

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

Jesper Bagger Francois Fontaine
University of Aarhus* Université de Strasbourg 3†

Fabien Postel-Vinay
University of Bristol
and

U. Paris 1 — Paris School of Economics‡

Jean-Marc Robin
U. Paris 1 — Paris School of Economics,§
and University College London, IFS and Cemmap

December 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 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. Our structural model permits a decomposition of monthly wage growth into contributions from human capital accumulation and from job search, within and between job spells. We find that the job-search-related within-job effects dominates between-job effects. In relative terms, human capital accumulation is quantitatively more important for wage growth early in workers' careers, and its quantitative importance increases in workers' educational attainment. Indeed, human capital accumulation is the primary source for early career wage growth among high-educated workers.

Keywords: Job Search, Human Capital Accumulation, Within-Job Wage Growth, Between-Job Wage Growth, Individual Shocks, Structural Estimation, Matched Employer-employee Data.

JEL codes:

*E-mail: jbagger@econ.dk.

†E-mail: Francois.Fontaine@urs.u-strasbg.fr.

‡Address: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, United Kingdom.
E-mail: Fabien.Postel-Vinay@bristol.ac.uk.

§Address: Université Paris 1 Panthéon Sorbonne, Centre d'Economie de la Sorbonne, 106/112 bd de l'Hôpital, 75647 Paris Cedex 13. E-mail: jmrobin@univ-paris1.fr.

1 Introduction

We present a tractable equilibrium job search model of individual worker careers allowing for human capital accumulation, employer heterogeneity and individual-specific productivity shocks. We estimate our structural model on a panel of Danish matched employer-employee data and use it to analyze the determinants of individual wage dynamics. Our main motivation for doing this is to quantify the respective roles of human capital accumulation and the forces of labor market competition in shaping the profile of individual labor earnings over the life cycle.

Our contribution adds to the empirical literature on wage equations in three significant ways. The first one refers to 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 U.S., 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 U.S. evidence and find support for both approaches, thus calling for the construction of a unified model. This paper offers such a model.

Earlier combinations of job search and human capital accumulation include the models of Bunzel, Christensen, Kiefer and Korsholm (2000), Rubinstein and Weiss (2005), Barlevy (2005) and Yamaguchi (2006). The model of Bunzel et al. is restrictive in a number of other ways. The human capital production function is linear, workers reap all benefit from human capital accumulation, and there is complete depreciation of workers' human capital upon lay-offs. Moreover, there is no individual-specific productivity shocks, and, although the model is estimated on Danish administrative data, it does not account for intrinsic firm and worker heterogeneity, or productivity shocks. Rubinstein's and Weiss model efficient wage contracts in a market of homogeneous firms, and their analysis of the model is essentially qualitative. Both Barlevy and Yamaguchi allow for determin-

¹Rubinstein 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.

istic human capital accumulation and stochastic productivity shocks. Barlevy uses Burdett and Mortensen's (1998) wage posting framework. Yamaguchi uses a sequential auction framework à la Postel-Vinay and Robin (2002), augmented with bargaining as in Dey and Flinn (2005) and Cahuc, Postel-Vinay and Robin (2006).²

Because Barlevy's and Yamaguchi's contributions are very much related to ours, a longer discussion of how these two models and empirical studies differ from our own work is in order.

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. With wage posting, designing optimal contract menus for changing environment is an extremely difficult task that is analytically tractable only in special cases (e.g. Stevens, 2004; Burdett and Coles, 2003; Harris and Holmstrom, 1982; Rubinstein and Weiss, 2005). With complete information and when contracts can only be modified by mutual agreement, what renders the model complicated is that a positive productivity shocks is fully captured by the employer unless the employer can credibly threaten his employer to quit; and when a negative shock occurs, the worker accepts a wage cut only if the shock is big enough to make firm profit negative. These bands of inaction turn Bellman equations for value functions into finite difference equations which are impossible to solve analytically except in very special cases (Postel-Vinay and Turon, 2005³).

Barlevy chooses not to trade heterogeneity and a realistic process of individual productivity shocks against theoretical generality and restricts the set of available wage contracts to piece rate contracts specifying the share of output received by the worker as a wage. We here also follow Barlevy's suggestion and assume piece rate contracts. However, our model and empirical analysis differ from Barlevy's in two main dimensions.

First, we use matched employer-employee data and put strong emphasis on both firm heterogeneity and individual (non i.i.d.) productivity shocks, whereas he uses NLSY data and thus cannot

²Yet, the strategic complete information game that he designs yields a slightly different outcome than in these two latter references.

³Postel-Vinay and Turon (2005) show that the combined assumptions of on-the-job search and wage renegotiation by mutual consent (taken up from Postel-Vinay and Robin's (2002) sequential auctions model) 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.

separate the different sources of heterogeneity. Second, wage-posting fails at describing the empirical relationship between wage and productivity convincingly because of the lack of firm competition for workers at the upper tail of the productivity distribution (Mortensen, 2003). As a consequence, to account for the long right tail of wage distributions, wage posting implies unrealistically long right tails for productivity distributions. By allowing firms to counter outside offer, the sequential auction model of Postel-Vinay and Robin (2002) increases firm competition in a natural way, and yields a wage equation that fits well the empirical relationship between observed firm output and wages (Cahuc *et al.*, 2006).

Yamaguchi extends Postel-Vinay and Turon's setup by allowing for bargaining in addition to Bertrand competition using a strategic bargaining game as in Cahuc *et al.* The model is extremely complex and is solved numerically. It is estimated on NLSY1979 data by indirect inference. Yamaguchi finds that on-the-job search, by improving the value of the outside option with respect to unemployment, raises the wage of workers with ten-year experience by 13%, which accounts for a quarter of total wage growth over that period. On the other hand, counterfactual simulations also show that when deterministic human capital accumulation is cut off, wages are reduced by an approximately equivalent amount of 15%.

Our model and empirical analysis differ from Yamaguchi's also in two main dimensions. First, the availability of matched employer-employee data allows us to account for firm productive heterogeneity. Second, we allow for individual productivity shocks that persist across job spells. By contrast, Yamaguchi's model only allows for match-specific productivity shocks. Every outside offer, irrespective of the time already spent on the job market, is a draw from the same sampling distribution. Hence, the stochastic component of human capital is entirely general in our model and it is entirely firm- or occupation-specific in Yamaguchi's model. Needless to say that our model is also considerably easier to simulate and estimate, thanks to the piece rate contract assumption.

With these restrictions imposed, our 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, that explain,

in particular, the frequent earnings cuts that we observe.⁴

A second important aim of the empirical literature on Mincer equations is to disentangle 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 recent contributions include the comparative study by Beffy et al. (2005) and Dustmann and Meghir (2005). Beffy et al. 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. Dustmann and Meghir, using German data, find positive return to experience and firm tenure for skilled workers; small returns to experience, zero returns to sector tenure and large returns to firm tenure for unskilled workers. In this paper we take a further step in the direction of structural modeling. In addition to general experience effects and exogenous individual productivity shocks, our model generates wage mobility that is potentially different within and between job spells because by moving jobs the employer's characteristics change independently of worker characteristics. (Note that this is only possible because of labor market imperfections.) We will test the capacity of our model to replicate standard measures of tenure and experience effects. This will tell us if, on the top of search frictions, one needs to allow for another source of within-firm earnings dynamics, such as firm-specific human capital, to explain the returns to tenure.

The third 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 (e.g. Hall-Mishkin, 1982; MaCurdy, 1982; Abowd and Card, 1989; Topel and Ward, 1992; Gottshalk and Moffitt, 2002; Alvarez, Browning and Eyrnæs, 2001; Meghir and Pistaferri, 2004; Guiso, Pistaferri and Schivardi, 2005). This literature assumes competitive labor markets and wage is thus always thought to be equal to productivity. Piece-rate contracts offer a nice and simple theoretical vehicle for transmitting individual productivity shocks, however persistent, to wages in presence of market imperfections.

⁴When we write wage, we mean annual earnings. Most data, and administrative data are no exception, generally do not distinguish between contractual wage and bonuses. Bonuses may absorb all observed earnings cuts.

