productivity and the reservation wage).

The situation is depicted in Figure 2. The left panel sketches the situation for Kuhn and Robert's (1989) static framework. The two continuous lines in the graph are the marginal revenue, $mr(z_0 - q)$, and respectively the wage schedule, $w(z_0 - q)$. The shaded area between them 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 $mr(z_0 - q)$ and the log reservation wage ω, neither the firm nor the workers have an 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 (5) implies efficient bargaining: all gains from trade are exploited. Although inframarginal workers earn a log wage $w(z_0 - q)$ above the log marginal revenue $mr(z_0 - n_0)$, the firm has no incentive to fire them, because it is obliged to fire the less expensive marginal workers first. In the dynamic framework à la Bentolila and Bertola (1990) there are some complications, see the right panel of Figure 2. Since investment I is irreversible, workers have an option value for the firm. Hence, the firm retains workers when marginal revenue has fallen just slightly below marginal cost, $w^r > mr(z_0 - n_0) > p^- - \pi$, since by firing the worker the firm loses the option of benefiting from the market index rising above z_0 at a later stage without having to pay the specific investment I again. 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, $w^r < w(z_0 - n_0) < \omega$, since by quitting the worker loses the option of benefitting when the market index z_t rising above z_0 at a later stage, and hence log wages rising above w^r. For some range of $z_t \in [z_f, z_0]$, marginal revenue

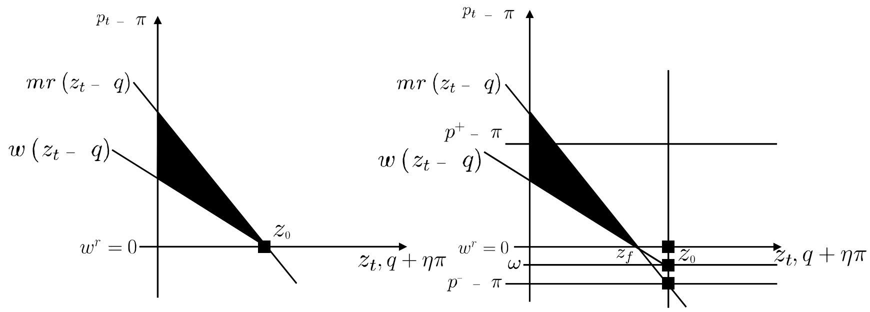

Figure 2: Static vs. dynamic framework

is therefore below marginal cost, $mr\,(z_t - n_0) < w^r$, and the wage for the marginal worker is below the outside wage, $w\,(z_t - n_0) < w^r$. Only when $w\,(z_t - n_0)$ is below ω the marginal worker quits, because the prospect of future wage increases no longer outweighs the cost of the current wage being below the outside wage. In the next sections we derive expressions for these option values, for the log-reservation wage ω, and for the hiring and firing threshold p^+ and p^-.

Define now the *seniority index*[5] r_t at time t as

$$r_t \equiv n_t - q. \tag{6}$$

Hence, the most senior worker has rank $q = 0$, and her seniority index is $r_t = n_t$, while the least senior worker at time t has $q = n_t$, and her seniority index is $r_t = 0$. The seniority index of worker i is the log of the ratio of the total number of workers

[5] *Seniority index* and *seniority* are used interchangeably throughout the text.

to the number of workers with greater or equal tenure than worker i; keeping total employment fixed, a 1% increase in your seniority is equivalent to the number of peer workers with greater or equal tenure decreasing by 1%. To see how wages are related to the seniority index r_t, equation (5) can be rearranged as

$$w(z_t - q) = \frac{\beta}{\eta} r_t + \beta \left(p_t - p^-\right) + \omega \tag{7}$$

Wages depend on the seniority index of the worker, r_t, and on the price of the firm's output, which varies between p^- and p^+. Hence, the return to seniority, β/η, is between zero and unity. As long as p_t is in between the upper and lower bounds p^- and p^+, the firm's employment, and hence the worker's seniority index, are constant. In this interval for p_t, variations in the market index z_t translate into fluctuations in p_t, or equivalently, in profits per worker.[6] Hence, there is a positive relation between wages and profits per worker, as documented e.g. by Abowd and Lemieux (1993). When $p_t = p^-$, a further decline in z_t induces the firm to fire some of its workers, so that n_t, and hence the seniority index r_t, go down. The lower r_t translates in a fall in the wage rate for the workers who are not laid off. The other way around, when $p_t = p^+$, a further increase in z_t induces the firm to hire additional workers, so that the seniority index of the incumbents goes up, which translates in an increase in the wage of incumbents.

[6] A 1 % increase in the price of output yields β % increase in the wage per worker, $0 < \beta < 1$. Since productivity is normalized to unity and since labor is the only factor of production, cost per unit of output increase also by β %. Hence, total revenue goes up more than total cost.

