A Feasible Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation

Jesper Bagger*
University of Aarhus

Francois Fontaine†
Université de Strasbourg 3

Fabien Postel-Vinay‡
University of Bristol
and
CREST-INSEE

Jean-Marc Robin§
EUREQua-Université de Paris 1,
University College London
and
Institute for Fiscal Studies

April 2006 — IN PROGRESS

Abstract

We present a tractable equilibrium job search model of individual worker careers allowing for human capital accumulation, employer heterogeneity and individual-level shocks. We estimate our structural model on a panel of Danish matched employer-employee data and use it to analyze the determinants of wage dispersion and individual wage dynamics. Our main motivation for doing this is to quantify the respective roles of human capital accumulation coming along with work experience and the forces of labor market competition activated by workers' job search behavior in shaping individual labor earnings dynamics over the life cycle. We shall pay particular attention to the definition of differences between the returns to work experience as opposed to the returns to tenure at a particular firm.

Keywords: Job Search, Human Capital Accumulation, Individual Shocks, Structural Estimation, Matched Employer-employee Data.
JEL codes:

* jbagger@econ.dk. Visiting EUREQua-Université de Paris I.

† Francois.Fontaine@urs.u-strasbg.fr. Also affiliated with IZA.

‡ Fabien.Postel-Vinay@bristol.ac.uk. Also affiliated with PSE (*Paris-Jourdan Sciences Economiques*, a CNRS-EHESS-ENPC-ENS joint research unit), CEPR, CMPO and IZA.

§ jmrobin@univ-paris1.fr. Also affiliated with CEPR and IZA.

1 Introduction

We present a tractable equilibrium job search model of individual worker careers allowing for human capital accumulation, employer heterogeneity and individual-level shocks. We estimate our structural model on a panel of Danish matched employer-employee data and use it to analyze the determinants of wage dispersion and individual wage dynamics. Our main motivation for doing this is to quantify the respective roles of human capital accumulation coming along with work experience and the forces of labor market competition activated by workers' job search behavior in shaping the profile of individual labor earnings over the life cycle. We shall pay particular attention to the definition of differences between the returns to work experience as opposed to the returns to tenure at a particular firm.

Our contribution is related to three major strands of the empirical labor literature. The first one covers the impressively large body of work building on Mincer's (1974) original specification of log-earnings as a function of individual schooling and experience. In their recent comprehensive review of the implications of Mincer's "stylized facts" for post-schooling wage growth in the US, Rubinstein and Weiss (2005) put human capital accumulation and job search forward as potential driving forces of the observed earnings/experience profile.[1] As these authors further note, the obvious differences between these two lines of explanation in terms of policy implications (concerning schooling and training on one hand and labor market mobility on the other) are enough to motivate a thorough assessment of their relative quantitative importance. Rubinstein and Weiss (2005) then go on to take a detailed look at the available US evidence and find support for both approaches, thus calling for the construction of a unified model. One of our contributions in this paper is to offer such a model.[2]

[1]Rubinstein and Weiss (2005) also point to learning about job, worker or match quality as a third potential line of explanation for the observed earnings/experience profile. Learning is formally absent from our structural model. However, as we shall briefly argue below, the impact of learning may be partly captured by a pattern of individual-level shocks that is consistent with our model.

[2]Earlier combinations of job search and human capital accumulation include Rubinstein and Weiss (2005) and Barlevy (2005). These two models are based on a wage posting approach which has restrictive implications for individual earnings dynamics, particularly within job spells (see the discussion below). Moreover, Rubinstein's and Weiss' analysis of their model is essentially qualitative, and Barlevy provides a structural estimation based on NLSY data. From this latter standpoint, our contribution is to use matched employer-employee data and to put strong emphasis on firm heterogeneity.

The second body of empirical work to which the present paper relates is the (equally large) literature on individual earnings dynamics. The long tradition of fitting flexible stochastic decompositions to earnings data has proved very useful in documenting the statistical properties of individual earnings from a dynamic perspective (Alvarez, Browning and Ejrnæs, 2001; Meghir and Pistaferri, 2004; Guiso, Pistaferri and Schivardi, 2005). However, the specific economic mechanisms underlying these empirical models remain vague, making it difficult to interpret the empirical evidence that they provide, or even to agree on what the "true" earnings process should be. The comparative analysis of Baker (1997) provides a very convincing illustration of this point. Also, a famous example where the lack of explicit structural foundations may have added to the heat and confusion of the academic debate is the attempts at disentangling the the effects of job tenure versus experience on wage growth. The available empirical evidence on this important question is mixed. Using data from the PSID, Altonji and Shakotko (1987) find small tenure effects, while Topel (1991) and Buchinsky et al. (2002) find large tenure and experience effects, also in PSID data. More recently, the comparative study by Beffy et al. (2005) reports that returns to tenure are large in the US and small in France, a discrepancy which they attribute to French-US differences in labor market mobility. In this paper we take a further step in the direction of structural modeling: our explicit formalization of the joint impact of experience (through human capital accumulation) and labor market mobility (through interfirm competition) on individual wages offers well-defined (if specific) measures of the returns to experience and tenure. We are thus able to base our discussion of the relative returns to tenure and experience on those well-defined theoretical concepts, for which we exhibit precise empirical measures.

Finally, the third strand of literature to which this paper contributes is that on equilibrium job search. Empirical job search models have so far been geared to the description of "cross-sectional" aspects of the data, notably wage dispersion or the wage-productivity relationship. Yet these models are inherently dynamic and have strong predictions about the process followed by individual wages over time. What little attention has been paid to those predictions has revealed that, in the absence of individual-level shocks, job search models fail to accommodate the observed downward wage flexibility.[3] Unfortunately, introducing individual shocks into a job search model with a wage setting mechanism that is both theoretically and descriptively appealing, while keeping the model empirically tractable turns out to be a difficult undertaking. In a recent attempt, Postel-Vinay and Turon (2005) investigate whether the combined assumptions (taken up from the Postel-Vinay and Robin (2002) sequential auctions model) of (on-the-job) search and wage renegotiation by mutual consent can act as a realistic "internal propagation mechanism" of i.i.d. productivity shocks. Using British household data from the BHPS, they find that these assumptions transform purely transitory productivity shocks into persistent wage shocks with a covariance structure that is consistent with the data. In this paper we consider a richer empirical specification by allowing for non-i.i.d. individual shocks, human capital accumulation, and by explicitly modelling employer heterogeneity (which our use of matched firm-worker data enables us to do). We circumvent the aforementioned theoretical difficulties by restricting wages to be set as *piece-rate* contracts specifying the share of output received by the worker as a wage. We further allow firms to respond to outside offers by increasing the piece rate following a Bertrand bidding game, in a similar way to Postel-Vinay and Robin (2002). With these restrictions imposed, the model delivers a structural wage equation similar to the standard human capital wage equation with worker and employer fixed effects, human capital effects and stochastic dynamics caused by (i) between-firm competition for the workers' services (activated by on-the-job search) and (ii) idiosyncratic individual productivity shocks.

We estimate our structural model on Danish matched employer-employee data using indirect inference (Gouriéroux, Monfort and Renault, 1993).[4] [OUTLINE OF THE FINDINGS].

The paper is organized as follows. In section 2 we spell out the details of the theoretical model

[3]More precisely, the original Burdett and Mortensen (1998) wage posting model essentially assumes strict downward wage rigidity. Extensions of this model by Barlevy (2005) and Rubinstein and Weiss (2005) attribute all within-spell wage changes (upward or downward) to exogenous individual shocks. The sequential auctions model of Postel-Vinay and Robin (2002) is consistent with endogenous wage increases within or between job spells and with endogenous wage cuts concomitant with job changes, but still predicts downward wage rigidity within a given job spell. See Postel-Vinay and Robin (2006) for a review of these issues.

[4]We argue below that indirect inference is particularly well suited for our purposes, due in particular to the freedom it leaves for the choice of the auxiliary model which is used to generate the moments that we ultimately aim to match with the structural model. Our choice of auxiliary models will reflect standard specifications in the three strands of literature mentioned earlier, thus establishing a natural link and a basis for comparison between these and our structural approach to modelling individual careers and earnings dynamics.

and in section 3 we present the data and the empirical strategy.

2 The model

2.1 The environment

Basics. We consider a labor market where a unit mass of workers face a continuum of identical firms producing a multi-purpose good which they sell in a perfectly competitive market. Time is discrete and the economy is at a steady state. For reasons that will shortly become clear, we consider individual (as opposed to calendar) time, which we index by t. Workers can either be unemployed or matched with a firm. They will transit between employment and unemployment as well as from job to job following a search process to be defined momentarily. Firms operate constant-return technologies and are modeled as a collection of job slots which can either be vacant and looking for a worker, or occupied and producing.

Production technology. Log-output per period, $y_t = \ln Y_t$, in a firm-worker match involving a worker with experience t is defined as

$$y_t = p + h_t, \tag{1}$$

where p is a fixed firm heterogeneity parameter and h_t is the amount of efficient labor the worker with experience t supplies in a period. It is defined as follows:

$$h_t = \alpha + g_t + \varepsilon_t, \tag{2}$$

where α is a fixed worker heterogeneity parameter reflecting permanent differences in individual productive ability, g_t is a state-dependent deterministic trend reflecting human capital accumulation on the job, and ε_t is a zero-mean shock. This latter shock is worker-specific, and we only restrict it to follow a first-order Markov process.[5] A useful benchmark may be to think of it as a linear AR(1) process, possibly with a unit root.