We estimate our structural model on Danish matched employer-employee data using indirect inference or simulated GMM. Our findings can be summarized like this: First, we find that our structural model encompasses commonly used reduced form models for analyzing earnings profiles (the “Mincer equation”) and earnings dynamics rather neatly. Second, with respect to the empirical decomposition of individual wage growth into human capital effects and job-search-related within- and between-job effects, we find that the within-job search effects dominates the between-job effects. In relative terms, human capital accumulation is quantitatively more important for wage growth in the early phases of workers’ careers, and its quantitative importance increases in workers’ educational attainment. Indeed, human capital accumulation is the primary source for early career wage growth among high-educated workers.

The paper is organized as follows. In section 2 we spell out the details of the theoretical model and in section 3 and 4 we present the data and the estimation protocol. In Sections 5 and 6 we show estimation results and analyze the decomposition of individual wage-experience profiles that motivated the paper. Section 7 concludes.

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 sold in a perfectly competitive market. Time is discrete and the economy is at a steady state. Workers can either be unemployed or matched with a firm. 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. Let t denote the number of periods that a worker has spent in employment since leaving school. Call it experience. 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:

$$h_t = \alpha + g(t) + \varepsilon_t, \quad (2)$$

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 that only change when the worker is employed. 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.

Timing of events within the period. The set of random events affecting a worker within a typical period includes job destruction, receipt of a job offer arrival, occurrence of a productivity shock, or exit from the labor force (this latter shock, to which we shall loosely refer as ‘retirement’, is there only to capture the observed attrition from the sample that we use for inference). 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 general human capital accumulation stops during unemployment spells, 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 (with $\mu + \delta + \lambda_1 \leq 1$). Similarly, any unemployed worker finds a new match with probability λ_0 , (such that $\mu + \lambda_0 \leq 1$). 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 $f(\cdot) = F'(\cdot)$,

⁵At 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.

with support $[p_{\min}, p_{\max}]$. For simplicity we neglect the possibility of two different events occurring simultaneously.

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 $R = e^r \leq 1$ is the endogenous contractual piece rate.

The rules governing the determination of the contractual piece rate are borrowed from the sequential auction 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 alternative employer with probability λ_1 at the end of the current period. The alternative 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 $\mathbb{E}_t V(\cdot)$ 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 $\mathbb{E}_t V(0, h_{t+1}, p)$ —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

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

condition:

$$\mathbb{E}_t V(r', h_{t+1}, p') = \mathbb{E}_t V(0, h_{t+1}, p). \quad (3)$$

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 $\mathbb{E}_t V(r', h_{t+1}, p) = \mathbb{E}_t V(0, h_{t+1}, p')$. 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:

$$\mathbb{E}_t V(r, h_{t+1}, p) = \mathbb{E}_t V(0, h_{t+1}, q(r, h_t, p)). \quad (4)$$

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 experience level t solves:

$$\mathbb{E}_t V(r_0, h_{t+1}, p) = V_0(h_t), \quad (5)$$

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:

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

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:

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

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

$$V(r, h_t, p) = \mathbb{E}_t \sum_{s=0}^{\infty} \left(\frac{1 - \mu - \delta}{1 + \rho} \right)^s \left\{ r + p + h_{t+s} + \frac{\delta}{1 + \rho} V_0(h_{t+s}) + \int_{q(r, h_{t+s}, p)}^p \frac{\lambda_1 \bar{F}(x)}{\rho + \mu + \delta} dx \right\}. \quad (8)$$

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

$$q(r, h_t, p) = \frac{\rho + \mu + \delta}{1 + \rho} r + p - \int \int_{q(r, h_{t+1}, p)}^p \frac{1 - \mu - \delta - \lambda_1 \bar{F}(x)}{1 + \rho} dx dM(h_{t+1}|h_t), \quad (9)$$

where $M(\cdot|h_t)$ 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:

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

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.

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

$$w_{it} = p_{j(i, t)} + \alpha_i + g(t) + \varepsilon_{it} - \int_{q_{it}}^{p_{j(i, t)}} \left(1 + \frac{\lambda_1 \bar{F}(x)}{\rho + \mu + \delta} \right) dx, \quad (11)$$

where $p_{j(i, t)}$ is worker i 's current employer's type and q_{it} 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: an experience effect $g(t)$,⁷ a worker fixed effect α_i , a transitory component ε_{it} , an employer fixed effect $p_{j(i,t)}$, and a random effect q_{it} relating to the most recent wage bargain.

The joint process governing the dynamics of $(p_{j(i,t)}, q_{it})$ can be characterized as follows:

$$\left(\begin{array}{c} p_{j(i,t+1)} \\ q_{i,t+1} \end{array} \right) \middle| \left(\begin{array}{c} p_{j(i,t)} \\ q_{it} \end{array} \right) = \left\{ \begin{array}{ll} \left(\begin{array}{c} p_{j(i,t)} \\ q_{it} \end{array} \right) & \text{with probability } 1 - \mu - \delta - \lambda_1 \bar{F}(q_{it}) \\ \left(\begin{array}{c} p_{j(i,t)} \\ p_{j(i,t)} > q' > q_{it} \end{array} \right) & \text{with density } \lambda_1 f(q') \\ \left(\begin{array}{c} p' > p_{j(i,t)} \\ p_{j(i,t)} \end{array} \right) & \text{with density } \lambda_1 f(p'). \\ \left(\begin{array}{c} \cdot \\ \cdot \end{array} \right) & \text{with probability } \mu \\ \left(\begin{array}{c} p' \\ b \end{array} \right) & \text{with probability } \delta f(p') \end{array} \right. \quad (12)$$

The last two rows characterize the following two possible events: first, the worker may retire (probability μ), in which case $(p_{j(i,t+1)}, q_{i,t+1})$ becomes unobserved forever, and second, the worker may become unemployed (probability δ), in which case $(p_{j(i,t+1)}, q_{i,t+1})$ 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 $(p_{j(i,t)}, q_{it})$ derived in Appendix A.⁸ Characterization of this steady-state distribution will be useful to simulate the model (see below section 4 and Appendix C).

3 Data

Background. The data used in the empirical analysis consist of a ten percent random sample of workers from the Danish register-based matched employer-employee dataset IDA, merged with detailed data on individual labor market histories covering the period 1986 to 1999. IDA contains annual socio-economic information on workers and background information on employers, and

⁷In the empirical analysis, $g(t)$ will be effectively made a function of experience as recorded in the survey.

⁸The other random components of wages appearing in (11) are exogenously distributed (α_i is just a fixed effect and ε_{it} follows an exogenous process of its own), and they are uncorrelated with $p_{j(i,t)}$ or q_{it} . 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 later in the analysis, though, there is assortative assignment based on experience.

covers the entire Danish population aged 16 to 69.⁹ The labor market history data is based on weekly reports on unemployment status and mandatory employer pension contributions. In the merging procedure, all labor market spells of a given worker in a given calendar 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 state 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 labor market history data distinguishes between four labor market states: Employment, temporary unemployment, unemployment and nonparticipation. Employment spells are associated with a firm identifier.¹⁰ 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. Likewise, nonparticipation spells shorter than 13 weeks are recoded as unemployment spells. Thus in the empirical analysis we distinguish between job spells, unemployment spells and nonparticipation spells. A job-to-job transition is a job transition with less than one week of work interruption.

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.¹¹

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

⁹IDA: *Integreret Database for Arbejdsmarkedsforskning* (Integrated Database for Labor Market Research) is constructed and maintained by Statistics Denmark.

¹⁰Employers 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.

¹¹An additional advantage of the indirect estimation technique applied in this paper (see section 4 below), is that it allows us to account for abnormal data features that might otherwise have caused serious problems in a straightforward fashion.

year. The experience information is available from 1 Jan 1964. We therefore discard workers born before 1 Jan 1948, since these cohorts might have accumulated experience before the measurement period was initiated. Since the period of observation ends ultimo 1999, the maximum age in the data is 51 years, effectively (and conveniently) ruling out effects of retirement considerations on the observed labor market behavior in our sample. 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.¹² Notice that we observe workers' *actual* (as opposed to potential) labor market experience.

Additional information on worker characteristics is annual and includes the standard covariates used in earnings regressions, of which we retain gender and years in education. From the employer side of the data we retain a public sector indicator. These variables are used for sample selection and stratification.

The analysis sample. We start out by discarding all workers with missing or inconsistent information on relevant variables. 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 in an otherwise stationary environment. 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 experience level of 5 years. Since we discarded all workers born prior to 1948, the workers in our analysis sample have at least 5 years of experience and are at most 51 years of age. 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 nonparticipa-

¹²ATP (*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.

tion are permanent). Fourth, we truncate a worker’s labor market history after the first observed transition into a public sector firm or into a firm with invalid firm identifier. Fifth, we stratify the initial sample into three levels of education, based on the number of years spent in education (9-11 years, 12 years and 13-18 years). This stratification is roughly in accordance with whether workers are unskilled, have a vocational or high school education, or have at least some higher education.