The firm size effect on wages has been extensively documented, see e.g. 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 r_t is driven by the evolution of the log firmsize n_t, since the value of her rank q is fixed over the duration of the job spell. 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 of a firm at the firing bound p^- satisfies

$$\begin{aligned} \mathrm{E}\left[w|p_t = p^-\right] &= \frac{1}{N_t}\int_{-\infty}^{n_t} w(z_t - q)e^q dq \\ &= \frac{1}{N_t}\int_{-\infty}^{n_t} \left[\omega + \beta/\eta\left(z_t - q - \eta p^-\right)\right] e^q dq = \omega + \beta/\eta, \end{aligned}$$

where in the first expression the factor e^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 of an increase in the market index z_t on the average log wage 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. A similar argument establishes that the average log wage of a firm at the hiring bound is also independent of firm size

$$E\left[w|p_t = p^+\right] = \omega + \beta/\eta + \beta\left(p^+ - p^-\right).$$

This reasoning implies that on the hiring and firing bounds both the average wage and profits per worker are fixed and fluctuations in the market index z_t translate into variation in the size of the workforce, while in between the two boundaries the reverse holds: employment is fixed, and fluctuations in z_t translate into variation in the average wage and profits per worker. Note that fluctuations in p_t (driving variation in profits per worker) and in the seniority index r_t are orthogonal, since when the one changes the other is constant, and the other way around. Hence, in a regression context, omitting either one as an explanatory variable does not affect the consistency of the parameter estimate for the other.

2.5 The worker's problem

The value of the parameter ω in equation (5) can be derived using the theory of option values, see Dixit and Pindyck (1994, 136–140)[7]. 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

[7] Dixit 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.

side of the problem, which yields expressions for the hiring and firing thresholds, p^- and p^+. 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

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

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

$$\begin{aligned} V(z_t - q) = & \frac{1}{r(\beta/\eta)} \exp\left[\omega + \frac{\beta}{\eta}\left(z_t - q - \eta p^-\right)\right] + A^- \exp\left[\lambda^-(z_t - q)\right] \\ & + A^+ \exp\left[\lambda^+(z_t - q)\right] \\ r(x) \equiv & \rho - \mu x - \frac{1}{2}(\sigma x)^2, \\ \lambda^{+,-} \equiv & \frac{-\mu \pm \sqrt{\mu^2 + 2\rho\sigma^2}}{\sigma^2}. \end{aligned} \tag{8}$$

where we substitute $w(z_t - q)$ from equation (5) 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 first inequality follows from equation (2). The first term on the right-hand side of equation (8) 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 reservation wage. The final two terms, $A^{+,-} \exp\left[\lambda^{+,-} z_t\right]$, 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. Hence, the option value of separation must converge to zero, which is the case for the negative root λ^-, since $\lim_{z\to\infty} A^- \exp\left[\lambda^- z\right] = 0$. 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 outside 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 separation after a small perturbation of z_t. This requires the first derivative of $V(z_t - q)$ to be zero. Elimination of A^- yields an expression for ω:

$$\omega = \ln r(\beta/\eta) - \ln\rho - \ln\left(1 - \frac{\beta}{\eta\lambda^-}\right). \tag{9}$$

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 outside 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.

2.6 The firm's problem

We now turn to the firm's optimal strategy. We observe that hiring a worker with rank q^* affects neither log marginal productivity[8] $mr(z_t - q)$, nor log wage $w(z_t - q)$, for workers with rank $q < q^*$ (the workers who have been hired before worker q^*).

[8] Note that the function $mr(z_t - q)$ 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 $t = 0$.

Similarly, whether or not the firm hires any workers $q > q^*$ after hiring worker q^* affects neither $mr\left(z_t - q^*\right)$, nor $w\left(z_t - q^*\right)$. 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 log marginal revenue $mr\left(z_t - q\right)$ and her log wage $w\left(z_t - q\right)$, 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 separately from 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 be the asset value of the firm for the N_t-th worker, that is, taking the optimal behavior of the firm for hiring workers with higher seniority as given. The Bellman equation for F satisfies

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

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

$$\begin{aligned} F\left(z_t - n_t\right) = & \frac{1}{r\left(\eta^{-1}\right)} \exp\left[\frac{1}{\eta}\left(z_t - n_t\right) - \pi\right] - \frac{1}{r\left(\beta/\eta\right)} \exp\left[w\left(z_t - n_t\right)\right] \\ & + B^- \exp\left[\lambda^-\left(z_t - n_t\right)\right]. \end{aligned} \tag{10}$$

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

$$G(z_t - n_t) = B^+ \exp\left[\lambda^+ (z_t - n_t)\right], \tag{11}$$

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 pasting conditions read

$$\begin{aligned} F\left(\eta p^-\right) &= G\left(\eta p^-\right), \\ F'\left(\eta p^-\right) &= G'(\eta p^-), \\ F\left(\eta p^+\right) &= G\left(\eta p^+\right) + I, \\ F'\left(\eta p^+\right) &= G'\left(\eta p^+\right). \end{aligned} \tag{12}$$

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 $z_t - n_t = \eta p^-$, 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 would not 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 $z_t - n_t = \eta p^+$: 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 (12) has a unique solution for* p^+, p^-, B^+, B^- *for which (i)* $p^+ - \pi > 0 > \omega > p^- - \pi, B^+ > 0, B^- > 0$*; (ii)* $\frac{\partial B^+}{\partial \beta} \leq 0$*; (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 outside wage, ρ^{-1}, the option value of a firm to hire an outside worker conditional on n_t and 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 (11). 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 quasi-rents should be shared between players according to their share in the cost of the specific investment. Since here the firm is assumed to bear the full cost, any share of the quasi-rents being assigned to workers leads to inefficiency. Hence, $\beta = 0$ is most efficient. If $\beta > 0$, firms will postpone hiring of new workers till the surplus of log marginal productivity $mr(z_t - q)$ above the log outside wage w^r is such that, even though the firm gets only part of the quasi-rents, 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 (12) yields

$$p^- - \pi = \ln r\left(\eta^{-1}\right) - \ln\rho - \ln\left(1 - \frac{1}{\eta\lambda^-}\right) + \ln\left(1 - \rho\frac{\lambda^+ - \lambda^-}{\lambda^-}B^+\exp\left[\eta\lambda^+ p^-\right]\right). \quad (13)$$

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 less attractive it is to fire a worker, see Proposition 3-(iv). Propositions 3-(iii) and (iv) imply that the higher is the workers' bargaining power β, the less volatile is employment, since it is insulated from shocks to the 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^-.