[5] At this point we do not attach any more specific interpretation to the ε_t shock. It reflects stochastic changes in measured individual productive ability that may come from actual individual productivity shocks (due to preference shocks, labor supply shocks, technological shocks and the like), or from public learning about the worker's quality.

Timing of events within the period. The set of random events affecting a worker within a typical period includes retirement, job destruction, job offer arrival, and productivity shock. These four shocks are revealed in the following order:

1. *Productivity shocks*: At the beginning of the period, for any employed worker, ε_t is revealed, the worker's experience increases from $t-1$ to t and her/his productivity is updated from h_{t-1} to h_t as per equation (2). We assume that unemployed workers do not accumulate experience, so that if a worker becomes unemployed at an experience level of $t-1$, her/his productivity stagnates at h_{t-1} for the duration of the ensuing spell of unemployment.

2. *Production and payments*: Then, production takes place and firms pay workers their salaries.

3. *Job mobility shocks*: At the end of the period any employed worker leaves the market for good with probability μ OR sees her/his match dissolved with probability δ OR receives an outside offer with probability λ_1 ($\lambda_1+\delta+\mu \leq 1$). Similarly, any unemployed worker finds a new match with probability λ_0. Upon receiving a job offer, any worker (regardless of her/his employment status or human capital) draws the type p of the firm from which the offer emanates from a continuous, unconditional sampling density $f(\cdot) = F'(\cdot)$, with support $[p_{\min}, p_{\max}]$.

2.2 The wage equation

Wage setting rules. Wages are defined as *piece-rate contracts*. If a worker supplies h_t units of efficient labor and produces $y_t = p + h_t$ (always in log terms), s/he receives a wage $w_t = r + p + h_t$, where $R = e^r$ is the endogenous contractual piece rate.

The rules governing the determination of the contractual piece rate are borrowed from the sequential auctions model of Postel-Vinay and Robin (2002). A brief sketch follows. Consider a worker with experience level t, employed at a firm of type p under a contract stipulating a piece rate of $R = e^r \leq 1$. Denote the value that the worker derives from being in that state as $V(r, h_t, p)$. This value is an increasing function of the worker's current and future wages and, as such, increases with the piece rate r and the employer's productivity p (see below for a formal confirmation of this statement). As described earlier, this worker contacts a potential alternate employer with probability λ_1 at the end of the current period. The alternate employer's type p' is drawn from the sampling distribution $F(\cdot)$. The central assumption is that the incumbent and outside employers Bertrand-compete over the worker's services, based on the information available at the end of the current period. The firm that values the worker most—i.e. the firm with higher productivity—wins the Bertrand game by offering the worker a piece rate corresponding to the maximum level of expected worker value $\mathbf{E}_t V(\cdot)$ that the other firm was prepared to offer.[6] This maximum value corresponds to the firm giving the worker the entire match surplus by setting the piece rate at $R = 1$ (or $r = 0$).

Formally, the outcome of the Bertrand game can be described as follows. First, if $p' > p$ (the poacher is more productive than the incumbent), then even if the incumbent employer offers a (log) piece rate of $r = 0$—with an associated expected worker value of $\mathbf{E}_t V(0, h_{t+1}, p)$—, the more productive poacher can still profitably attract the worker by offering marginally more than the latter value. This corresponds to a piece rate $r' < 0$ at the type-p' firm defined by the indifference condition:

$$\mathbf{E}_t V\left(r', h_{t+1}, p'\right) = \mathbf{E}_t V(0, h_{t+1}, p). \tag{3}$$

Second, if $p' \leq p$ (the poacher is less productive than the incumbent), then the situation is a priori symmetric in that the incumbent employer is able to profitably retain the worker by offering a piece rate r' such that $\mathbf{E}_t V(r', h_{t+1}, p) = \mathbf{E}_t V(0, h_{t+1}, p')$. Note, however, that p' may be so low that this would not even correspond to a wage increase. This is indeed the case whenever the poacher's type p' falls short of the threshold value $q(r, h_t, p)$, defined by a similar indifference condition:

$$\mathbf{E}_t V(r, h_{t+1}, p) = \mathbf{E}_t V[0, h_{t+1}, q(r, h_t, p)]. \tag{4}$$

In those cases, the worker simply discards the outside offer from p'.

The above describes the rules following which the piece rate of an employed worker is revised over time. Concerning unemployed workers, we consistently assume that firms make take-it-or-leave-it offers to workers. As a result, the piece rate r_0 offered to an unemployed worker with

[6] $\mathbf{E}_t$ designates the expectation operator conditional on the available information at experience t, i.e. conditional on the realized productivity shock ε_t.

experience level t solves:

$$\mathbf{E}_t\, V\left(r_0, h_{t+1}, p\right) = V_0\left(h_t\right), \tag{5}$$

where $V_0(h_t)$ is the lifetime value of unemployment at experience t.

Worker values. The workers' flow utility function is logarithmic and all workers have a common rate of future discount of ρ. The typical employed worker's value function $V(r, h_t, p)$ is then defined recursively as:

$$V\left(r, h_t, p\right) = w_t + \frac{\delta}{1+\rho} V_0\left(h_t\right) + \frac{1}{1+\rho}\mathbf{E}_t \Bigg\{ \left[1 - \mu - \delta - \lambda_1 \overline{F}\left(q\left(r, h_t, p\right)\right)\right] \cdot V\left(r, h_{t+1}, p\right) + \lambda_1 \overline{F}\left(p\right) \cdot V\left(0, h_{t+1}, p\right) + \lambda_1 \int_{q(r,h_t,p)}^{p} V\left(0, h_{t+1}, x\right) dF\left(x\right) \Bigg\}, \tag{6}$$

where the threshold $q(\cdot)$ is defined as in (4). Because the maximum profitable piece rate is $r = 0$, it follows that $q(0, h_t, p) \equiv p$. The worker's value function at this maximum piece rate is then easily deduced from (6). The following is a useful characterization:

$$\mathbf{E}_t\, V\left(0, h_{t+1}, p\right) = \frac{(1+\rho)\,p}{\rho+\mu+\delta} + \mathbf{E}_t \left\{ \sum_{s=0}^{+\infty} \left(\frac{1-\mu-\delta}{1+\rho}\right)^s \cdot \left(h_{t+1+s} + \frac{\delta}{1+\rho} V_0\left(h_{t+1+s}\right)\right) \right\}. \tag{7}$$

Using this latter expression together with integration by parts in (6), we obtain a slightly simpler definition of the worker's generic value function:

$$V\left(r, h_t, p\right) = r + p + h_t + \frac{\delta}{1+\rho} V_0\left(h_t\right) + \mathbf{E}_t \left\{ \frac{1-\mu-\delta}{1+\rho} V\left(r, h_{t+1}, p\right) + \int_{q(r,h_t,p)}^{p} \frac{\lambda_1 \overline{F}\left(x\right)}{\rho+\mu+\delta} dx \right\}. \tag{8}$$

Piece-rate wages. A combination of (4), (7) and (8) leads to the following alternative definition for $q(\cdot)$:

$$\begin{aligned} q\left(r, h_t, p\right) &= \frac{\rho+\mu+\delta}{1+\rho} r + p - \mathbf{E}_t \left\{ \int_{q(r,h_{t+1},p)}^{p} \frac{1-\mu-\delta-\lambda_1 \overline{F}\left(x\right)}{1+\rho} dx \right\} \\ &= \frac{\rho+\mu+\delta}{1+\rho} r + p - \int \int_{q(r,h_{t+1},p)}^{p} \frac{1-\mu-\delta-\lambda_1 \overline{F}\left(x\right)}{1+\rho} dx\, dH\left(h_{t+1} \mid h_t\right), \end{aligned} \tag{9}$$

where $H(\cdot \mid h_t)$ is the law of motion of h_t. Note that this latter is essentially (i.e., up to the deterministic drift g_t) the transition distribution of the first-order Markov process followed by ε_t, as this latter shock is the only stochastic component in h_t. Clearly, (9) has a simple, deterministic (indeed constant), consistent solution $q(r,p)$ implicitly defined by:

$$r = -\int_{q(r,p)}^{p} 1 + \frac{\lambda_1 \overline{F}(x)}{\rho + \mu + \delta} dx. \tag{10}$$

Now even though (9) implies no direct dependence of $q(\cdot)$ on h_t, other, nondeterministic solutions to (9) may still exist because of the autoregressive component in the process of productivity shocks ε_t. Indeed if workers expect future values of the threshold $q(\cdot)$ to be conditioned on future values of their productivity h, then this makes the current threshold $q(r, h_t, p)$ a function of their current productivity h_t because of the latter's persistence.