The strata-specific wage distributions are inflated to the 1999-level using Statistics Denmark’s consumer price index and trimmed at the 2.5 percentile and the 97.5 percentile to exclude abnormal wage observations. Despite the fact that our structural model does not explicitly feature aggregate technological progress, we do recognize the potential importance of such effects on individual earnings, and we therefore seek to purge the wage data of macro trends. Even though the raw data is a representative sample of the Danish labor force, the sample selection procedure implies that there is an important cohort-element in our analysis samples, simply because workers in the initial cross-section (in 1986) all have 21 or less years of experience. The ageing of this initial cross-section during our window of observation (from 1986 to 1999) makes it impossible to separate wage growth stemming from exogenous technological progress and endogenous experience accumulation using the successive cross-sections in the analysis panels. We circumvent this problem by trending up real wages to the 1999 level using the trend in earnings computed from a sequence of cross-sections of workers with 21 or less years of experience. This approach has the advantage over de-trending using, say, the trend in real GDP over the analysis period, that it allows us to compute strata-specific trends.

The stratified analysis samples are thus (unbalanced) panels where workers are followed in a period of up to 14 years in the private sector, containing information on earnings, the labor market states occupied, and experience. We will refer to these as the “master panels”. Our structural estimation procedure requires the calculation of a number of auxiliary statistics, which are obtained from different subsamples of the master panels. We provide relevant descriptive statistics on these subsamples below, as we discuss the structural estimation procedure in detail.

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 the technique of indirect inference. 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 a set of auxiliary models, which differs from the original structural model that we aim to estimate.

The technical details of indirect inference are spelled out in e.g. Gouriéroux, Monfort and Renault (1993), but for completeness and because we report the asymptotic standard errors of the estimated structural parameters, we recapitulate the asymptotic theory of our estimator in Appendix B.

4.2 Empirical specification

Indirect inference only requires that the structural model can be simulated given a value of the vector of structural parameters. To that end, functional form assumptions about the sampling distribution of firm productivity $F(\cdot)$, the distribution of worker heterogeneity $H(\cdot)$, the specification of the idiosyncratic productivity shock process ε_t 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 Weibull distribution:

$F(p) = 1 - e^{-[\nu(p-b)]^\omega}$ for $p \geq b$.¹³ We next assume that log-worker effects α are normally distributed

¹³The Weibull distribution is rather flexible and can resemble the normal distribution, while it contains the truncated exponential as a special case ($\omega = 1$), which in turn is equivalent to a Pareto distribution for productivity in

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 $H(\cdot) = \mathcal{N}(0, \sigma^2)$. The idiosyncratic productivity shocks are assumed to follow an AR(1) process with normal innovations: $\varepsilon_t = \eta\varepsilon_{t-1} + u_t$, $u_t \sim \mathcal{N}(0, \sigma_u^2)$. Finally, we specify the deterministic trend in individual productivity as a piecewise linear function with knots $\{t_1^*, t_2^*, t_3^*\} = \{5, 10, 15\}$ years of experience¹⁴, i.e. $g(t) = \sum_{k=1}^3 \gamma_k(t - t_k^*)\mathbf{1}(t > t_k^*)$, thereby introducing an additional set of structural parameters $(\gamma_1, \gamma_2, \gamma_3)'$. The structural parameter vector can thus be spelled out as $\theta = (\mu, \delta, \lambda_0, \lambda_1, \omega, \nu, b, \sigma, \eta, \sigma_u, \gamma_1, \gamma_2, \gamma_3)'$.¹⁵

The period length is set to one month, 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 given a value of θ are given in Appendix C.

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. As we argued in the introduction, we relate our analysis to the empirical labor literature concerned with wage equations and wage distributions. Our selection of auxiliary models reflects this link in that it will borrow from the specifications commonly used in the literature. Specifically, we will combine the following three auxiliary models: a duration model based on job search theory, a Mincer wage equation with worker and firm fixed-effects, and a first-differenced version of the Mincer equation as a model of within-job 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.

Spell durations: a job search model. The auxiliary duration model is derived from a discrete-time, partial-equilibrium job search model. The basic environment of this latter model is much the

levels.

¹⁴Recall all workers in the sample have at least five years of experience.

¹⁵We do not estimate the discount rate ρ , which is fixed at a monthly value of $\rho = 0.005$.

same as the one of our structural model: jobseekers 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 . Upon receiving a job offer, workers draw the associated wage from a continuous wage offer distribution F^A .

Under this set of assumptions, employed workers accept any offer higher than their current wage and unemployed workers accept any job offer above their reservation wage. Because this model is not structural, we shall further force the lower bound of the wage offer distribution to be at least equal to the unemployed workers' 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 $\lambda_1^A \bar{F}^A(w)$. The job-to-unemployment hazard rate is δ^A and the hazard rate into nonparticipation is μ^A .

Note 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. Thus in particular the model does not permit job-to-job transitions with wage cuts.

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 a given sampling date (viz. November 1991) until their first observed transition (if any). For each worker we 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. Inspection of the data revealed that a non-trivial fraction of jobs end in the last week in December in each of the sample-years. While transitions patterns are likely to exhibit some seasonality, the clustering of job terminations is concentrated only in the last week of December, and is so pervasive that we decided to exclude all job spells that end in a job-to-job transition in the last week of December in any year.¹⁶ Descriptive statistics for this sample are given in Table 1.

¹⁶We only exclude jobs ending in a job-to-job transition because the the dating of unemployment periods is very reliable in our data. We do not exclude spells ending in a job-to-job transition the last week of December from the analysis samples used for estimation of the other auxiliary models. The reason for this seeming inconsistency is that the November stock sampling scheme applied makes the parameters of the wage posting model especially sensitive

	Ed. 9-11	Ed. 12	Ed. 13-18
Number of spells	7,068	12,334	2,396
<i>U-spells (in months):</i>			
Number of U-spells	509	687	78
Number of UJ transitions	485	667	75
Number of UN transitions	21	20	3
Number of censored U-spells	3	0	0
Mean (<i>S.D.</i>) noncens. res. duration./mts.	7.69 (<i>9.03</i>)	7.35 (<i>7.76</i>)	9.87 (<i>13.37</i>)
<i>J-spells (in months):</i>			
Number of J-spells	6,559	11,647	2,318
Number of JU transitions	1,605	2,471	247
Number of JJ transitions	2,344	4,837	1,072
Number of JN transitions	279	366	55
Number of censored J-spells	2,331	3,973	944
Mean (<i>S.D.</i>) noncens. res. duration/mts.	32.27 (<i>26.01</i>)	31.73 (<i>26.21</i>)	32.71 (<i>25.43</i>)
<i>Cross-Section Log-Wages:</i>			
Mean log-wage	5.0605	5.1471	5.4623
Std. dev. of log-wages	0.2204	0.2186	0.2656

Table 1: Sample descriptive statistics—Cross-section residual spell durations and wages

We estimate $\mu^A, \delta^A, \lambda_0^A, \lambda_1^A$ using the two-step semi-nonparametric estimation procedure developed in Bontemps, Robin and Van den Berg (2000). These estimates provide the first set of moments to match.

A Mincer wage equation for matched employer-employee data. The second auxiliary model is a Mincer 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).

The data used for the estimation of Mincer equation is extracted from the master panels of earnings, seniority and experience described above. Seniority in the job is constructed using the labor market history data, which implies that we cannot compute an accurate seniority measure for jobs that are ongoing at the start of the data period (Jan. 1 1986). To get round this problem we (somewhat unconventionally) introduce left-censored seniority as a distinct set of controls in

to this particular recording error. When simulating the estimation sample for the wage posting model we drop one fourth of all job spells ending in a job-to-job transition in “December”, i.e. in any period immediately following a recording date.