2.7 Unemployment and the welfare cost of hold up

The model implies that new jobs are rationed. This can be seen most easily by considering the wage of a worker who is just hired. This wage is higher than that of a worker who is at the borderline of being laid off, that is

$$w\left(\eta p^+\right) > w\left(\eta p^-\right) = \omega.$$

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 outside wage, $1/\rho$, the asset value of a worker who is just hired must be higher than $1/\rho$. To close the model, we therefore have to explain who gets hired by a firm and who does not. 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 must spend all her time searching and hence, she cannot produce as self employed. For simplicity, we assume that leisure has no value.[9] New jobs at firms arrive at a rate λ per unit of the labor force and are distributed randomly among the unemployed. Hence, the asset value of unemployment, V^U, satisfies

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

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. $V\left(\eta p^+\right) - V^U$ is the asset gain of getting a job offer. The level of unemployment follows from the no-arbitrage condition between self employment and unemployment

$$u = \lambda\left[V\left(\eta p^+\right) - \frac{1}{\rho}\right], \tag{14}$$

where we use $V^U = 1/\rho$, the asset value of self employment[10]. The higher the asset gain of getting a job at a firm, the higher is unemployment. Hence, there are two types

[9] Allowing 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 increase in unemployment.

[10] We assume: $u < 1$. If not, everybody would be self-employed, since even if everyone were searching for a job, firms would not open vacancies.

of inefficiency.[11] 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 costs of hiring a worker in net present value terms exceed the social costs 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 costs of rationing that dissipate workers' surplus. The no-arbitrage condition (14) implies that the workers as a group spoil their whole share of the quasi-rents in wasteful unemployment.

As discussed before, the inefficiency is due to a violation of the Hosios condition that the costs and quasi-rents 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 costs 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 costs of the specific investment as they get from its revenues, again satisfying the Hosios condition. It is useful to consider this

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

case more closely. The asset 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, independent of her rank q. Seniors getting higher rents than juniors at a particular point in time reflects the fact that they were able to realize the upside of the risky returns on their share in the costs of specific investment, I. Hence, in that case, the LIFO separation rule can be interpreted as a protection of the 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 costs 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 is so high that their share in this present value suffices to cover the costs of investment. These arguments imply that as long as we do not know what share of the costs of specific investment is born by workers, empirical evidence showing that there is a return to seniority in wages is inconclusive on whether employment is below its inefficient level or not. 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: 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.

2.8 Extensions

What would be the effect of firing costs in this world? We model firing costs as a wealth transfer L from the firm to the worker, at the moment of firing. This wealth transfer raises the value of the outside option of the worker. The value matching condition reads $V(\eta p^-) = \rho^{-1} + L$. Hence, the expression for ω reads (*cf.* equation (9))

$$\omega = \ln r(\beta/\eta) - \ln \rho - \ln\left(1 - \frac{\beta}{\eta\lambda^-}\right) + \ln(1 + \rho L). \tag{15}$$

Firing costs raise ω. The wealth transfer has two counteracting effects on the firing bound p^-: the direct effect of firing costs makes layoffs less attractive to the firm, while the indirect effect via higher wages makes layoffs more attractive, since workers are more costly. The first-order condition for optimal firing now reads (*cf.* equation (12))

$$F\left(\eta p^-\right) = G\left(\eta p^-\right) + L,$$

where we use the value of ω from equation (15) to account for the effect of firing cost on wages. Some calculations, see the Appendix, show that the value for p^- remains the same as in equation (13). The direct and the indirect effect cancel therefore exactly, except for the indirect effect via B^+, the option value of rehiring. On the hiring side, firing cost has two effects with the same sign: first, it raises wages, 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.

The trade unions are often viewed as a necessary condition to provide workers bargaining power.[12] However, the ability of workers to refuse to transfer their tacit knowledge to new hires is an alternative source of power. 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.[13] Hence, a LIFO scheme enables workers to exploit their individual bargaining power without having to solve their collective action problem. Whether the LIFO model is really a more appropriate description of a non-unionized than of a unionized environment is an empirical issue. For the United States, some scattered evidence is mixed. On the one hand, Medoff (1979) or Abraham and Medoff (1984) show that explicit LIFO layoff rules are far

[12] This problem has been suggested to us by Kevin Murphy.

[13] The 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.

more common in unionized firms, suggesting that the model discussed here applies to unionized rather than non-unionized firms. On the other hand, Teulings and Hartog (1998: 225) show that tenure profiles are flatter in unionized firms. To the extent that tenure proxies seniority, this suggests the opposite. This paper will not bring much clarity on this particular issue either, since our data refer to Denmark and Portugal, both countries where more or less all firms are unionized.

3 Empirical framework

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

1. The Last-in-First-Out separation rule: the workers hired last, leave the firm first.

2. A return to seniority in wages: a worker's wage depends on her seniority in the firm, that is, her tenure relative to that of her colleagues.