Neglecting such expectational mechanisms for now, we concentrate on the deterministic solution (10), under which the (log) wage w_{it_i} earned by worker i hired at firm $j(i,t_i)$ at experience level t_i—so that $j(i,t_i)$ is the function mapping worker identifiers and experience levels into employer identifiers—is defined as follows:

$$w_{it_i} = p_{j(i,t_i)} + \alpha_i + g_{t_i} + \varepsilon_{it_i} - \int_{q_{it_i}}^{p_{j(i,t_i)}} 1 + \frac{\lambda_1 \overline{F}(x)}{\rho + \mu + \delta} dx, \tag{11}$$

where q_{it_i} is the type of the last firm from which worker i was able to extract the whole surplus in the offer-matching game. This wage equation implies a decomposition of individual wages into five components: a deterministic trend g_{t_i}, a worker fixed effect α_i, a transitory component ε_{it_i}, an employer fixed effect $p_{j(i,t_i)}$, and a random effect q_{it_i} relating to the most recent wage bargain. The joint process governing the dynamics of $\left[p_{j(i,t_i)}; q_{it_i}\right]'$ can be characterized as follows:

$$\begin{pmatrix} p_{j(i,t_i+1)} \\ q_{i,t_i+1} \end{pmatrix} \mid \begin{pmatrix} p_{j(i,t_i)} \\ q_{it_i} \end{pmatrix} = \begin{cases} \begin{pmatrix} p_{j(i,t_i)} \\ q_{it_i} \end{pmatrix} & \text{with probability } 1 - \mu - \delta - \lambda_1 \overline{F}(q_{it_i}) \\ \begin{pmatrix} p_{j(i,t_i)} \\ p_{j(i,t_i)} > q' > q_{it_i} \end{pmatrix} & \text{with density } \lambda_1 f(q') \\ \begin{pmatrix} p' > p_{j(i,t_i)} \\ p_{j(i,t_i)} \end{pmatrix} & \text{with density } \lambda_1 f(p'). \end{cases} \tag{12}$$

We also have to consider two additional possibilities: first, the worker may retire (probability μ), in which case $\left[p_{j(i,t_i+1)}; q_{i,t_i+1}\right]'$ becomes unobserved forever, and second, the worker may become unemployed (probability δ), in which case $\left[p_{j(i,t_i+1)}; q_{i,t_i+1}\right]'$ is only observed as s/he re-enters employment, and is then equal to $[p'; b]'$, where p' is a random draw from $F(\cdot)$ and b is productivity in nonemployment.

This process is associated with a steady-state cross-sectional distribution of the pair $\left(p_{j(i,t_i)}, q_{it_i}\right)$ derived in Appendix A.[7] Characterization of this steady-state distribution will be useful to simulate the model (see below section 4 and Appendix B).

3 Data

Background. The data used in the empirical analysis consist of a ten percent random sample of workers from the Danish matched employer-employee dataset IDA, which have been merged with detailed data on individual labor market histories covering the period 1986—1999. IDA contains annual socio-economic information on workers and (some) background information on employers, and covers the entire Danish population aged 16-69.[8] The labor market history data (LMH) data is based on weekly reports on unemployment status and mandatory employer pension contributions and provide very reliable information on an individual's labor market status. In the merging procedure, all labor market spells of a given worker in a given calender year are linked to the annual worker-specific IDA information (except for earnings information—see below) for that given year. Hence, the structure of the dataset is such that a worker who occupies, say, three different labor market states during a given calendar year will have three observations associated with that calendar year conveying information on the duration of stay in each of the states (within the calendar year) along with socio-economic information. As this latter piece of information is obtained from the IDA data it is constant over the three observations relating to that given worker for the given calendar year.

The LMS data distinguishes between four labor market states: Employment, temporary un-

[7]The other random components of wages appearing in (11) are exogenously distributed (α_i is just a fixed effect and ε_{it_i} follows an exogenous process of its own), and they are uncorrelated with $p_{j(i,t_i)}$ or q_{it_i}. In other words, the set of assumptions we have adopted implies that there is no assortative assignment of workers to firms based on those unobserved worker characteristics. As will become clear shortly, though, there will be assortative assignment based on experience.

[8]IDA: *Integreret Database for Arbejdsmarkedsforskning* (Integrated Database for Labor Market Research) is constructed and maintained by Statistics Denmark.

employment, unemployment and nonparticipation. Employment spells are associated with a firm identifier.[9] We treat temporary unemployment as employment and aggregate job spells that are interrupted by temporary unemployment into a single job spell of duration equal to the sum of durations of actual employment periods and of periods of temporary unemployment. Thus in the empirical analysis we distinguish between job spells, unemployment spells and nonparticipation spells. Very short nonemployment spells (viz. in practice unemployment or nonparticipation spells shorter than 5 weeks) are likely to be periods of transition before an already obtained job can be initiated rather than "real" nonemployment and are discarded. Consequently, a job-to-nonemployment transition followed by a nonemployment-to-job transition within four weeks is considered a job-to-job transition in the empirical analysis.

Earnings information consists of the annual average hourly wage in the job occupied in the last week of November. This implies that job spells that do not overlap with the last week of November in any year—which likely includes a sizeable proportion of short-term jobs—will have no wage information. Likewise, if the worker was unemployed in the last week of November there is no record of earnings for that worker in the corresponding year. The wage distribution is trimmed from below with a set of "effective" minimum wages, inflated to 1999 levels using Statistics Denmark's consumer price index, and trimmed from above at the 0.975 percentile. Trimming is carried out separately on each stratified sample.

Besides information on labor market transitions and earnings, the most important piece of information for the purpose of this study is workers' labor market experience. This information is available on an annual basis (from IDA) and refers to the workers experience at the end of a calendar year. Experience obtained before 31 Dec 1979 is measured in years, while experience obtained after 1 Jan 1980 is measured in 1/1000 of a year's full-time work, and is constructed from workers' mandatory pension payments, ATP.[10] Notice that we observe workers' *actual* (as opposed to potential) labor market experience.

[9] Employers are identified both at the firm and plant level. We construct job spells using the firm-level identifiers, i.e. we do not treat job changes within the same firm as labor market transitions.

[10] ATP (*Arbejdsmarkedets Tillægs Pension*) is a mandatory pension for all salaried workers aged 16-66 who work more than eight hours per week. ATP-savings are optional for the self-employed. ATP effectively covers the entire Danish labor force.

Additional information on worker characteristics is annual and includes the standard covariates used in earnings regressions, of which we retain gender, age and educational attainment for sample selection and stratification.

The analysis sample. In estimating the model we will assume that the data is drawn from the steady state distributions of earnings, spell durations and experience obtained from the theoretical model (see Appendix A). We thus consider that our theory pertains to workers who are long enough into their working lives to be only seeking to improve their earnings through job search and experience accumulation. To obtain an empirical counterpart of this group of workers we impose a number of sample selection criteria on the data. First, we only select male workers in order not to confound the empirical analysis with fertility and household production issues which are usually believed to have important bearings on female labor market outcomes, but which our model is not well suited to deal with. Second, we trim individual labor market histories at the minimal age of 30. Third, we truncate a worker's labor market history after the first observed transition into a non-participation spell (that is, we consider that transitions into nonparticipation are permanent). Fourth, we stratify the initial sample into three levels of education, based on the number of years spent in education (9-12 years, 13-14 years and 15-18 years). Finally, we discard all workers with missing or inconsistent information on relevant variables. Relevant descriptive statistics are provided below, as we discuss the estimation procedure.

4 The estimation protocol

In this section we discuss estimation protocol that we apply to the data just described in order to obtain structural parameter estimates for the model developed in section 2.

4.1 General approach

The structural model fails to deliver easily tractable, closed-form expressions of the distributions of important endogenous variables (notably wage levels and growth rates), effectively ruling out standard likelihood-based inference. We thus resort instead to indirect inference techniques. The technical details of indirect inference are spelled out in e.g. Gouriéroux, Monfort and Renault (1993), and here we only provide a brief account of the underlying intuition, after which we go on to describe the specific details of our implementation.

Indirect inference is a generalization of the method of simulated moments. The underlying idea is to find values of the structural parameters that minimize the distance between a given set of moments of the real data and the model-predicted counterparts of these moments based on artificial data obtained by simulation of the structural model. The set of moments that are matched in this fashion can be viewed in all generality as the (vector of) parameter(s) of an auxiliary model, which differs from the original structural model that we aim to estimate.

To state things more formally, let us begin by introducing the following notation. Let θ denote the vector of structural parameters, the true value of which is θ_0. For a given value of θ, we further designate by DGP(θ) the structural model under consideration. Finally, let $\mathbf{Y}_N$ designate our estimation sample (the observed data). We work under the maintained identifying assumption that our structural model is correct, i.e. that the data generating process of the observed sample $\mathbf{Y}_N$ is DGP(θ_0), which makes $\mathbf{Y}_N$ a function of the structural parameter set at its true value, θ_0. We further assume that DGP(θ) can be simulated for any given value of θ.

Indirect inference then works through the following steps. First, an estimable auxiliary model characterized by the auxiliary parameter β is formulated and estimated on the observed data $\mathbf{Y}_N$, the resulting estimate of the auxiliary parameter being denoted by $\widehat{\beta}_N(\theta_0)$. The notation purposely emphasizes that $\widehat{\beta}_N(\theta_0)$ is a function of the structural parameter at the true value θ_0, even though the auxiliary model is misspecified since it will generally differ from the original (true) structural model DGP(θ_0).