		Number of employers per worker						
	# Workers	1	2	3	4	5	6+	Average
Ed. 9-11	13,310	7,442	3,585	1,444	594	175	70	1.70
Ed. 12	23,934	11,868	7,116	3,101	1,277	390	182	1.82
Ed. 13-18	5,552	2,866	1,608	709	244	92	33	1.77

		Number of observations per worker						
	# Workers	1	2	3	4	5-10	11-14	Average
Ed. 9-11	13,310	2,454	1,425	1,129	924	3,875	2,982	6.49
Ed. 12	23,934	3,115	2,619	2,111	1,808	7,688	5,569	6.84
Ed. 13-18	5,552	676	674	523	470	1,861	1,105	6.55

		Observed firm size						
	# Firms	2-5	6-10	11-20	21-50	51-100	100+	Average
Ed. 9-11	4,995	4,204	477	193	86	27	8	4.53
Ed. 12	9,318	7,878	859	356	171	33	21	4.68
Ed. 13-18	1,896	1,497	220	117	49	12	1	5.19

Table 2: Sample descriptive statistics—Worker mobility—Mincer equation

the Mincerian equation. Since the raw data is a ten percent sample of workers appearing in IDA in the observation period, the observed firm sizes are not likely to equal actual firm sizes, and in fact, a substantial number of firms are only represented in the data by one worker. Clearly, separate identification of firm and worker effects is not possible in these cases. Hence, we exclude all firms that are only represented by one single worker in the sample, thereby effectively trimming the left tail of the distribution of firm effects in the data, since the eliminated firms are likely to be of low productivity. A similar problem arises for workers that are only represented in the data by one observation, rendering separation of worker effects and idiosyncratic noise impossible. We circumvent this problem by restricting calculation of the worker effects to the set of workers that are observed at least twice in the panel (after firms represented by only one worker have been excluded). Summing up, the Mincerian wage equation is estimated on an unbalanced matched employer-employee panel where each firm is represented by at least two workers. The panel features I workers (some of which are only observed once) and J firms, for a total of N observations. Table 2 provides a brief statistical summary of the sample.

Let s_{it} be individual i 's observed seniority in the firm $j(i, t)$ he is currently working at; t is the worker's experience, ψ_i a worker fixed-effect, ϕ_j a firm fixed-effect. The indicator z_{it} takes the value one if the observed seniority is left-censored (i.e. the job spell has started before the observation period), and the value zero otherwise. Then we consider the following regression:

$$w_{it} = \zeta_1(s_{it}) z_{it} + \zeta_2(s_{it}) (1 - z_{it}) + \zeta_3(t) + \phi_{j(i,t)} + \psi_i + u_{it}, \quad (13)$$

where u_{it} is a statistical residual, $\zeta_1(s_{it})$ and $\zeta_2(s_{it})$ are two parametric functions of seniority and $\zeta_3(t)$ is a parametric function of experience t . Specifically, we control for seniority and experience through a set of piecewise linear functions with knots at $\{0, 5\}$ years (seniority) and $\{5, 10, 15\}$ years (experience).

The structural wage equation (11) decomposes wages into a firm heterogeneity component $p_{j(i,t)}$, a fixed worker heterogeneity component α_i , a human capital component $g(t)$, idiosyncratic productivity shocks ε_{it} , and labor market frictions through the last integral term in (11). The auxiliary wage equation (13) assumes 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} in the notation of the structural model) is unobserved in the data, and the integral term of the structural wage equation (11) is therefore absent from the conditioning set in the auxiliary model. Seniority $s_{i,j(i,t)}$ can be viewed as a proxy for this factor.

Imposing the restriction that the worker-specific effect ψ_i has mean zero and is orthogonal to all other components of the wage equation,¹⁷ we can estimate (13) by first regressing log-earnings on the piecewise linear functions in seniority and experience as well as a full set of firm dummies.¹⁸ Worker effects are subsequently recovered as the within-worker average residuals.

From the auxiliary wage equation we include in the set of moments to match the estimated seniority and experience effects $\widehat{\zeta} = (\widehat{\zeta}_0, \widehat{\zeta}_1, \widehat{\zeta}_2)$, the average firm effect, the standard error of the firm effects, the standard error of the worker effects and finally the standard error of the residuals.

¹⁷Note that this restriction is also imposed in the structural model, where fixed worker heterogeneity is orthogonal to all other stochastic components in the model.

¹⁸We run the full regression in one step using the "SparseSolve" routine in GAUSS, but the estimator is equivalent to the within-firm estimator of (13).

Within-job wage growth equation. Using the auxiliary wage equation (13) 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 on (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 and $t+1$. Taking first differences in equation (13) under this restriction yields the following auxiliary model for within-job wage growth (when seniority and experience, both measured in years, enter in piecewise linear functions as explained above):

$$\Delta w_{it} = \xi_0 + \xi_1 \Delta(t-10)\mathbf{1}(t > 10) + \xi_2 \Delta(t-15)\mathbf{1}(t > 15) + v_{it}. \quad (14)$$

First-differencing within job spells eliminates the firm and worker fixed heterogeneity components. Moreover we only include experience in the r.h.s. of 14 as, within a job spell, experience and seniority are undistinguishable.

We estimate (14) directly rather than using the estimated residuals \hat{u}_{it} from (13) for two reasons. First, contrary to \hat{u}_{it} , the residuals \hat{v}_{it} from (14) are not affected by estimation errors on the firm and worker effects. Second, the estimation of (14) provides us with additional slope parameters $\hat{\xi} = (\hat{\xi}_0, \hat{\xi}_1, \hat{\xi}_2)$ which convey information and can be incorporated into the set of moments to match.

To obtain the estimation sample of the auxiliary wage growth equation we impose only one selection criteria on the original master panel: for a job spell to be included in the estimation sample it must contain at least two consecutive annual wage observations, so as to make first differencing is possible. If a job spell has several “disconnected” stretches of consecutive annual wage observation, we include all such stretches into the estimation sample.¹⁹ More consecutive observations will be needed when we later compute residual autocovariances from (14). We report autocovariances up to order 3, so that autocovariances are computed from the subset of jobs with at least five observations. Table (3) contains descriptive statistics on the estimation sample for the auxiliary wage growth equation.

Equation (14) is estimated by OLS. Using the resulting vector of estimated residuals \hat{v}_{it} we next compute the first three within-job autocovariances for the residuals of each of the J jobs in the analysis sample. The sample autocovariances are then obtained as the average job-specific

¹⁹This situation might arise if a job is observed for, say, 11 years with the first five years having valid wage information, the sixth wage observation being invalid, and the last five years again having valid wage information.

	Ed. 9-11	Ed. 12	Ed. 13-18
Number of within-job wage cycles with ≥ 2 consecutive obs.	22,301	40,900	9,906
Number of within-job wage cycles with 5 consecutive obs.	2,162	3,905	885
Number of within-job wage cycles with 6 consecutive obs.	1,403	2,569	600
Number of within-job wage cycles with 7 consecutive obs.	1,047	1,840	449
Number of within-job wage cycles with 8 consecutive obs.	871	1,541	340
Number of within-job wage cycles with > 8 consecutive obs.	3,329	5,722	1,186
Average log-wage; 5-10 years of experience	5.02	5.10	5.39
Average log-wage; 10-15 years of experience	5.07	5.16	5.51
Average log-wage; 15-20 years of experience	5.11	5.19	5.56
Average annual within-job log-wage growth	0.0090	0.0096	0.0222

Table 3: Sample descriptive statistics—Within-job wage growth equation

autocovariances taken over the J jobs in the analysis sample. From the auxiliary wage growth model (14) we include the estimated slope parameters $\hat{\xi}$ and residual autocovariances up to order three.

Summary. We thus have a set of 21 moments that we seek to match using our structural model: the four transition parameters of the auxiliary job search model $(\hat{\mu}^A, \hat{\delta}^A, \hat{\lambda}_0^A, \hat{\lambda}_1^A)$, the seven slope parameters of the wage equation (13) $(\hat{\zeta}_1, \dots, \hat{\zeta}_7)$, the first and second moments of the firm effects and the second moment of the worker effects and residuals in that same equation, and finally the three slope parameters $(\hat{\xi}_0, \hat{\xi}_1, \hat{\xi}_2)$ and the three autocovariances of residuals from the wage growth equation (14).