We shall analyze these implications empirically.[14] 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. Below we give a brief description of both data sets; an accompanying

[14] In the companion web appendix of this paper, we also verify that Gibrat's law for the evolution of the firm sizes holds in both countries, in particular for firms with more than 20 employees. This implication is not specific to 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.

online appendix has details on data cleaning and structuring, as well as information on the relevant labour market institutions in the two countries. Subsequently, we proceed with discussing the tests of the model's two implications.

3.1 Data description

For Denmark, we use the *Integrated Database for Labor Market Research* (IDA) for 1980-2001, from the official statistics bureau, Statistics Denmark, which has been often used previously, e.g. by Mortensen (2003). IDA tracks every single individual between 15 and 74 years old. The labor market status of each person is recorded at November 30 each year. 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. We use information on hourly gross 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 1964-1980 the tenure can be calculated from a second dataset on the contribution histories to ATP, a mandatory pension program. Tenure in job spells started before 1964 is left censored (less than 3% of the observations). Potential experience is age-schooling-6.

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	21.90	5.41	20.25	12.05	37.79	0.35	0.63	4.68
					(7.99)	(5.58)	(12.19)	(3.05)	(11.95)		(0.72)	(2.44)
PT 1991-2000	11420191	3211990	4268149	330270	3.68	8.43	20.26	6.81	36.73	0.40	0.87	4.04
					(2.52)	(8.61)	(11.39)	(3.65)	(11.18)		(0.85)	(2.14)
Cross-section year 2000												
DK 2000	1087922	1087922	1087922	79188	23.53	5.26	21.10	12.69	38.98	0.36	0.61	4.86
					(9.02)	(5.78)	(11.68)	(2.74)	(11.75)		(0.73)	(2.48)
PT 2000	1403686	1403686	1403686	184957	4.06	7.93	20.47	7.59	37.14	0.42	0.75	3.64
					(2.75)	(8.54)	(11.42)	(3.83)	(11.08)		(0.81)	(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 parentheses.)

For Portugal, we use *Quadros de Pessoal* for 1991-2000, provided by the Ministry of Labour and Social Solidarity, also often used before, e.g. by Cabral and Mata (2003). It is based on a compulsory survey of firms, establishments and all their workers; compulsory participation enhances the data quality. The information available is similar to Denmark's except that workers' tenure is directly reported; the industry is that industry with the highest share of firm's sales or, when allocation by sales is not possible, the industry with the highest employment share. We use all full-time employees in their main job, aged 15-74, and working for a firm located in Portugal's mainland.

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[15] are presented in Table 1, both for the pooled data and for 2000 separately. There are several obvious differences between the two countries. For instance in 2000, 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.

Table 2 reports some summary statistics and correlation patterns for the main

[15] In the online appendix of the paper summary statistics for broad category industries are also presented.

variables of interest, tenure, log firmsize, rank and seniority. For the purpose of our analysis, it is important to note that rank is not perfectly related to either tenure or log firmsize. We return to this issue in Section 3.3, when discussing the identification of the return to seniority.

Table 2: Correlations tenure, logfsize, rank, seniority

	Denmark				Portugal			
	tenure	logfsize	rank	seniority	tenure	logfsize	rank	seniority
A: All observations								
Mean	5.41	4.68	4.05	0.63	8.43	4.04	3.17	0.87
Std dev	5.58	2.44	2.45	0.72	8.61	2.14	2.17	0.85
Corr rank	-0.09	0.96	1.00	-0.16	0.48	0.92	1.00	-0.23
Corr seniority	0.77	0.13	-0.16	1.00	-0.46	0.16	-0.23	1.00
B: Deviations from individual means within job spells								
Std dev	2.64	0.22	0.33	0.37	1.94	0.19	0.35	0.30
Corr rank	-0.75	0.17	1.00	-0.80	0.02	0.63	1.00	-0.87
Corr seniority	0.76	0.46	-0.80	1.00	-0.02	-0.16	-0.87	1.00
C: Individual first differences within job spells								
Mean	1.00	0.01	-0.14	0.16	1.00	-0.04	0.10	-0.14
Std dev	0.00	0.23	0.19	0.23	0.00	0.20	0.41	0.32
Corr rank		0.43	1.00	-0.40		0.62	1.00	-0.87
Corr seniority		0.64	-0.40	1.00		-0.15	-0.87	1.00

Labels: Corr=correlation; Std dev=standard deviation. For Denmark the pooled sample is 1980-2001, for Portugal 1991-2000. The correlation rows with the rank and seniority are computed using all observations with common nonmissing values for all four variables.

3.2 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 wether or not the LIFO rule applies. Analogous to Section 2.4, q_{ijt} is defined as the log number of workers employed at firm j at time t with a tenure greater or equal than that of worker i; hence q_{ijt} includes worker i herself. The seniority index r_{ijt} is defined as in equation (6):

$$r_{ijt} \equiv n_{jt} - q_{ijt}. \tag{16}$$

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 category but not necessarily for other skill categories. This may disrupt a strict application of the LIFO separation rule. Second, workers separate not only due to shocks to 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. When establishing that there is a return to seniority in Section 3.3, we benefit from this independent source of variation by using q_{ijt} as an instrument for r_{ijt}. In this section, we just have to account for the fact that the LIFO rule does not perfectly explain the order of separation. We intend to show that r_{ijt} has an impact on the job separation rate, over and above the impact of tenure. Therefore, we model the transition process by a mixed proportional hazard model with discrete time periods. The probability of leaving the firm conditional on an elapsed incomplete tenure of T_{ijt} years (i.e. the hazard rate) can be written as:

$$\theta(r_{ijt}, Z_{ijt}, T_{ijt}, v_i) = \frac{\exp\left(\gamma_0 r_{ijt} + \gamma_1 Z_{ijt} + \psi_{T_{ijt}} + v_i\right)}{1 + \exp\left(\gamma_0 r_{ijt} + \gamma_1 Z_{ijt} + \psi_{T_{ijt}} + v_i\right)} \tag{17}$$

where Z_{ijt} is a vector of observed characteristics of the worker and the job (education, experience, occupation, region and industry indicators) 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 γ_0 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 γ_0 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.[16] In addition, short spells are underrepresented since a worker has to stay at least till the next period of observation for a spell to be recorded. With the data at hand, we cannot correct for these problems.