Second, given a parameter value θ, the structural model DGP(θ) is simulated S times in order to produce S simulated data sets of size $\widetilde{N}$, $\left\{\mathbf{Y}_{\widetilde{N}}^{(s)}(\theta)\right\}_{s=1}^{S}$, where $\widetilde{N}$ needs not equal N. The auxiliary model is next estimated on each of the S simulated data sets to produce a series of S estimates $\left\{\widehat{\beta}_{\widetilde{N}}^{(s)}(\theta)\right\}_{s=1}^{S}$, of which we consider the mean: $\widehat{\beta}_{\widetilde{N},S}(\theta) = \frac{1}{S}\sum_{s=1}^{S}\widehat{\beta}_{\widetilde{N}}^{(s)}(\theta)$. Third and finally, we seek the value of θ that minimizes the distance between $\widehat{\beta}_{\widetilde{N},S}(\theta)$ and $\widehat{\beta}_N(\theta_0)$. Formally, the indirect inference estimator $\widehat{\theta}_N$ is defined by

$$\widehat{\theta}_N = \arg\min_{\theta} \left\| \widehat{\beta}_{\widetilde{N},S}(\theta) - \widehat{\beta}_N(\theta_0) \right\|, \tag{13}$$

where $\|\cdot\|$ denotes the Euclidean distance. The choice of an auxiliary model can thus be seen as a choice of metric with which to measure the distance between the real data and the data simulated from the structural model. Gouriéroux, Monfort and Renault (1993) show that (under a set of regularity assumptions) the indirect inference estimator $\widehat{\theta}_N$ is consistent and asymptotically normal.

4.2 Empirical specification

Indirect inference only requires that the structural model DGP(θ) can be simulated given a value of the parameter vector θ. To that end, functional form assumptions about the sampling distribution of firm productivity $F(\cdot)$, the distribution of worker heterogeneity $H(\cdot)$ and the (deterministic component of the) relationship between productivity and experience g_t are needed.

The sampling distribution of firm productivity $F(\cdot)$ is truncated from below at b, the level of non-market productivity. In the empirical analysis we assume a truncated exponential distribution: $F(p) = 1 - e^{\nu(b-p)}$ for $p \geq b$.[11] We next assume that log-worker effects α are normally distributed among workers. As the productivity of a match equals the sum of the firm and the worker effect $(p + \alpha)$ the two distributions are only identified up to a normalization. We normalize the mean worker effect to zero, hence assuming $H(\cdot) = \mathcal{N}(0, \sigma^2)$. Finally, we specify the deterministic trend in individual productivity as a quadratic function of experience t, i.e. $g_t = \gamma_1 t + \gamma_2 t^2$, thus introducing a new pair of structural parameters (γ_1, γ_2).

The full vector of structural parameters can thus be spelled out as $\theta = (\mu, \delta, \lambda_0, \lambda_1, \nu, b, \sigma, \gamma_1, \gamma_2)'$ and consists of the transition parameters μ, δ, λ_0 and λ_1, the parameters of the distribution of firm productivity, ν and b, the variance of fixed worker heterogeneity, σ, and finally the parameters linking worker productivity to experience, γ_1 and γ_2.[12] We set the period length to one month,

[11] Recalling that p and b designate log-productivity levels, our assumption on $F(\cdot)$ is effectively that the sampling distribution of productivity levels is Pareto(ν).

[12] In the current version of the paper we ignore the idiosyncratic shock ε_{it} (see equation (2)) and do not estimate the discount rate ρ, which is fixed at a monthly value of $\rho = 0.005$.

so that our simplifying assumption that at most one mobility shock occurs within a period (see section 2) can be deemed a reasonable approximation.

Finally, details of the procedure used to simulate the structural model DGP(θ) given a value of θ are given in Appendix B.

4.3 Auxiliary Models

Outline. The choice of an auxiliary model, i.e. the choice of "which moments to match" is a crucial step in the indirect inference approach. Again the auxiliary model determines the metric within which we measure the distance between real and simulated data.

As we argued in the introduction, we relate our analysis to three different strands of the empirical labor literature, namely the "human capital" literature, the "wage dynamics" literature and the job search literature. Our selection of auxiliary models reflects this three-fold link in that it will borrow from the specifications commonly used in each of these strands of literature. Specifically, we will combine the following three auxiliary models: a duration model based on job search/wage posting theory, a Mincerian wage equation with standard human capital variables included as regressors which we supplement with firm fixed-effects, and finally the first-differenced version of this latter Mincerian equation as a model of (within-firm) wage dynamics. Because these auxiliary models are fairly standard reduced-forms used for the analysis of labor market transitions and earnings dispersion/dynamics, our indirect inference procedure has the additional benefit of explicitly linking our structural approach to well-known results from the reduced-form literature.

We now give details of our three auxiliary models in turn. We distinguish between the structural and auxiliary parameters by adding a superscript A to the auxiliary parameters wherever confusion might arise.

Duration: a job search/wage posting model. The auxiliary duration model is derived from a discrete time version of the job search/wage posting model of Burdett and Mortensen (1998). The basic environment of this latter model is much the same as the one of our structural model: job-seekers receive a job offer with per-period probability λ_0^A when unemployed and λ_1^A when employed.

Employed workers further face a per-period job destruction probability of δ^A and retire/die with probability μ^A. The difference with the model considered in this paper is that in the wage-posting model, firms "post wages", i.e. commit to paying a constant wage over the duration of the match and do not engage in offer-matching. As a consequence, from the workers' standpoint, the value of a job is uniquely characterized by the wage w it pays. Upon receiving a job offer, workers draw the associated wage from a continuous wage offer distribution $F^A(\cdot)$.[13]

Under this set of assumptions it can be shown that workers optimally follow a reservation wage strategy, whereby employed workers accept any offer higher than their current wage and unemployed workers accept any job offer (with the lower bound of the wage offer distribution being at least equal to the unemployed workers' common reservation wage). Hence, the hazard rate out of an unemployment spell is simply λ_0^A and the job-to-job transition probability of a worker with current wage w is $\lambda_1^A \overline{F}^A(w)$. The job-to-unemployment hazard rate is δ^A and the hazard rate into nonparticipation is μ^A.

The transition pattern in the auxiliary model is very close to that of the structural model, with the only difference pertaining to job-to-job transitions which occur on receipt of a higher wage offer in the wage posting model, and on receipt of an offer from a more productive (higher p) firm in our structural offer-matching/piece-rate model.[14]

Finally, the pattern of worker mobility in the wage-posting model implies the following useful steady-state relationship between the wage offer distribution $F^A(\cdot)$ and the distribution of earned wages in a cross-section of workers, denoted by $G^A(\cdot)$:

$$F^A(w) = \frac{\left(\mu^A + \delta^A + \lambda_1^A\right) G^A(w)}{\mu^A + \delta^A + \lambda_1^A G^A(w)}. \tag{14}$$

This equation is easily obtained from flow-balance equations and parallels the relationship (A9) between the sampling and cross-sectional distributions of firm types established in Appendix A for

[13]The existence of such a non-degenerate and continuous wage offer distribution is an equilibrium outcome of wage posting models in which firm behavior is explicitly formalized (see Burdett and Mortensen, 1998). However we shall restrict our auxiliary model—which we see as a duration model primarily aimed at identifying the transition parameters of our structural model—to describe the workers' side of the market, thus taking $F^A(\cdot)$ as a primitive.

[14]Thus in particular the wage-posting model does not permit job-to-job transitions with wage cuts. It would be straightforward to extend the auxiliary model to allow for such job-to-job transitions, but we choose not to do so for parsimony's sake, and given the fact that the purpose of the auxiliary model is not to maximize descriptive accuracy, but rather to provide us with informative and easy-to-estimate moments of the data.

the structural model.

The data used for estimation of the auxiliary duration model consist of an initial cross-section of I employed or unemployed workers whom we follow from the initial sampling date (viz. November 1991) until their first observed transition (if any). For each worker we record record the duration until completion/censoring of the initial spell, and in case the spell is completed we record the type of transition that completed the spell (unemployment-to-job, unemployment-to-nonparticipation, job-to-job, etc.). For employed workers we also record the wage earned at the initial sampling date. Descriptive statistics for this sample are given in table 1.

< **Table 1 about here.** >

Let d denote spell durations, ι_μ (resp. ι_δ, ι_λ) an indicator for transition into non-participation (resp. unemployment, employment). Finally, let c be a censoring indicator, i.e. $c = 1 - \iota_\mu - \iota_\delta - \iota_\lambda$ for employed workers and $c = 1 - \iota_\mu - \iota_\lambda$ for unemployed workers. We follow the two-step semi-nonparametric estimation procedure developed in Bontemps, Robin and Van den Berg (2000). The first step consists of using (14) to back out from the (observed) cross-sectional distribution of wages $G^A(\cdot)$ an estimator of the unobserved distribution of wage offers $F^A(\cdot)$ as a function of the transition parameters:

$$\widehat{F}^A(w) = \frac{\left(\mu^A + \delta^A + \lambda_1^A\right)\widehat{G}^A(w)}{\mu^A + \delta^A + \lambda_1^A \widehat{G}^A(w)}, \tag{15}$$

where $\widehat{G}^A(w) = \frac{1}{I}\sum_{i=1}^{I} \mathbf{1}(w_i \leq w)$ is the empirical cdf of wages in the population of initially employed workers, a non-parametric estimator of $G^A(\cdot)$. With the estimator for the wage offer distribution in hand, we then proceed to the second step in which we estimate the auxiliary transition parameters $\left(\mu^A, \delta^A, \lambda_0^A, \lambda_1^A\right)$ using maximum likelihood. Given the spell hazard rates derived above, the log-likelihood contribution of an initially unemployed worker is:

$$\begin{aligned}\ell_u^A\left(\mu^A, \lambda_0^A\right) &= c \times d \ln\left(1 - \mu^A - \lambda_0^A\right) \\ &\quad + (1-c) \times \left[\iota_\mu \ln \mu^A + \iota_\lambda \ln \lambda_0^A + (d-1)\ln\left(1 - \mu^A - \lambda_0^A\right)\right], \end{aligned} \tag{16}$$

while that of an initially employed worker reads as:

$$\ell_e^A\left(\mu^A,\delta^A,\lambda_1^A\right)=c\times d\ln\left(1-\mu^A-\delta^A-\lambda_1^A\widehat{\overline{F}}^A(w)\right)$$
$$+(1-c)\times\left[\iota_\mu\ln\mu^A+\iota_\delta\ln\delta^A+\iota_\lambda\ln\left(\lambda_1^A\widehat{\overline{F}}^A(w)\right)+(d-1)\ln\left(1-\mu^A-\delta^A-\lambda_1^A\widehat{\overline{F}}^A(w)\right)\right]. \tag{17}$$