We tested our indirect inference procedure using a Monte-Carlo study on 100 replications of small simulated samples (1,000 workers followed over 168 periods, i.e. fourteen years). The design of our (admittedly small-scale) Monte-Carlo study is as follows: We take the estimated structural model for high-educated workers as our data generating process, and compute the set of moments just described. Next, we seek to recover the true structural parameters by applying our estimation protocol, starting the iterative optimization scheme in the same parameter values as in our real-data estimation, but using a set of simulated worker trajectories different from that used to generate the true set of moments. This is repeated 100 times, keeping the simulation of the true model fixed

	True value	Mean est.	Est. ± 1.96 s.d.
<i>Transition parameters</i>			
μ			[;]
δ			[;]
λ_0			[;]
λ_1			[;]
<i>Firm productivity</i>			
ν			[;]
b			[;]
ω			[;]
<i>Worker productivity</i>			
σ			[;]
<i>Idiosyncratic shocks</i>			
η			[;]
σ_u			[;]
<i>Experience</i>			
γ_1			[;]
γ_2			[;]
γ_3			[;]

Table 4: Monte-Carlo study of the estimation procedure

across repetitions. Results of our MC study are reported in Table 4.

5 Estimation 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 duration model. Table 5 reports estimates of the first set of auxiliary parameters—namely, transition parameters of the auxiliary wage posting model—obtained from the real data. It also reports the corresponding estimates based on data generated by the structural model.

Estimates on the real data basically reflect the average spell durations and transition rates shown

	Ed. 9-11		Ed. 12		Ed. 13-18	
	Real	Sim.	Real	Sim.	Real	Sim.
μ^A	0.0010	0.0015	0.0009	0.0014	0.0009	0.0021
δ^A	0.0057	0.0060	0.0057	0.0051	0.0040	0.0031
λ_0^A	0.1192	0.1116	0.1325	0.1040	0.0977	0.0962
λ_1^A	0.0173	0.0091	0.0209	0.0081	0.0210	0.0096

Table 5: The auxiliary duration model

in the descriptive Table 1 (only adding a correction for right-censoring). Based on the numbers in Table 5, the predicted average unemployment spell durations are 8.3, 7.5 and 10.1 months for the low, medium and high-education groups, respectively while the corresponding average employment spell durations are 8, 7.8 and 10.1 years and the related average monthly probabilities of job-to-job transition are 0.0052, 0.0058 and 0.0052.²⁰

Turning to the simulation-based estimates, our structural model seems to replicate the estimates of δ^A and λ_0^A very accurately, while at the same time overestimating μ^A and rather strongly underestimating λ_1^A . Based on our estimation of the auxiliary search model on simulated data, we would predict average unemployment spell durations of 8.8, 9.5 and 10.2 months, average job spell durations of 8.1, 9.3 and 10.8 years and average monthly job-to-job transition probabilities of 0.0034, 0.0030 and 0.0032 (all these numbers pertain to the low, medium and high-education categories, respectively). The structural model thus notably understates job-to-job transition probabilities.

Rosholm and Svarer (2004) ran into similar problems in their analysis of a version of the wage

²⁰These average durations and probabilities are derived as follows (we remove the superscript A for simplicity). The average job spell duration conditional on wage w is $d(w) = 1/(\mu + \delta + \lambda_1 \bar{F}(w))$. It follows from the steady-state assumption that the cross-section distribution of w among employed workers is

$$G(w) = \frac{(\mu + \delta) F(w)}{\mu + \delta + \lambda_1 \bar{F}(w)}$$

(see equation A9 in appendix A for a derivation). The average job spell duration is then obtained by integrating w out of the conditional duration:

$$\bar{d} = \int d(w) dG(w) = \frac{\mu + \delta + \lambda_1/2}{(\mu + \delta)(\mu + \delta + \lambda_1)}.$$

The average unemployment spell duration is obviously equal to $1/(\mu + \delta)$. Average job-to-job transition probabilities are derived following similar steps and equal

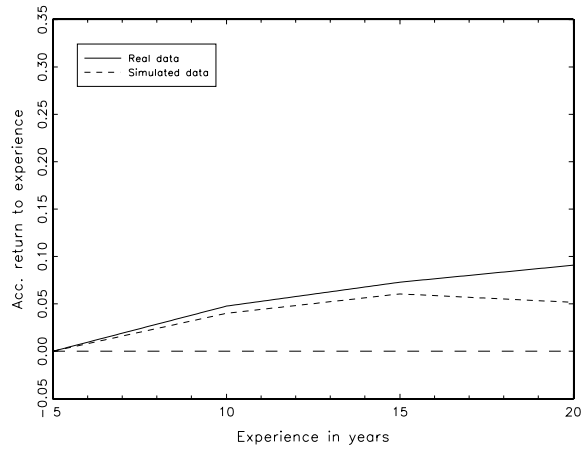
$$\int \lambda_1 \bar{F}(y) dG(w) = (\mu + \delta) \left[\left(1 + \frac{\mu + \delta}{\lambda_1} \right) \ln \left(1 + \frac{\lambda_1}{\mu + \delta} \right) - 1 \right].$$

Note that these expressions are independent of the nature of the scalar index used by workers to rank jobs, and are therefore equally valid for the auxiliary duration model and for the structural model.

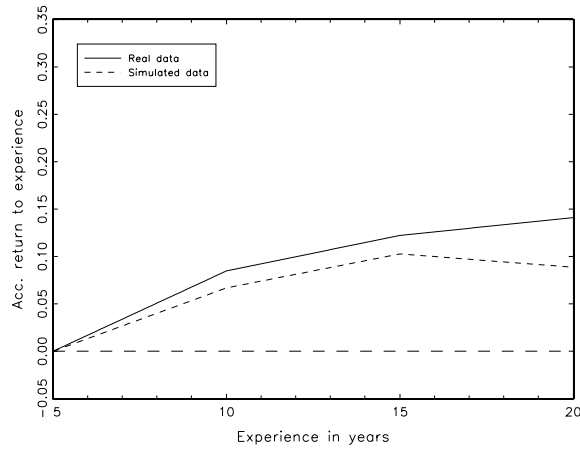
posting model with training investment estimated on the IDA data, overestimating average job durations and obtaining an overall poor fit to job durations. Rosholm and Svarer tend to find more mobility in the Danish data than we do, but their reported estimates relate to a sample period covering 1981-1990 (while ours relate to the period 1991-1999), and their sensitivity analysis reveals that job offer arrival rates decline in the latter years of their sample.

The auxiliary wage equation. Figures 1 and 2 show the experience and seniority profiles of individual wages as estimated from the auxiliary wage equation (13). In Figures 1 and 2, the solid line depicts the profile based on real data, while the dashed line relates to model-generated data. Finally, moments of the firm and worker heterogeneity distributions—again based on the auxiliary wage equation (13)—are reported in Table 6.

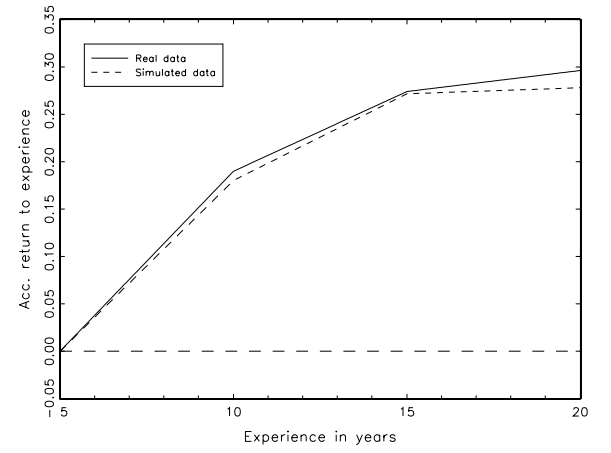
We first review estimates based on the real data. The auxiliary wage equation indicates positive returns to experience in all three subsamples (Figure 1). These are quantitatively rather modest for the low-educated group (who benefit from a 5 percent wage increase as they go from 5 to 10 years of experience, followed by a further 5 percent as they go from 10 to 20 years of experience), and become more substantial as we look at more educated workers (workers in the highest education group see their wages increase by 20 percent between 5 and 10 years of experience, and then by another 10 percent over the following 10 years of their careers). Notice that all workers in the analysis samples have at least five years of experience. Hence, to compare our estimated experience profiles with profiles estimated without conditioning on a minimum level of experience, our profiles should be topped with the wage growth due to experience accumulation during the first five years of employment. Returns to seniority, on the other hand, are predicted by equation (13) to be very small in all three subsamples (Figure 2). Workers typically enjoy a 2-3 percent pay increase in the first 5 years of a job spell, with a wage-tenure profile that remains essentially flat (if not slightly downward sloping) thereafter. This pattern of experience and seniority effects is consistent with previous studies based on IDA data: after correcting for unobserved heterogeneity, Bingley and Westergaard-Nielsen (2003) find annual returns to seniority increasing from 0.001 to 0.007 from 1886 to 1997, while Buhai, Portela, Teulings and van Vuuren (2006) find virtually no returns to tenure. Because the profiles plotted in Figures 1 and 2 emerge from (13), which is a mere auxiliary