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 run separate regressions 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. Deleting the left-censored spells implies that we have a maximum of 22 years of tenure in Denmark and 10 for Portugal.

Table 3 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.[17] 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 order, since the firm's workforce is likely to

[16] For Portugal, tenure is reported in months. We use this information in the estimation. For the rest, the modelling is identical to that for Denmark.

[17] The detailed estimation results are presented in the accompanying web appendix of this paper.

be heterogeneous, with separate LIFO ordering applying 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 γ_0, 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 index. For example, some separations might be driven by the worker and the firm learning about the quality of the match, see e.g. Jovanovic (1979). These separations do not fit the LIFO pattern.

Table 3: Main results LIFO test

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

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

3.3 The return to seniority

3.3.1 General specification

The return to seniority in wages 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}

$$w_{ijt} = \alpha + \chi X_{ijt} + \gamma T_{ijt} + \upsilon Q_{jt} + \delta r_{ijt} + \zeta n_{jt} + \varepsilon_{ijt}, \tag{18}$$

where X_{ijt} is experience and where Q_{ijt} is the industry (or: occupation) specific experience, cf. Neal (1995). We omit higher order terms in experience and tenure, and other controls (including time effects) from equation (18) 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 effect[18]

$$\varepsilon_{ijt} = \varphi_{ij} + \psi_j + \mu_i + \nu_{ijt}. \tag{19}$$

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). Learning and search theories imply that 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 positive correlation between $\varphi_{ij} + \psi_j + \mu_i$ and T_{ijt}. Search theories imply 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 positive correlation between $\varphi_{ij} + \psi_j$ and T_{ijt}. There are two obvious solutions to this problem, either

[18] This 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.

within-job first differencing (FD) or adding fixed effects for every job spell (FE), eliminating the observables X_{ijt} and Q_{ijt} and the unobservables ϕ_{ij}, ψ_j, and μ_i from the specification. First differencing yields

$$\Delta w_{ijt} = \chi + \gamma + \upsilon + \delta \Delta r_{ijt} + \zeta \Delta n_{jt} + \Delta \nu_{ijt}. \tag{20}$$

Adding fixed effects per job spell is equivalent to estimating (18) by taking deviations from the mean over time, within a job spell:

$$\widetilde{w}_{ijt} = (\chi + \gamma + \upsilon)\,\widetilde{T}_{ijt} + \delta \widetilde{r}_{ijt} + \zeta \widetilde{n}_{jt} + \widetilde{\nu}_{ijt}, \tag{21}$$

where the upper tilde denotes deviations from the mean per job spell, e.g. $\widetilde{w}_{ijt} = w_{ijt} - \overline{w}_{ijt}$, with $\overline{w}_{ijt}$ the mean of w_{ijt} for a job spell. We exclude $\widetilde{X}_{ijt}$ from (21) because it is perfectly collinear with $\widetilde{T}_{ijt}$. In both specifications above, it is immediately clear that the first-order effects of tenure (either firm, or industry, or occupation specific) 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 within-spell variation. Hence, researchers had to revert to between job spell variation. 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}, as noted above in Table 2. 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 the estimation of the return to tenure.

There is a second identification issue. In the theoretical model, workers live forever and thus the log number of workers with tenure greater or equal than the tenure of worker i, q_{ijt}, does not depend on t, see equation (6). In that case, the only source of within job spell variation in the seniority index r_{ijt} is log firm size n_{jt}. Hence, the coefficients of these variables, δ and respectively ζ, are not separately identified when using only within job spell variation. As discussed in Section 3.2 and Table 2, q_{ijt} does depend on t empirically, mainly because more senior workers than worker i retire at some point. Hence, we can use q_{ijt} as an instrument for r_{ijt}. Substituting the definition of r_{ijt}, equation (16), in equation (20) yields

$$\Delta w_{ijt} = \chi + \gamma + \upsilon - \delta \Delta q_{ijt} + (\zeta + \delta)\, \Delta n_{jt} + \Delta \nu_{ijt}.$$

This equation makes clear that the effect of the seniority index on wage growth is identified from variation in q_{ijt}. *Mutatis mutandis*, the same argument applies for the FE method, see equation (21).

The choice between the FE and FD estimators above depends on the error structure of υ_{ijt}. The closer is υ_{ijt} to a unit root, the more efficient is the FD method; the closer υ_{ijt} is to being serially uncorrelated, the more efficient is the FE method.

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 (20) 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 (21) 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 (21) than from using (20).[19]

Table 4: Residual Autocovariances for Within-Job LogWage Innovations

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

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

[19] We report robust standard errors, so that correlation between the residuals over time does not affect the validity of the standard errors.

First, we check the characteristics of the dynamic process of v_{ijt}. Table 4 reports the variance-covariance of Δv_{ijt}. 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 of 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.[20] These numbers are of the same order of magnitude as found for the United States.