Conditional on the worker's initial employment status ($u=1$ for initially unemployed workers), the log-likelihood contribution of a generic worker thus writes as $\ell^A\left(\mu^A,\delta^A,\lambda_0^A,\lambda_1^A\right)=u\cdot\ell_u^A\left(\mu^A,\lambda_0^A\right)+(1-u)\cdot\ell_e^A\left(\mu^A,\delta^A,\lambda_1^A\right)$.

From the auxiliary duration model we include the estimated transition parameters $\left(\widehat{\mu}^A,\widehat{\delta}^A,\widehat{\lambda}_0^A,\widehat{\lambda}_1^A\right)$ in the vector of moments used in the indirect inference procedure.[15]

A Mincer wage equation for matched employer-employee data. The second auxiliary model is a Mincerian wage equation augmented to incorporate firm specific effects, as is typically done when applying such equations to matched employer-employee data (Abowd, Kramarz and Margolis, 1999). Formally, we consider:

$$w_{it_i}=\zeta_1 s_{i,j(i,t_i)}+\zeta_2 s_{i,j(i,t_i)}^2+\zeta_3 t_i+\zeta_4 t_i^2+\phi_{j(i,t_i)}+\psi_i+u_{it_i}, \tag{18}$$

where $s_{i,j(i,t_i)}$ is individual i's seniority in the firm $j\left(i,t_i\right)$ they currently work at, ψ_i is a worker fixed-effect, ϕ_j is a firm fixed-effect and u_{it_i} is a statistical residual. The link between this auxiliary wage model and the structural wage equation (11) is as follows. The structural wage equation decomposes wages into a firm heterogeneity component $p_{j(i,t_i)}$, a fixed worker heterogeneity component α_i, a quadratic in experience g_{t_i}, idiosyncratic productivity shocks ε_{it_i} (not included in the current empirical implementation), and labor market frictions through the last integral term in (11). The auxiliary wage equation (18) predicates a similar decomposition, up to the difference that the type of the last employer from which the worker was last able to extract surplus (q_{it_i} in the notation of

[15] As a final remark about the identification of the structural transition parameters μ, δ, λ_0 and λ_1, it should be noted that the theory put forth in section 2 would in principle allow for direct identification of these transition parameters. Indeed it is conceivable to maximize the true likelihood—i.e. the likelihood based on the structural model DGP(θ)— of observed spell durations treating firm types $p_{j(i,t)}$ as unobserved heterogeneity (thus applying Ridder's and van den Berg's (2003) *unconditional inference* approach). However, in practice, we found that this method did not perform very well on the data we are using, which led us to estimate the transition parameters jointly with the rest of the structural parameter vector using indirect inference.

the structural model) is unobserved in the data, and is therefore absent from the conditioning set in the auxiliary model. We include seniority $s_{i,j(i,t_i)}$ as a proxy for this factor. Considering that we have stratified the data according to workers' educational attainments, we claim that we have controlled for the most important human capital variables: education, seniority and experience.

< **Table 2 about here.** >

The data used for the estimation of (18) is a matched employer-employee panel of data on earnings, seniority and experience which we stratify by educational attainment. The panel features I workers and J firms, for a total of N observations. Table 2 provides a brief statistical summary of the sample. Imposing the restriction that the worker-specific effect ψ_i has mean zero and is orthogonal to all other components of the wage equation,[16] we can estimate (18) by first regressing log-earnings on quadratics in seniority and experience as well as a full set of firm dummies.[17] Worker effects are subsequently recovered from the residuals. Specifically, write (18) in matrix form:

$$\mathbf{w} = \mathbf{X}\zeta + \mathbf{F}\phi + \mathbf{D}\psi + \mathbf{u}, \tag{19}$$

where $\mathbf{w}$ is the $N \times 1$ vector of log-wages, $\mathbf{X}$ is the $N \times 4$ matrix of regressors, $\mathbf{F}$ is the $N \times J$ design-matrix of firm indicators, $\mathbf{D}$ is the $N \times I$ design matrix of worker indicators and $\mathbf{u}$ is the $N \times 1$ vector of residuals. The restriction placed on worker effects implies that $[\mathbf{F} : \mathbf{X}]' \mathbf{D} = \mathbf{0}$ so that $\mathbf{D}\psi$ can be subsumed within the residual. We then obtain estimates of $(\zeta', \phi')'$ as:

$$\left(\widehat{\zeta}', \widehat{\phi}'\right)' = \left([\mathbf{X} : \mathbf{F}]' [\mathbf{X} : \mathbf{F}]\right)^{-1} [\mathbf{X} : \mathbf{F}]' \mathbf{w} \tag{20}$$

and the worker effects are recovered as the within-worker average residuals:

$$\widehat{\psi} = \left(\mathbf{D}'\mathbf{D}\right)^{-1} \mathbf{D}'\mathbf{M}_{[\mathbf{X}:\mathbf{F}]}\mathbf{w} \tag{21}$$

where $\mathbf{M}_{[\mathbf{X}:\mathbf{F}]} = \mathbf{I} - [\mathbf{X} : \mathbf{F}] \left([\mathbf{X} : \mathbf{F}]' [\mathbf{X} : \mathbf{F}]\right)^{-1} [\mathbf{X} : \mathbf{F}]'$.

From the auxiliary wage equation we include the estimated slope parameters $\left(\widehat{\zeta}_1, \cdots, \widehat{\zeta}_4\right)$, the

[16]This restriction is also imposed in the structural model, where fixed worker heterogeneity is orthogonal to all other stochastic components in the model. Our treatment of worker specific effects as (uncorrelated) fixed effects in the auxiliary wage equation (18) is therefore consistent with the structural model.

[17]We run the full regression in one step using the "SparseSolve" routine in GAUSS, but the estimator is equivalent to the within-firm estimator of (18).

average firm effect $\overline{\phi} = \frac{1}{J}\sum_{j=1}^{J}\widehat{\phi}_j$, the standard error of estimated firm effects $\widehat{\sigma}_{\phi} = \sqrt{\frac{1}{J}\sum_{j=1}^{J}\left(\widehat{\phi}_j - \overline{\phi}\right)^2}$ and the standard error of worker effects $\widehat{\sigma}_{\psi} = \sqrt{\frac{1}{I}\sum_{i=1}^{I}\widehat{\psi}_i^{\,2}}$.

Dynamic moments. Using the auxiliary wage equation (18) we can consider the autocorrelation structure of within-job wage growth, which is what the estimation of statistical models of earnings dynamics is typically based (see e.g. Alvarez, Browning and Ejrnæs, 2001). For simplicity, we condition the analysis on worker i staying in the same firm between experience levels t_i and $t_i + 1$. Taking first differences in equation (18) under this restriction yields the following auxiliary model for within-job wage growth:

$$\Delta w_{it_i} = \xi_0 + \xi_1 \Delta s^2_{i,j(i,t_i)} + \Delta u_{it_i}. \tag{22}$$

First-differencing eliminates the firm and worker fixed heterogeneity components. Moreover we only include experience in the r.h.s. of (22) as, within a job spell, experience and seniority are undistinguishable.

We estimate (22) directly rather than using the estimated residuals $\widehat{u}_{it}$ from (18) for two reasons. First, contrary to $\widehat{u}_{it}$, the residuals from (22) are not affected by estimation errors on the firm and worker effects. Second, the estimation of (22) provides us with thwo additional slope parameters $\left(\widehat{\xi}_0, \widehat{\xi}_1\right)$ which convey information and can be incorporated into the set of moments to match. Note that we thus do *not* impose consistency of coefficient estimates between the auxiliary wage equations in levels (18) and growth rates (22). According to our structural model, this pair of equation is a misspecified representation of the individual earnings process and one should therefore not expect it to be consistent in any particular way.