Ed. 9-11

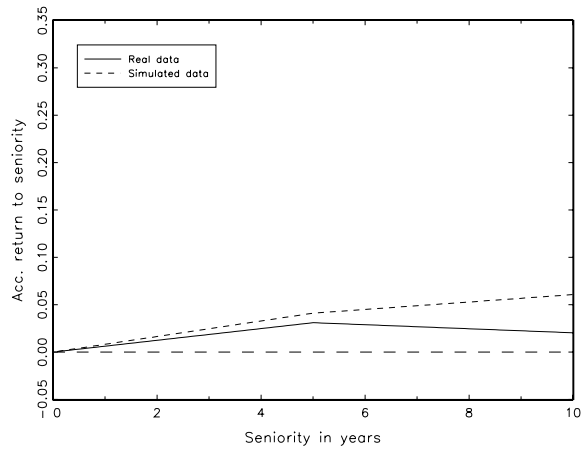


Ed. 12

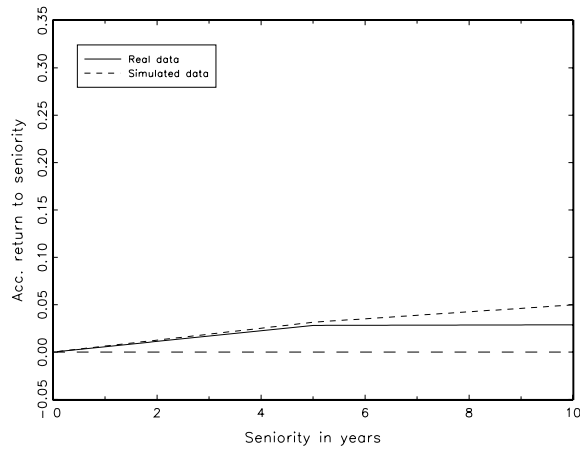


Ed. 13-18

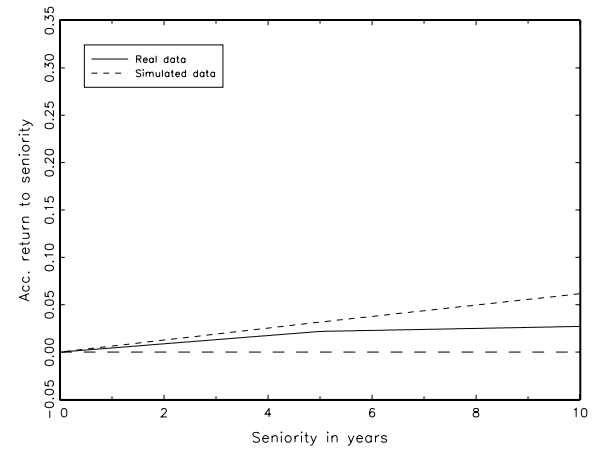
Figure 1: Model fit—Mincer equation experience profiles



Ed. 9-11



Ed. 12



Ed. 13-18

Figure 2: Model fit—Mincer equation tenure profiles

	Ed. 9-11		Ed. 12		Ed. 13-18	
	Real	Sim.	Real	Sim.	Real	Sim.
<i>Firm effects</i>						
Mean	5.0008	5.0102	5.0511	5.0617	5.3108	5.3380
Std. dev.	0.1696	0.1530	0.1640	0.1495	0.1856	0.1747
<i>Worker effects</i>						
Std. dev.	0.1186	0.1265	0.1251	0.1244	0.1699	0.1576
<i>Residual</i>						
Std. dev.	0.1120	0.1167	0.1170	0.1164	0.1159	0.1234
<i>Variance decomposition</i>						
$V(w)$	0.0453		0.0486		0.0700	
$V(X\zeta)$ (exp. & ten. effects)	0.0012		0.0023		0.0095	
$V(\phi)$ (firm effect)	0.0192		0.0193		0.0230	
$V(\psi)$ (person effect)	0.0120		0.0133		0.0253	
Corr ($X\zeta, \phi$)	0.0108		-0.0187		-0.0525	
Explained var.	0.0325		0.0347		0.0563	
Explained var. (% of $V(w)$)	71.5		71.8		80.4	

Table 6: The auxiliary wage equation (13): moments of heterogeneity distributions

model, we do not want to push the analysis of this pattern any further than saying that it is correctly picked up by our structural model, albeit with a slight tendency to understate experience effects and overstate tenure effects.

Next turning to Table 6, comparison of the firm and worker effect distributions across education groups hints at some degree of positive assortative matching on education, whereby more educated workers tend to be hired at firms with higher mean unobserved heterogeneity parameter. (This particular interpretation is of course conditional on the normalization of the mean worker effect at zero in all samples.) Moreover, dispersion of worker- (and, to an even smaller extent, firm-) effects tends to be slightly higher among more educated groups. All numbers in Table 6 are very accurately replicated by the structural model.

Equation (13) finally allows for a simple cross-sectional variance decomposition of log wages. The share of total log wage variance explained by (13) is about 70% for the two less educated groups and about 80% for the high-educated category. The decomposition (keeping in mind that it is based on a misspecified auxiliary model) reveals that individual characteristics are more important

	Ed. 9-11		Ed. 12		Ed. 13-18	
	Real	Sim.	Real	Sim.	Real	Sim.
<i>Autocovariances</i>						
Order 0	0.0132	0.0145	0.0130	0.0140	0.0112	0.0134
Order 1	-0.0044	-0.0054	-0.0042	-0.0052	-0.0033	-0.0044
Order 2	-0.0008	-0.0007	-0.0009	-0.0007	-0.0008	-0.0009
Order 3	-0.0004	-0.0004	-0.0005	-0.0004	-0.0004	-0.0006

Table 7: The auxiliary wage growth equation (14): autocovariance structure of residuals

in explaining wage dispersion among workers with a higher level of education where about 50% of log wage variance can be attributed to individual characteristics ($X\zeta + \psi$) and about 30% is due to firm effects (ϕ). Among workers in the two lower education groups the corresponding figures are 30% (individual characteristics) and 40% (firm effects). 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 (13) to U.S. and French data.

The auxiliary wage growth equation. Finally, results from the auxiliary wage growth equation (14) are reported in Figure 3, which plots the wage-experience profiles estimated from (14) both on real (solid line) and simulated (dashed line) data, and Table 7 which reports the autocovariance structure of unexplained wage growth based on (14).

We first look at the profiles, which in fact combine the returns to tenure and experience within a job spell. As one would expect based on estimation results for the wage equation in levels (13), these profiles are upward sloping for all education groups and steeper for more educated workers. Again this pattern is very well captured by the structural model.

We next turn to Table 7, which informs the stochastic dynamics of the residual in equation (14). Residual autocovariances decline sharply between one and two lags, and are essentially zero at longer lags. As is typically found in studies of individual earnings dynamics based on individual