This evidence suggests that in terms of efficiency we prefer FD, while in terms of allowing for a lagged effect of Δr_{ijt} on Δw_{ijt}, we prefer FE. Hence, we report both the FD and FE estimation results. Our regressions control for up to quartic terms in tenure and experience, log firmsize, and industry, occupation, and region dummies[21]. In Table 5 we report the results. We present the estimation results for two specifications, one excluding 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

[20] Let 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, $\mathrm{Var}(\Delta v_{ijt}) = \mathrm{Var}(u_{ijt}) + 2\mathrm{Var}(q_{ijt})$ and $\mathrm{Cov}(\Delta v_{ijt}, \Delta v_{ij,t-1}) = -\mathrm{Var}(q_{ijt})$, so that $\mathrm{Var}(u_{ijt}) = \mathrm{Var}(\Delta v_{ijt}) + 2\mathrm{Cov}(\Delta v_{ijt}, \Delta v_{ij,t-1})$.

[21] Time effects are also accounted for in all subsequent wage analyses, by time-detrending wages prior to the regressions.

the coefficients for tenure and log firmsize by $5 - 30\%$. The coefficients for experience are hardly affected by including seniority[22]. The effect of tenure and log firmsize on wages is at least partly a proxy for the effect of seniority. Seniority is expected to be measured with greater measurement error than tenure and firmsize. 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. Still, 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 feeds into seniority. Furthermore, seniority is affected by measurement errors due to the fact that separate seniority statistics might apply for subgroups of the workforce. The actual effect of seniority on wages is therefore likely to be underestimated, and 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.

[22] In the FD and FE specifications from above these effects can be seen for tenure and experience in the higher order polynomial terms. In the accompanying web appendix of this paper we also present the OLS estimates, with and without including seniority index in the regressions, where differences under the two specifications in the linear experience and tenure terms can be noticed.

Table 5: FE and FD Wage Regressions for the Entire Private Sector in Denmark and Portugal

	Denmark				Portugal			
	FD1	FD2	FE1	FE2	FD1	FD2	FE1	FE2
sen. index		.003***		.008***		.016***		.022***
		(.0003)		(.0003)		(.0005)		(.0005)
logfsize	.013***	.011***	.026***	.021***	.025***	.015***	.040***	.028***
	(.0002)	(.0003)	(.0003)	(.0003)	(.0004)	(.0005)	(.0004)	(.0004)
tenure+exper	.047***	.045***	.010***	.007***	.068***	.065***	.059***	.055***
	(.0003)	(.0003)	(.0001)	(.0002)	(.0005)	(.0005)	(.0003)	(.0003)
$tenure^2$	.191***	.199***	-.052***	-.036***	-.086***	-.069***	-.083***	-.067***
	(.002)	(.002)	(.002)	(.002)	(.003)	(.003)	(.002)	(.002)
$tenure^3$	-.101***	-.105***	.014***	.008***	.027***	.021***	.024***	.019***
	(.001)	(.001)	(.099)	(9.88e-07)	(.001)	(.001)	(.0007)	(.0007)
$tenure^4$	.002***	.002***	-.0009***	-.0002	-.003***	-.002***	-.003***	-.002***
	(.0002)	(.0002)	(.0002)	(.0002)	(.0002)	(.0002)	(.00009)	(.00009)
$exper^2$	-.224***	-.223***	.099***	.100***	-.204***	-.204***	-.149***	-.147***
	(.002)	(.002)	(.0006)	(.0006)	(.004)	(.004)	(.002)	(.002)
$exper^3$	.039***	.039***	-.039***	-.039***	.043***	.043***	.030***	.029***
	(.0007)	(.0007)	(.0002)	(.0002)	(.001)	(.001)	(.0007)	(.0007)
$exper^4$	-.003***	-.003***	.004***	.004***	-.003***	-.003***	-.002***	-.002***
	(.00007)	(.00007)	(.00003)	(00003)	(.0001)	(.0001)	(.00007)	(.00007)
N obs	14907897		22364083		5758655		10743244	
Workers	2116307		2771627		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 seniority index 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).

Are there other theories that might explain our results? For the answer to this question, it is important to realize that the effect of the seniority index r_{ijt} is identified from variation in the log number of workers with tenure greater or equal than the tenure of worker i, q_{ijt}. Any alternative theory explaining our findings must therefore be able to explain why the wage of a worker goes up when q_{ijt} goes down. This rules out a host of theories able to explain why r_{ijt} shows up significantly, but not able to explain why q_{ijt} does so. For example, the effect of q_{ijt} cannot be rationalized by learning models such as Jovanovic's (1979), or search models like Postel-Vinay and Robin (2002) and Burdett and Coles (2003), because these theories relate in no way the wage of worker i to the number of more senior workers in the firm.