Under the maintained assumption that Δu_{it_i} is independent of $s_{i,j(i,t_i)}$ and t_i, (22) can be estimated by OLS. Doing this and using the resulting vector of estimated residuals $\widehat{\Delta \mathbf{u}}$ we next compute the first three within-job autocorrelation coefficients:

$$\widehat{\rho}_k = \frac{\widehat{\Delta \mathbf{u}}' \mathbf{L}^k \widehat{\Delta \mathbf{u}}}{\widehat{\Delta \mathbf{u}}' \widehat{\Delta \mathbf{u}}}, \qquad k = 1, 2, 3, \tag{23}$$

where $\mathbf{L}$ is the lag operator and include these in the set of moments to be matched, along with the slope parameters from (22), $\left(\widehat{\xi}_0, \widehat{\xi}_1\right)$.[18]

Summary. We thus have a set of 16 moments that we seek to match using our structural model: the four transition parameters of the auxiliary job search model $\left(\widehat{\mu}^A, \widehat{\delta}^A, \widehat{\lambda}_0^A, \widehat{\lambda}_1^A\right)$, the four slope parameters of the wage equation (18) $\left(\widehat{\zeta}_1, \cdots, \widehat{\zeta}_4\right)$, the first and second moments of the worker and firm effects in that same equation $(\overline{\phi}, \widehat{\sigma}_\phi, \widehat{\sigma}_\psi)$, and finally the two slope parameters and the three autocorrelation coefficients of residuals from the wage growth equation (22), $\left(\widehat{\xi}_0, \widehat{\xi}_1\right)$ and $(\widehat{\rho}_1, \cdots, \widehat{\rho}_3)$.

< Table 3 about here. >

We tested our indirect inference procedure using a Monte-Carlo study on 200 replications of small simulated samples (1,000 workers followed over 120 period, i.e. ten years). Results of our MC study are reported in table 3. For all parameters, the true value lies well within the 95 percent confidence bounds and the mean estimate is reasonably precise and close to the true value. Overall, and given the small size of our simulated samples, these results seem reassuring both about the accuracy of our indirect inference estimator and about its small sample properties.

5 Results

We begin this section with a brief look at the results pertaining to our three auxiliary models, as these will be useful for later comparison with the structural model. At that point we also comment on our structural model's capacity to replicate those results. We then turn to structural parameter estimates, and comment on the structural model's account of individual earnings dynamics within and between job spells.

5.1 Auxiliary models

The auxiliary job search/wage posting model. Table 4 reports estimates of the first set of auxiliary parameters—namely, transition parameters of the auxiliary wage posting model—obtained

[18]To compute all three autocorrelations we need five observations within the job. Hence, the entire analysis of earnings autocorrelations is restricted to jobs for which we have at least five observations.

from the real data. The corresponding estimates based on data generated by the structural model are shown in table 5.

< Tables 4 and 5 about here. >

Estimates on the real data basically reflect the average spell durations and transition rates shown in the descriptive table 1 (only adding a correction for right-censoring), and as such they do not bring about specific comments.

[RESULTS FROM THE STRUCTURAL MODEL NOT YET AVAILABLE.]

The auxiliary wage equation (18). Parameter estimates of the auxiliary wage equation (18) are gathered in tables 6 (data) and 7 (model), together with moments of the firm and worker heterogeneity distributions.

< Tables 6 and 7 about here. >

We first review estimates based on the real data (table 6). The auxiliary wage equation indicates positive returns to both tenure and experience in all three subsamples. For each education group, estimated returns to experience and tenure are of similar orders of magnitude, and they are rather modest (1.5 to 4 log points after five years of experience/tenure). Perhaps one puzzling feature of this set of estimates is that it implies substantially lower returns to both tenure and experience for the high-educated group than for the other two. Exactly what drives this result is not entirely clear.[19] Nor does it really matter for our purposes, though, given the status of equation (18) as a mere auxiliary model. We shall only take the coefficients displayed in table 6 as features of the estimation samples that we seek to match using our structural model.

Comparison of the firm and worker effect distributions across education groups hints at a small degree of positive assortative matching on education, whereby more educated workers tend to be hired at firms with (slightly) higher mean unobserved heterogeneity parameter. (This particular

[19]One potential suspect would be our assumption about the lack of correlation between the worker effect ψ_i and the other regressors. Assortative matching (i.e. a nonzero covariance of worker and firm effects), for example, would affect the estimates of the tenure coefficients as tenure with a given firm is likely (positively) correlated with the firm effect, as any search model would predict. Difference across education groups in the extent of assortative matching could potentially explain the unintuitive ranking of these groups in terms of returns to tenure and experience that is implied by OLS estimation of (18).

interpretation if of course conditional on the normalization of the mean worker effect at zero in all samples.) Moreover, dispersion of firm and worker effects tends to be slightly higher among more educated groups.

The bottom of table 6 finally provides the the cross-sectional variance decomposition of log wages that results from (18). About 80% of total wage variance is explained by (18) in every education group, of which roughly 50% is due to individual characteristics ($X\zeta + \psi$) and 30% is attributable to firm effects (ϕ). Interestingly, the correlation between firm effects and observed individual characteristics (i.e. tenure and experience) is virtually zero in all groups (while the correlation between firm effects and unobserved person effects is zero by construction). Overall, this decomposition of $V(\ln w)$ is in the same ballpark as that obtained by Abowd, Kramarz, Lengermann and Roux (2003) applying (a slightly enriched version of) equation (18) to US and French data.

[RESULTS FROM THE STRUCTURAL MODEL NOT YET AVAILABLE.]

The auxiliary wage growth equation (22). Finally, tables 8 (data) and 9 (model) report parameter estimates for the auxiliary wage growth equation (22).

< **Tables 8 and 9 about here.** >

Consistency would require the constant term in the wage growth equation (22) (which is labeled as the linear effect of experience in tables 8 and 9) to equal the sum of the linear terms in experience and tenure in the specification in levels (18). Comparison of the relevant numbers in tables 6 and 8 reveals that the latter is systematically larger (by 20 to 60 percent, depending on the education group) than the former. This may be taken to reflect the upward bias than is generally thought to affect OLS estimates of the returns to seniority (Altonji and Williams, 2004).[20] Once again, however, this inconsistency among auxiliary models—which, by assumption, misrepresent of the data—does not really matter for our purposes.

Table 8 also reports the autocovariance and autocorrelation structures of unexplained wage

[20] Note that the first-difference estimates of the combined returns to experience and tenure are not affected by the assumption of uncorrelated worker effects ψ_i discussed earlier in footnote 19.

growth based on (22). Interestingly, the residual autocovariance seems to decline smoothly over successive lags (at least up to 4 lags, which is the highest reported order), in a way that is not frankly suggestive of a low-order MA structure, as would typically be found in studies of individual earnings dynamics based on individual or household data. Whether our structural model is able to replicate this and other features of the data is the question to which we turn next.

[RESULTS FROM THE STRUCTURAL MODEL NOT YET AVAILABLE.]

5.2 Structural model

[RESULTS FROM THE STRUCTURAL MODEL NOT YET AVAILABLE.]

6 Conclusion

[FORTHCOMING]

References

[1] Abowd, J. M., F. Kramarz, P. Lengermann and S. Roux (2003),"Interindustry and Firm-Size Wage Differentials in the United States and France", Cornell University working paper.

[2] Abowd, J. M., F. Kramarz, and D. N. Margolis (1999), "High Wage Workers and High Wage Firms", *Econometrica*, 67, 251-333.

[3] Altonji, J. G. and R. Shakotko (1987), "Do Wages Rise With Job Seniority?", *Review of Economic Studies*, LIV (3), 437-60.

[4] Altonji, J. G. and N. Williams (2004), "Do Wages Rise With Job Seniority? A Reassessment", *Industrial and Labor Relations Review*. Forthcoming.

[5] Alvarez, J., M. Browning and M. Ejrnæs (2001), "Modelling Income Processes with Lots of Heterogeneity", CAM discussion paper 2002-01.

[6] Baker, M. (1997), "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings", *Journal of Labor Economics*, 15 (2), 338-75.

[7] Barlevy, G. (2005), "Estimating Models of On-the-Job Search using Record Statistics", Federal Reserve Bank of Chicago Working Paper (http://www.chicagofed.org/publications/workingpapers/papers/wp2003-18.pdf).

[8] Beffy, M., M. Buchinsky, D. Fougère, T. Kamionka and F. Kramarz (2005), "The Returns to Seniority in France (and Why they are Lower than in the United States)", paper presented at the 9th Econometric Society World Congress (http://eswc2005.econ.ucl.ac.uk/ESWC/2005/prog/getpdf.asp?pid=1440&pdf=/papers/ESWC/2005/1440/BBFKK.pdf).

[9] Bontemps, C., J.-M. Robin, and G. J. Van den Berg (2000), "Equilibrium Search with Continuous Productivity Dispersion: Theory and Non-Parametric Estimation", *International Economic Review*, 41 (2), 305-58.

[10] Buchinsky, M., D. Fougère, F. Kramarz and R. Tchernis (2002), "Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.", CREST Working Paper 2002-29.

[11] Burdett, K. and D. T. Mortensen (1998), "Wage Differentials, Employer Size and Unemployment", *International Economic Review*, 39, 257-73.

[12] Gouriéroux, C., A. Monfort and E. Renault (1993), "Indirect Inference", *Journal of Applied Econometrics*, 8, 85-118.

[13] Guiso, L., L. Pistaferri and F. Schivardi (2005), "Insurance Within the Firm", forthcoming *Journal of Political Economy.*

[14] Meghir, C. and L. Pistaferri (2004), "Income Variance Dynamics and Heterogeneity", *Econometrica*, 72, 1-32.

[15] Mincer, J. (1974), *Schooling, Experience and Earnings*, New-York: Columbia University Press.

[16] Postel-Vinay, F. and J.-M. Robin (2002), "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity", *Econometrica*, 70 (6), 2295-350.