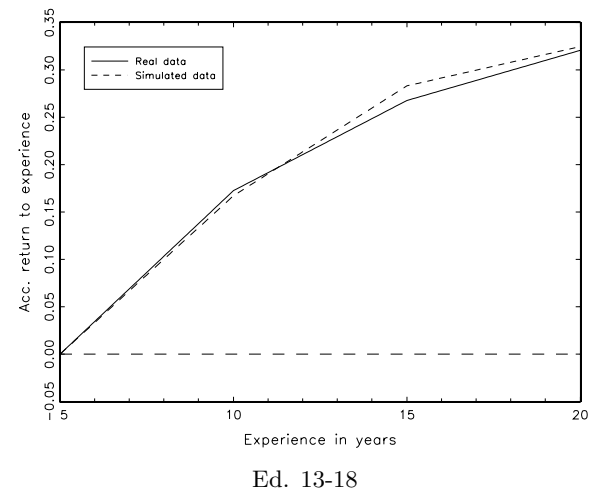
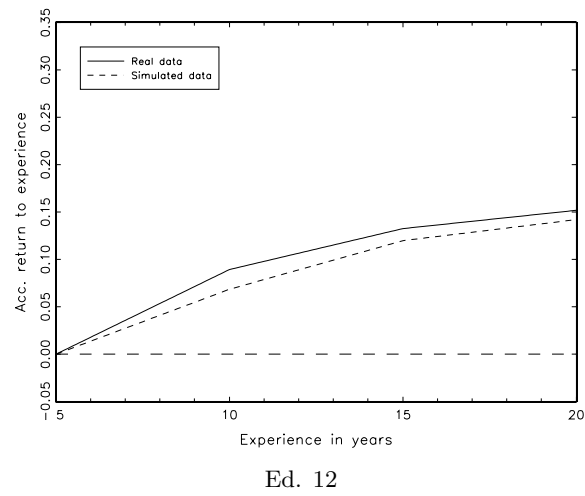
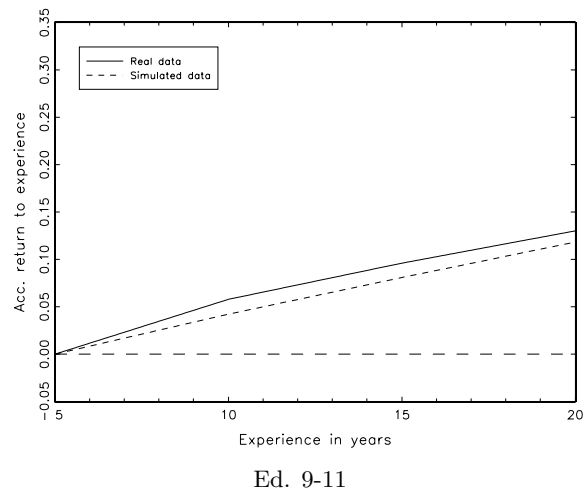


Figure 3: Model fit—Wage growth equation experience profiles

	Ed. 9-11	Ed. 12	Ed. 13-18
<i>Transition parameters</i>			
μ (attrition/birth)	0.0014 (0.0000)	0.0012 (0.0000)	0.0018 (0.0000)
δ (job-to-unemployment)	0.0055 (0.0000)	0.0047 (0.0000)	0.0028 (0.0000)
λ_0 (unemployment-to-job)	0.1203 (0.0000)	0.1259 (0.0002)	0.1056 (0.0000)
λ_1 (job-to-job)	0.0111 (0.0000)	0.0103 (0.0000)	0.0143 (0.0000)
<i>Firm productivity, $F(p) = 1 - e^{-[\nu(p-b)]^\omega}$</i>			
ν (scale)	3.8431 (0.0172)	3.9756 (0.0113)	3.7499 (0.0170)
b (location)	5.0622 (0.0020)	5.1242 (0.0016)	5.4026 (0.0021)
ω (shape)	1.5065 (0.0480)	1.7481 (0.0287)	1.6382 (0.0289)
<i>Worker productivity, $H(\alpha) = \mathcal{N}(0, \sigma^2)$</i>			
σ	0.0690 (0.0032)	0.0807 (0.0014)	0.0737 (0.0053)
<i>Idiosyncratic shocks, $\varepsilon_t = \eta\varepsilon_{t-1} + u_t, u_t \sim \mathcal{N}(0, \sigma_u^2)$</i>			
η	0.7997 (0.0015)	0.7963 (0.0006)	0.8385 (0.0006)
σ_u	0.0498 (0.0001)	0.0494 (0.0000)	0.0421 (0.0001)
<i>Experience, $g_t = \sum_{k=1}^3 \gamma_k(t - \tau_k)\mathbf{1}(t > \tau_k)$</i>			
γ_1 (knot $\tau_1 = 5$ years)	-0.0015 (0.0002)	0.0043 (0.0002)	0.0170 (0.0001)
γ_2 (knot $\tau_2 = 10$ years)	-0.0005 (0.0004)	-0.0044 (0.0002)	-0.0063 (0.0002)
γ_3 (knot $\tau_3 = 15$ years)	0.0011 (0.0002)	-0.0037 (0.0001)	-0.0117 (0.0001)

Table 8: Structural parameter estimates

or household data, this is suggestive of a low-order MA structure. Our structural model is once again able to replicate this feature of the data.

5.2 Structural parameter estimates

Estimates of the structural model parameters are reported in Table 8. The very low standard errors not only reflect the huge number of observations but they also tell that the indirect inference procedure succeeds in identifying structural parameters precisely.

Job mobility. Parameters relating to labor market mobility (i.e, offer arrival, layoff and attrition probabilities) are reported in the top four rows of Table 8. They look quite similar to the corresponding estimates obtained from the auxiliary wage posting model (see Table 5): compared to these latter, the structural parameters are suggestive of slightly lower layoff rates and slightly

higher job offer probabilities. The translation of these structural parameter estimates into average durations and turnover rates is as follows: predicted average unemployment spell durations are 8.2, 7.9 and 9.3 months for the low, medium and high education category, respectively; average employment spell durations are 8.4, 9.6 and 11.3 years for these same groups; finally, the average monthly probabilities of experiencing a job-to-job transition are 0.0038, 0.0035 and 0.0040. To sum up, education protects from unemployment risk, increases the occurrence of job offers to employees but, maybe surprisingly, reduces the occurrence of job offers to unemployed.

The sampling distribution of firm productivity, $F(\cdot)$. Estimates of the parameters of $F(\cdot)$ are reported in rows 5 to 7 of Table 8. Perhaps more directly informative are the implied mean and variances of the relating sampling distributions. The mean sampled productivity is 5.30 for workers with 9-11 years of schooling, 5.35 for workers with 12 years of schooling and 5.64 for workers with 13-18 years of schooling. The corresponding variances are .0025, .0018 and .0022. Finally, the lower support of $F(\cdot)$ is the b parameter, which is directly available from Table 8.

There appears to be a clear and statistically significant ranking of the three education groups in terms of mean sampled productivity, which is also reflected in the lower supports of the sampling distributions. Indeed plotting all three sampling distributions on the same graph reveals more generally that $F_{\text{ed. 9-11 yrs}} \prec F_{\text{ed. 12 yrs}} \prec F_{\text{ed. 13-18 yrs}}$ in the sense of first-order stochastic dominance. A similar (unreported) plot of the corresponding cross-sectional distributions of firm types $L(p)$, which are deduced from the estimated sampling distributions $F(\cdot)$ and transition parameters μ , δ and λ_1 using equation (A9), shows that the same FOSD-ordering holds true for these cross-sectional distributions, thus confirming the presence of positive sorting by educational attainment.

Worker heterogeneity. Row 8 of Table 8 contains the estimated standard deviation of the distribution of worker fixed, innate ability, α . These do not differ much between education groups. Indeed the only two variances that are statistically significantly different are those for the low and medium education groups. Interestingly, the structural model estimates a much lower person-effect variance than the auxiliary Mincer equation (in level). This is certainly due to the fact that the person effect in the auxiliary equation captures the persistence generated by the AR(1) idiosyncratic

shock ε_{it} .

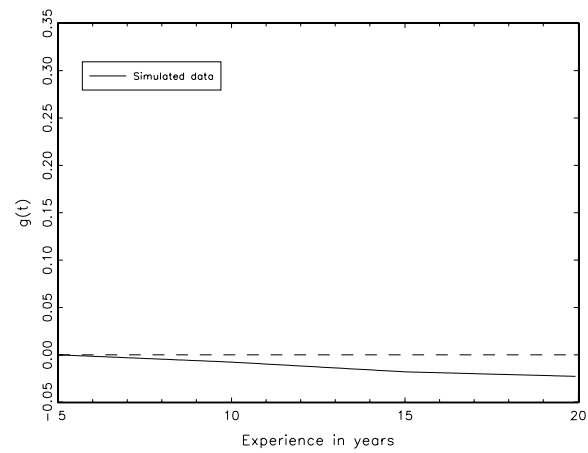
The stochastic component of individual productivity, ε_t . Parameters of the assumed monthly AR(1) process followed by ε_{it} are reported in rows 9 and 10 of Table 8. These parameters suggest that the dynamics of ε_{it} are quite similar across groups, with a tendency towards less dispersed innovations and more persistence in the highest education group. Up to these small differences, the autoregression coefficient η hovers around 0.8 for all three groups. This, however, is based on a period length of one month, and translates into a very small annual coefficient of 0.226 which is small but not negligible.²¹

The deterministic trend of human capital accumulation, $g(t)$. Finally, the bottom 3 rows of Table 8 contain the parameters of the deterministic trend in individual human capital accumulation, $g(t)$. For added legibility, these trends are also plotted in Figure 4.