One might attempt to explain our findings by tournament models like Lazear and Rosen (1981) or Malcomson (1984). In Malcomson's model, workers' output is non-verifiable, so that the firm's promise to award a performance pay is non-credible. This credibility problem can be resolved by the firm promising e.g. to promote the best one out of every four workers to a better paid senior position within two years. Hence, running a tournament allows the firm to implement a credible system of performance pay. In its original form, Malcomson's model yields an association between tenure and wages, but does not yield an association between q_{ijt} and wages. However, an amended version of the model might do. The firm would state its promise somewhat differently, namely, that it has some well-paid top positions in the hierarchy. It promises that as soon as one of them becomes vacant (e.g. by retirement), it will be filled by the best worker in the firm at a lower position. On average, this compensation system yields a positive association between q_{ijt} and wages. From a theoretical perspective, the problem with this version of the tournament model is that it is inferior to Malcomson's original model, since the promotion-bonus is conditional on the retirement of a senior worker, which is a more risky bonus than the unconditional bonus, and hence this conditional bonus is a less efficient incentive when juniors are risk averse. One explanation is that the firm seeks to combine both models, LIFO layoff and the tournament: if the firm has to pay its workers some rents, than it can better use these payments to provide better incentives. Empirically, we do not see an easy way to discriminate between this specific form of the tournament model and the LIFO layoff model. A related explanation suggested to us argues that a firm needs "juniors" and "seniors" in fixed proportions, for example, because juniors do the production work, while seniors do the planning and coordination. The firm fills its vacancies for seniors by simply "promoting" some of its juniors to senior positions. The problem with this explanation is that it does not explain why the firm does not hire "seniors" on the outside market, and, even more problematic, why it would pay higher wages to these "seniors", apart from providing incentives for juniors, which, however, brings us back to Malcomson's tournament model.

Commentators have suggested other, more informal, theories to us. Maybe the firm has an "iron stock" of workers who are never fired, and a fluctuating stock of temporary workers, receiving lower pay. Were this true, it would generate a positive association between your seniority and your wage. However, this theory would not generate a positive association between the *change* in the seniority of an individual worker and the *change* in her wage. Stating the same argument differently, controlling for fixed worker effects as we do in both the FD and FE regressions deals with this alternative. Similarly, any theory that starts from the argument that firms are heterogeneous, some firms requiring more specific investments than others, fails to explains our results. Firms are indeed heterogeneous, but fixed firm effects should deal with these theories.

3.3.2 Returns to seniority for subgroups

The LIFO layoff rule is unlikely to apply unconditionally. One would expect the firm to apply separate layoff rules for different subgroups of its workforce. For example, a construction firm is unlikely to fire its secretaries if it has an excess supply of bricklayers, whatever the seniority of both groups of workers. One can therefore expect the theory to work better using separate seniority indices for broad groups of workers in our regression. One could calculate separate indices for each occupation. Apart from the fact that we do not have a good classification of occupations in our data, we hesitate to distinguish workers by their occupation, since that might change endogenously. The wage increase due to a rise in the worker's seniority might go hand in hand with a relabelling of her occupation. However, we can distinguish workers by typically time-invariant characteristics, unrelated to their actual job. Hence, we repeat the analysis separately for males and females, and respectively for low- and high-educated workers. The results are reported in Table 6. 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. 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 FE estimate is even significantly negative for Denmark, though small in absolute value. The impact of seniority for the highly educated workers 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 6: 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
sen. index	.005***	.005***	.005***	.010***	.015***	.019***	.014***	.019***
	(0.0004)	(0.0005)	(0.0004)	(0.0004)	(0.0007)	(0.0006)	(0.0007)	(0.0006)
logfsize	.002***	.014***	.014***	.025***	.015***	.028***	.019***	.031***
	(0.0005)	(0.0005)	(0.0004)	(0.0004)	(0.0007)	(0.0006)	(0.0006)	(0.0006)
(ten+exp)	.032***	.009***	.052***	.007***	.053***	.042***	.080***	.073***
	(0.0004)	(0.0002)	(0.0004)	(0.0002)	(0.0007)	(0.0005)	(0.0007)	(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
sen. index	.010***	.020***	.002***	-.002***	.029***	.032***	.013***	.016***
	(.0003)	(.0004)	(.0004)	(.0004)	(.002)	(.002)	(.0005)	(.0005)
logfsize	.007***	.016***	.014***	.024***	.026***	.026***	.016***	.031***
	(.0004)	(.0004)	(.0005)	(.0005)	(.002)	(.002)	(.0005)	(.0004)
(ten+exp)	.040***	.006***	.031***	.006***	.116***	.099***	.056***	.049***
	(.0004)	(.0002)	(.0007)	(.0002)	(.002)	(.001)	(.0005)	(.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. Seniority index 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).

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 reduction in the number of workers more senior than you raises your wage by 0.01 to 0.02%. This return is higher for higher educated workers, and it is twice as high in Portugal than it is in Denmark. Some 5 to 30% of what has been know as the return to tenure and to firmsize is in fact a return to seniority. A number of familiar theories explaining the tenure profiles in wages, like Jovanovic's (1979) learning model or the search models by Postel-Vinay and Robin (2002) and Burdett and Coles (2003), cannot explain this finding. Our results are not at odds with these theories, but these earlier theories cannot be the full story, because they 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. Our results might be explained by an amended version of Malcomson's (1984) tournament model, where the number of prizes in the tournament depends upon the number of seniors that leave the firm. However, it is hard to explain why the firm would use this type of conditioning on the number of prizes, without referring to some sort of rent sharing, as in our LIFO layoff model. A return to seniority implies that a worker is to some extent shareholder in her own firm. Hence, the theory here links labor economics to 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. This conclusion is conditional on the assumption that the firm bears the full cost of specific investment. This assumption has not been tested 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. Efficiency requires the costs and the benefits of the specific investment to be born by the same party, which is the Hosios condition. When workers are risk averse, efficiency can only be obtained when this is the firm, since any other allocation assigns a risky return to a risk averse player. Hence, 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.

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. On the one hand, returns to seniority might be largely due to legal institutions, and these institutions vary across countries. On the other hand, the economic mechanism for having a LIFO layoff rule exists everywhere and the legal institutions might just be a formalization of rules of conduct and implicit contracts that would have emerged anyway. There is only one way to investigate this, which is to repeat our empirical analysis for the United States.