[17] Postel-Vinay, F. and J.-M. Robin (2006), "Microeconometric Search-Matching Models and Matched Employer-Employee Data", *in* R. Blundell, W. Newey and T. Persson, editors, *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress*, Cambridge: Cambridge University Press (forthcoming).

[18] Postel-Vinay, F. and H. Turon (2005), "On-the-job Search, Productivity Shocks, and the Individual Earnings Process", CEPR Discussion Paper No. 5593.

[19] Ridder, G. and Van den Berg, G. J. (2003) "Measuring Labor Market Frictions: A Cross-Country Comparison", *Journal of the European Economic Association*, 1(1), 224-44.

[20] Rubinstein, Y. and Y. Weiss (2005), "Post-Schooling Wage Growth: Investment, search and Learning", manuscript, Tel-Aviv University.

[21] Topel, R. (1991), "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority", *Journal of Political Economy*, 99 (1), 145-76.

APPENDIX

A Derivation of steady-state distributions

In this appendix we derive the joint steady-state cross-sectional distribution of two of the random components of wages appearing in (11), namely $\left(p_{j(i,t_i)}, q_{it_i}\right)$. This derivation will be useful to simulate the model, which we will need to do when implementing our estimation procedure based on simulated moments.

The steady state assumption implies that inflows must balance outflows for all stocks of workers defined by a status (unemployed or employed), a level of experience t, a piece rate r, and an employer type p. This Appendix spells out the relevant flow-balance equations and the ensuing characterizations of steady-state distributions.

Unemployment rate. Assuming that all labor market entrants start off at zero experience as unemployed job seekers and equating unemployment inflows and outflows immediately leads to the following definition of the steady-state unemployment rate, u:

$$u = \frac{\mu + \delta}{\mu + \delta + \lambda_0}. \tag{A1}$$

Distribution of experience levels. Denote the steady-state fraction of employed (resp. unemployed) workers with experience *equal to* t by $a_1(t)$ (resp. $a_0(t)$). For any positive level of experience, $t \geq 1$, these two fractions are related by the following pair of difference equations:

$$(\lambda_0 + \mu)\, u a_0(t) = \delta(1-u)\, a_1(t) \tag{A2}$$

$$(1-u)\, a_1(t) = (1 - \mu - \delta)(1-u)\, a_1(t-1) + \lambda_0 u a_0(t-1). \tag{A3}$$

with the fact $a_1(0) = 0$ stemming from the assumed within-period timing of events, which implies that employed workers always have strictly positive experience. Moreover, the fraction of "entrants", i.e. unemployed workers with no experience $a_0(0)$, is given by:

$$(\mu + \lambda_0)\, u a_0(0) = \mu. \tag{A4}$$

Jointly solving those three equations, one obtains:

$$a_1(t) = \left(\mu + \frac{\mu\delta}{\mu + \lambda_0}\right)\left(1 - \mu - \frac{\mu\delta}{\mu + \lambda_0}\right)^{t-1}. \tag{A5}$$

The corresponding cdf is obtained by summation:

$$A_1(t) = \sum_{\tau=1}^{t} a_1(\tau) = 1 - \left(1 - \mu - \frac{\mu\delta}{\mu + \lambda_0}\right)^{t}. \tag{A6}$$

(Note that, as a result of the adopted convention regarding the within-period timing of events, no employed worker has zero experience.) $A_0(t)$ is then deduced from summation of (A2): $A_0(t) = \frac{\mu(\mu+\delta+\lambda_0)}{(\mu+\delta)(\mu+\lambda_0)} + \frac{\delta\lambda_0}{(\mu+\delta)(\mu+\lambda_0)} A_1(t)$.

Conditional distribution of firm types across employed workers. Let $L(p \mid t)$ denote the fraction of employed workers with experience level $t \geq 1$ working at a firm of type p or less. For $t = 1$ workers can only be hired from unemployment, implying that $L(p \mid t = 1) = F(p)$. For $t > 1$ workers can come from both employment and unemployment and the flow-balance equation determining $L(p \mid t)$ is given by:

$$L(p \mid t)\, a_1(t) = \left(1 - \mu - \delta - \lambda_1 \overline{F}(p)\right) L(p \mid t-1)\, a_1(t-1) + (\mu + \delta)\, a_0(t-1)\, F(p). \tag{A7}$$

Using (A2), and since (A5) implies:

$$\frac{a_1(t-1)}{a_1(t)} = \left(1 - \mu - \frac{\mu\delta}{\mu + \lambda_0}\right)^{-1} = \frac{\mu + \lambda_0}{\mu + \lambda_0 - \mu(\mu + \delta + \lambda_0)},$$

one can rewrite (A7) as $L(p \mid t) = \Lambda_1(p)\, L(p \mid t-1) + \Lambda_2 F(p)$, with:

$$\Lambda_1(p) = \frac{\left(1 - \mu - \delta - \lambda_1 \overline{F}(p)\right)(\mu + \lambda_0)}{\mu + \lambda_0 - \mu(\mu + \delta + \lambda_0)} \quad \text{and} \quad \Lambda_2 = \frac{\delta\lambda_0}{\mu + \lambda_0 - \mu(\mu + \delta + \lambda_0)}.$$

This last equation solves as:

$$L(p \mid t) = \left[\Lambda_1(p)^{t-1} + \Lambda_2 \frac{1 - \Lambda_1(p)^{t-1}}{1 - \Lambda_1(p)}\right] F(p). \tag{A8}$$

Summing over experience levels, we obtain the unconditional cdf of firm types:

$$L(p) = \frac{(\mu + \delta)\, F(p)}{\mu + \delta + \lambda_1 \overline{F}(p)}. \tag{A9}$$

Conditional distribution of piece rates. Equation (10) states that piece rates are of the form $r = r(q, p)$. Thus the conditional distribution of piece rates within a type-p firm is fully characterized by the distribution of threshold values q in a type-p firm, $G(q \mid p, t)$, which we now derive. For $t > 1$, the flow-balance equation determining $G(q \mid p, t)$ is given by:

$$\begin{aligned} G(q \mid p, t)\, \ell(p \mid t)\, a_1(t) = {} & \left(1 - \mu - \delta - \lambda_1 \overline{F}(q)\right) G(q \mid p, t-1)\, \ell(p \mid t-1)\, a_1(t-1) \\ & + \lambda_1 L(q \mid t-1)\, a_1(t-1)\, f(p) + (\mu + \delta)\, a_0(t-1)\, f(p), \end{aligned} \tag{A10}$$

where $\ell(p \mid t) = L'(p \mid t)$ is the conditional density of firm types in the population of employed workers corresponding to the cdf in (A8). Rewriting this last equation in the case $q = p$, so that $G(q \mid p, t) = 1$, yields the differential version of (A7):

$$\begin{aligned} \ell(p \mid t)\, a_1(t) = {} & \left(1 - \mu - \delta - \lambda_1 \overline{F}(p)\right) \ell(p \mid t-1)\, a_1(t-1) \\ & + \lambda_1 L(p \mid t-1)\, a_1(t-1)\, f(p) + (\mu + \delta)\, a_0(t-1)\, f(p). \end{aligned} \tag{A11}$$

Dividing (A10) and (A11) by $f(p)$ throughout shows that $G(q \mid p, t)\, \ell(p \mid t)\, a_1(t) / f(p)$ and $\ell(q \mid t)\, a_1(t) / f(q)$ solve the same equation. Hence:

$$G(q \mid p, t) = \frac{\ell(q \mid t) / f(q)}{\ell(p \mid t) / f(p)} \qquad \text{for } q \in [p_{\min}, p],\ t > 1. \tag{A12}$$

The unconditional version, (A13), obtains by similar reasoning:

$$G(q \mid p) = \frac{\ell(q) / f(q)}{\ell(p) / f(p)} = \left(\frac{\mu + \delta + \lambda_1 \overline{F}(p)}{\mu + \delta + \lambda_1 \overline{F}(q)}\right)^2, \qquad \text{for } q \in [p_{\min}, p]. \tag{A13}$$

B Details of the simulation procedure

This Appendix describes the procedure that we implement in order to simulate a panel of I workers over T periods given values of the structural model's parameters. In practice, we have used $I = 5,000$ and $T = 120$ months (ten years) in the main estimation routine.

We begin with a sample of N workers with zero experience for which we draw individual (log) heterogeneity parameters α from $\mathcal{N}(0, \sigma^2)$. All workers start out unemployed. We thus begin out of the steady state. We first simulate the model for a large number of periods until the various conditional cross-sectional distributions of firm types among workers become (approximately) confounded with the corresponding predicted steady-state distributions (see Appendix A).[21] We then use the output of this preliminary simulation as the initial state of the labor market for a second, T-period simulation which produces the final simulated data set.

At each new simulated period we append the following to the record of each individual worker: the worker's status (employed or unemployed), the worker's experience level, and if employed, the worker's tenure, employer type p and threshold value $q(\cdot)$ determining the worker's piece rate.