5 References

Abowd, J.M. and D. Card (1989), "On the covariance of earnings and hours changes", Econometrica 57(2), 411–445.

Abowd, J.M. and T. Lemieux (1993), "The effects of product market competition on collective bargaining agreements: the case of foreign competition in Canada", Quarterly Journal of Economics, 108(4), 983–1014.

Abraham, K.G. and J.L. Medoff (1984), "Length of service and layoffs in union and nonunion work groups", Industrial and Labor Relations Review, 38(1), 87-97

Altonji, J.G. and R.A. Shakotko (1987), "Do wages rise with seniority?", Review of Economic Studies, 54(3), 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, 111(5), 1043–1075.

Brown, C. and J. Medoff (1989), "The employer size wage effect", Journal of Political Economy, 97(5), 1027–1059.

Buhai, I.S. and C.N. Teulings (2006), "Tenure profiles and efficient separation in a stochastic productivity model", Tinbergen Institute DP 05–099/3.

Burdett, K. and M. Coles (2003), "Equilibrium wage-tenure contracts", Econometrica, 71(5), 1377–1404.

Cabral, L.M.B. and J. Mata (2003), "On the evolution of the firm size distribution: facts and theory", American Economic Review 93(4), 1075–1090.

Dixit, A.K. and R.S. Pindyck (1994), "Investment under uncertainty", Princeton University Press, New Jersey.

Dixit, A.K. (1989), "Entry and exit decisions under uncertainty", Journal of Political Economy, 97(3), 620–638.

Hosios, A.J. (1990), "On the efficiency of matching and related models of search and unemployment", Review of Economic Studies, 57(2), 279–298.

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(7), 473–508.

Kuhn, P. and J. Robert (1989), "Seniority and distribution in a two-worker trade union", Quarterly Journal of Economics, 104(3), 485–505.

Lazear, E. and S. Rosen (1981), "Rank order tournaments as optimal salary schemes", Journal of Political Economy, 89(5), 841-864.

Lindbeck, A. and D. Snower (1990), "Do cooperation and harassment explain involuntary unemployment?", American Economic Review, June, 80(3), 167–188.

Malcomson, J.M. (1984), "Work incentives, hierarchy, and internal labor markets", Journal of Political Economy, 92(3), 486–507.

Medoff, J.L. (1979), "Layoffs and alternatives under trade unions in U.S. manufacturing", American Economic Review, 69(3), 380-395

McLaughlin, K. J.(1991), "A theory of quits and layoffs with efficient turnover," Journal of Political Economy, 99(1), 1–29.

Mortensen, D.T. (2003), Wage dispersion: why are similar workers paid differently?, MIT Press, Cambridge.

Neal, D. (1995), "Industry-specific human capital: Evidence from displaced workers", Journal of Labor Economics, 14(4), 653-677

Postel-Vinay, F. and J.M. Robin (2002), "Equilibrium wage dispersion with worker and employer heterogeneity", Econometrica, 70(6), 2295–2350.

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, New York.

Topel, R.H. (1991), "Specific capital, mobility and wages: wages rise with job seniority", Journal of Political Economy, 99(1), 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.

A Derivation

Substitution of equation (5), (9), (10), and (11) in equation (12) 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

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

Elimination of C^- from the first two equations of this system yields equation (13). 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 a solution with $\Delta > 0$. The second equation of this system can be written as

$$\begin{aligned} I \cdot R\left(\Delta\right) &= \psi \cdot S\left(\Delta, \beta\right), \qquad (22) \\ R\left(\Delta\right) &\equiv -\left(\mu^+ - \mu^-\right) E + \left(\mu^+ - 1\right) \mu^- G + \left(1 - \mu^-\right) \mu^+ H, \\ S\left(\Delta, \beta\right) &\equiv \left(\mu^+ - 1\right)\left(\beta - \mu^-\right)\left(FG + EH\right) + \left(1 - \beta\right)\left(\mu^+ - \mu^-\right)\left(EF + GH\right) \\ &\quad - \left(1 - \mu^-\right)\left(\mu^+ - \beta\right)\left(FH + EG\right), \\ R\left(0\right) &= 0, R_\Delta\left(0\right) = \left(1 - \mu^-\right)\left(\mu^+ - 1\right)\left(\mu^+ - \mu^-\right) > 0, R\left(\Delta\right) \geq 0, \\ S\left(0, \beta\right) &= 0, S_\Delta\left(0, \beta\right) = 0, S_{\Delta\Delta}\left(0, \beta\right) > 0, S\left(\Delta, \beta\right) \geq 0. \end{aligned}$$

Hence

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

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

$$\begin{aligned}
R &= \frac{-(\mu^+-\mu^-)\beta F + (\mu^+-\beta)\mu^- G + (\beta-\mu^-)\mu^+ H}{R(\Delta)}\psi, \qquad (23)\\
C^- &= \frac{-(\mu^+-\beta)E + (\mu^+-1)\beta F + (1-\beta)\mu^+ H}{R(\Delta)}\psi,\\
C^+ &= \frac{(\beta-\mu^-)E - (1-\mu^-)\beta F + (1-\beta)\mu^- G}{R(\Delta)}\psi.
\end{aligned}$$

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 (22) with respect to Δ and β yields an expression for $d\Delta/d\beta$. Totally differentiating equation (23) 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 $\frac{\partial p^-}{\partial\beta} < 0$.■