In each period, a worker can receive an offer (probability λ_0 or λ_1, depending on the worker's current status), become unemployed (probability δ) or leave the sample (probability μ). Each time an unemployed worker receives an offer, we record a change of status, the productivity of the new employer (p'), an increase in experience and we set the worker's tenure to one. When an employed worker (with employer type p) receives an offer, this results in a job-to-job transition if $p' > p$, in which case we record the productivity p' of the new employer, set $q(\cdot) = p$, the worker's tenure at the new firm to one and increment the worker's experience. In case $q(\cdot) < p' \leq p$, the does not change firms. However we need to update the worker's productivity threshold $q(\cdot)$ to p', and also increment the worker's seniority and experience. Finally, workers who leave the sample (probability μ are automatically (i.e. deterministically) replaced by newborn unemployed workers with zero experience and new values of α drawn from $H(\cdot) \equiv \mathcal{N}(0, \sigma^2)$.

We give the following tweak to the simulations. Upon receiving an offer, workers theoretically draw the offering firm's type p' from the continuous sampling distribution $F(p') = 1 - e^{\nu(b-p')}$. Because the theoretical $F(\cdot)$ is continuous, a rigorous implementation of this would invariably produce simulated (finite) samples with at most one worker observation per firm (where a firm is defined by its value of p), thus making the identification of firm effects in the auxiliary wage equation (18) impossible. To get round this problem, we discretize $F(\cdot)$ in the following way. We take a fixed number J of active firms (e.g. the number of firm observations in the actual sample), give each of them a rank $j = 1, \cdots, J$ and assign corresponding productivity levels of $p_j = F^{-1}(j/J)$. Firm ranks j are then drawn uniformly by workers as they receive offer.

[21] Various experiments suggested that a 600-period initial simulation was on the safe side to ensure this convergence. This turned out to be faster to run than direct draws from the predicted steady-state distributions, notably because of the conditioning on the workers' experience levels which makes the number of different distributions we would have to take draws from very large.

The simulated data set is split into two auxiliary data sets: one for the estimation of the auxiliary job search model and one for the estimation of the wage equations (18) and (22). Simulated data sets, which have monthly wage observations (computed using (11) and the information recorded for each worker), are remodeled to replicate the structure of the actual data set (which only has yearly wage observations for the active job spell at the end of November—see section 3).

	Educ. 9-12 years	Educ. 13-14 years	Educ. 15-18 years
Number of spells	25,466	30,086	10,191
U-spells (in months):			
Number of U-spells	2,639	2,082	395
Number of UJ transitions	2,048	1,716	338
Number of UN transitions	581	364	55
Number of censored U-spells	10	2	2
Mean (*S.D.*) noncensored residual dur.	11.23 (*13.09*)	9.97 (*10.72*)	10.87 (*11.81*)
J-spells (in months):			
Number of J-spells	22,827	28,004	9,796
Number of JU transitions	3,494	3,550	545
Number of JJ transitions	9,621	12,494	4,209
Number of JN transitions	3,236	3,386	1.053
Number of censored J-spells	6,386	8,574	3.989
Mean (*S.D.*) noncensored residual dur.	34.12 (*29.11*)	35,50 (*29.34*)	36.25 (*29.41*)
Cross-Section Wages (in DKK):			
Minimum/Maximum	77.39/341.20	77.39/340.03	77.39/511.22
10%-percentile	124.29	130.15	152.43
25%-percentile	140.70	146.57	173.53
50%-percentile	162.98	172.36	215.74
75%-percentile	193.47	205.19	276.72
90%-percentile	230.99	246.23	341.20
Mean (*S.D.*) wages	171.36 (*44.51*)	180.40 (*46.01*)	233.40 (*76.58*)

Table 1: Sample descriptive statistics—Initial spell durations and wages

	# Workers	Total number of employers						Average
		1	2	3	4	5	6+	
Educ. 9-12 years	40,417	19,012	10,312	5,665	3,095	1,323	1,010	2.04
Educ. 13-14 years	38,302	16,756	10,062	5,634	3,099	1,471	1,280	2.14
Educ. 15-18 years	13,829	5,255	3,829	2,461	1,275	593	416	2.24

	# Workers	Number of years of presence in sample							Average
		1	2	3	4	5	6-10	11-14	
Educ. 9-12 years	40,417	5,231	4,585	4,025	3,747	3,399	13,360	6,070	5.90
Educ. 13-14 years	38,302	4,463	3,064	2,912	2,815	2,689	13,925	8,434	6.76
Educ. 15-18 years	13,829	995	1,231	1,110	1,140	1,092	5.291	2,970	6.91

	# Firms	Firm size						Average
		1-5	6-10	11-20	21-50	51-100	100+	
Educ. 9-12 years	29,429	27,484	1,096	465	253	86	45	2.74
Educ. 13-14 years	31,367	29,521	1,055	444	238	67	42	2.56
Educ. 15-18 years	9,870	9,029	430	228	125	35	23	3.07

Table 2: Sample descriptive statistics—Worker mobility

	μ	δ	λ_0	λ_1	ν	b	σ	γ_1	γ_2
True	.0030	.0050	.070	.020	2.00	5.00	.500	.050	−.0008
Mean	.0035	.0055	.070	.018	2.06	4.92	.407	.072	−.0012
S.D.	.0006	.0006	.008	.007	.21	.098	.116	.014	.0003
$P_{2.5}$	.0025	.0044	.051	.007	1.63	4.72	.071	.043	−.0017
P_5	.0028	.0045	.054	.008	1.73	4.76	.169	.049	−.0017
P_{10}	.0029	.0047	.059	.011	1.80	4.79	.257	.056	−.0016
P_{25}	.0031	.0050	.064	.013	1.92	4.85	.338	.063	−.0014
P_{50}	.0034	.0055	.070	.017	2.06	4.91	.429	.072	−.0013
P_{75}	.0037	.0058	.076	.022	2.20	4.99	.484	.081	−.0011
P_{90}	.0041	.0062	.079	.028	2.31	5.03	.540	.087	−.0009
P_{95}	.0046	.0064	.081	.030	2.37	5.06	.556	.093	−.0008
$P_{97.5}$	.0048	.0065	.083	.032	2.47	5.11	.584	.100	−.0007

Table 3: Monte-Carlo study of the estimation procedure

	μ^A		δ^A		λ_0^A		λ_1^A	
Educ. 9-12 years	.0042	*.0001*	.0039	*.0001*	.0682	*.0015*	.0193	*.0003*
Educ. 13-14 years	.0035	*.0001*	.0034	*.0001*	.0830	*.0019*	.0200	*.0003*
Educ. 15-18 years	.0033	*.0001*	.0016	*.0001*	.0779	*.0041*	.0146	*.0004*

Table 4: Parameter estimates—Auxiliary wage posting model—Real data

[NOT YET AVAILABLE]

Table 5: Parameter estimates—Auxiliary wage posting model—Simulated data

	Educ. 9-12 years		Educ. 13-14 years		Educ. 15-18 years	
Tenure (years)	.0103	*.0004*	.0096	*.0004*	.0042	*.0009*
Tenure^2 (years^2/100)	−.0519	*.0040*	−.0356	*.0038*	−.0228	*.0080*
Experience (years)	.0095	*.0003*	.0042	*.0003*	.0352	*.0004*
Experience^2 (years^2/100)	−.0198	*.0007*	−.0035	*.0007*	−.0635	*.0012*
Firm effects						
Mean	4.96		5.02		5.05	
Std. dev.	.25		.23		.31	
10%-percentile	4.66		4.75		4.66	
25%-percentile	4.79		4.87		4.86	
50%-percentile	4.93		5.00		5.06	
75%-percentile	5.11		5.16		5.25	
90%-percentile	5.30		5.32		5.43	
Worker effects						
Mean	—Normalized—		—Normalized—		—Normalized—	
Std. dev.	.14		.13		.17	
10%-percentile	−.15		−.14		−.19	
25%-percentile	−.07		−.06		−.08	
75%-percentile	.05		.05		.07	
90%-percentile	.15		.15		.19	
Log-wage variance decomposition						
$V(\ln W)$	.0641		.0591		.0984	
$V(X\zeta)$	.0008		.0009		.0127	
$V(\phi)$	.0345		.0309		.0471	
$V(\psi)$	.0147		.0137		.0236	
$\text{Cov}(X\zeta, \phi)$	.00007		.00002		−.00183	
$\text{Corr}(X\zeta, \phi)$	.012		.003		−.075	
Explained var. (*% of* $V(\ln w)$)	.0501	(*78.2 %*)	.0455	(*77.0 %*)	.0798	(*81.1 %*)

Table 6: Parameter estimates—Auxiliary wage equation (18)—Real data

[NOT YET AVAILABLE]

Table 7: Parameter estimates—Auxiliary wage equation (18)—Simulated data

	Educ. 9-12 years		Educ. 13-14 years		Educ. 15-18 years	
Experience (years)	.0124	*.0012*	.0103	*.0012*	.0336	*.0016*
Experience^2 (years^2/100)	−.0157	*.0028*	−.0076	*.0026*	−.0488	*.0040*
Autocovariances:						
Order 0	.0143		.0130		.0151	
Order 1	−.0035		−.0032		−.0034	
Order 2	−.0011		−.0011		−.0011	
Order 3	−.0010		−.0008		−.0011	
Order 4	−.0006		−.0005		−.0007	
Autocorrelations:						
Order 1	−.2431		−.2424		−.2246	
Order 2	−.0794		−.0837		−.0757	
Order 3	−.0684		−.0630		−.0737	
Order 4	−.0408		−.0414		−.0480	

Table 8: Parameter estimates—Auxiliary wage growth equation (22)—Real data

[NOT YET AVAILABLE]

Table 9: Parameter estimates—Auxiliary wage growth equation (22)—Simulated data