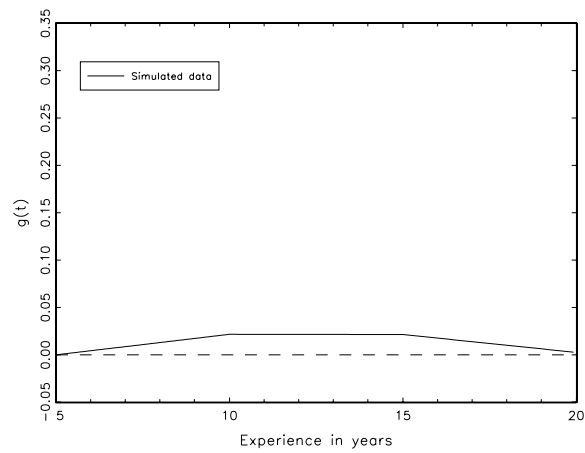
The differences between education categories in terms of human capital accumulation patterns are striking. Low-educated workers do not accumulate any productive skills as they become more experienced (if anything, they *lose* some productivity along the way), whereas at the other extreme, workers with more than 13 years of schooling grow about 15 percent more productive between the 5th and 15th year of their careers. The human capital profile then tapers off (and even declines slightly) for these high-educated workers towards the end of their working lives. Workers in the intermediate education group seem to accumulate a small amount of human capital between 5 and 10 years of experience (making their individual productivity increase by a modest 2.5 percent), which they eventually start losing after 20 years of experience. We should emphasize once more that the estimated profiles do not cover the first five years of employment in an individual's working life. Any human capital accumulation that takes place in the first five years of employment is thus picked up by the worker-effects. One can therefore not exclude that human capital accumulation mostly takes place (possibly at a very high rate) at early stages of the low-educated workers' careers. In any case, an immediate consequence of our findings is that labor market competition,

²¹If $\varepsilon_t = \eta\varepsilon_{t-1} + u_t$, the correlation of $\varepsilon_t + \dots + \varepsilon_{t-k}$ with $\varepsilon_{t-k-1} + \dots + \varepsilon_{t-2k}$ can be shown to be

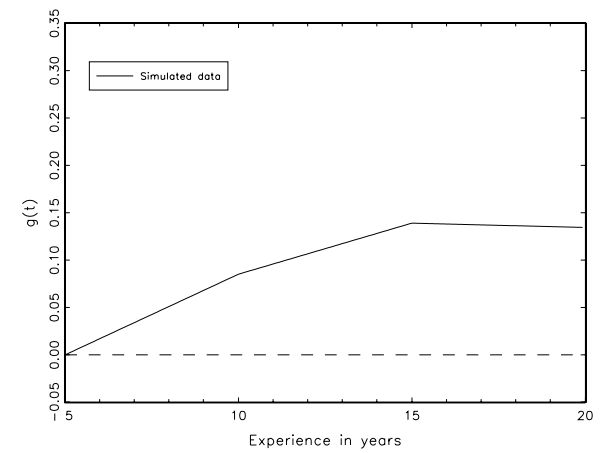
$$\frac{\eta(\eta^k + \eta^{k-1} + \dots + \eta + 1)}{2(\eta^k + 2\eta^{k-1} + 3\eta^{k-2} + \dots + k\eta) + k + 1}.$$



Ed. 9-11



Ed. 12



Ed. 13-18

Figure 4: Structural human capital-experience profile $g(t)$

not human capital accumulation, is going to be identified by our structural model as the primary cause of post-schooling wage growth for low-educated worker category. We now look at this specific issue in more detail.

6 The decomposition of post-schooling wage growth

The main motivation for this paper was the analysis of the underlying processes driving individual wage growth. With the theoretical apparatus in place and the estimated parameters in hand, we are now in a position to address this question in details.

Analysis. Starting from the wage equation (11) and the characterization of wage dynamics in (12), period-to-period wage growth goes as follows:

$$\begin{aligned}
w_{t+1} - w_t &= h_{t+1} - h_t \quad \text{with probability } 1 - \mu - \delta - \lambda_1 \bar{F}(q_t) \\
&= h_{t+1} - h_t + \int_{q_t}^{q'} \left(1 + \frac{\lambda_1 \bar{F}(x)}{\rho + \delta + \mu} \right) dx \quad \text{with density } \lambda_1 f(q'), \quad q' \in (q_t, p_t) \\
&= h_{t+1} - h_t + p' - p_t - \int_{p_t}^{p'} \left(1 + \frac{\lambda_1 \bar{F}(x)}{\rho + \delta + \mu} \right) dx \\
&\quad + \int_{q_t}^{p_t} \left(1 + \frac{\lambda_1 \bar{F}(x)}{\rho + \delta + \mu} \right) dx \quad \text{with density } \lambda_1 f(p'), \quad p' > p_t \\
&= \cdot \quad \text{(unobserved) with probability } \mu + \delta.
\end{aligned}$$

Hence, conditional on staying employed between experience levels t and $t + 1$, the following holds true (where $\frac{\lambda_1}{1-\mu-\delta}$ includes this latter conditioning):

$$\begin{aligned}
\mathbb{E}[w_{t+1} - w_t | p_t, q_t, t] &= \underbrace{\mathbb{E}(h_{t+1} - h_t | t)}_{\text{human capital}} \\
&\quad - \underbrace{\frac{\lambda_1}{1 - \mu - \delta} \int_{p_t}^{+\infty} f(x) \int_{p_t}^x \frac{\lambda_1 \bar{F}(z)}{\rho + \delta + \mu} dz dx + \frac{\lambda_1}{1 - \mu - \delta} \bar{F}(p_t) \int_{q_t}^{p_t} \left(1 + \frac{\lambda_1 \bar{F}(x)}{\rho + \delta + \mu} \right) dx}_{\text{search, between-job spells}} \\
&\quad - \underbrace{\tilde{\lambda}_1 \bar{F}(p_t) \int_{q_t}^{p_t} \left(1 + \frac{\lambda_1 \bar{F}(x)}{\rho + \delta + \mu} \right) dx + \tilde{\lambda}_1 \int_{q_t}^{p_t} \bar{F}(x) \left(1 + \frac{\lambda_1 \bar{F}(x)}{\rho + \delta + \mu} \right) dx}_{\text{search, within-job spell}}. \tag{15}
\end{aligned}$$

Note that two of the middle terms cancel out if one seeks simplification. Further note that the first term—reflecting the contribution of human capital accumulation to wage growth—simply equals the deterministic trend $g(t + 1) - g(t)$. Finally, the conditioning variables q_t and p_t can be integrated

out using the conditional distributions derived in equations (A8) and (A12) (see Appendix A):

$$\mathbb{E}(w_{t+1} - w_t|t) = \int_{p_{\min}}^{+\infty} \int_{p_{\min}}^p \mathbb{E}[w_{t+1} - w_t|p, q, t] dG(q|p, t) dL(p|t). \quad (16)$$

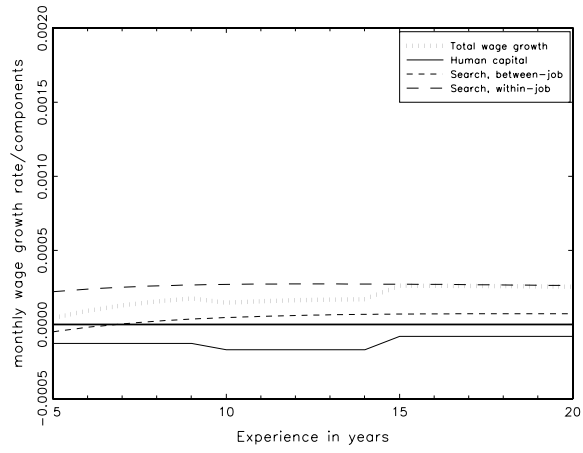
We thus end up with a “natural” additive decomposition of monthly wage growth (conditional on experience) into a term reflecting the contribution of human capital accumulation and two terms reflecting the impact of job search, within and between job spells.²²

Results. The decomposition of monthly wage growth given by (15) and (16) is rendered graphically as a function of work experience in Figure 5. In all three panels, the dotted line represents overall monthly wage growth given experience, $\mathbb{E}(w_{t+1} - w_t|t)$, while the solid, short-dashed and long-dashed lines represent the contributions to wage growth of human capital accumulation, the “between-job spells” component of job search and the “within-job spells” component of job search, respectively, as described in (15).

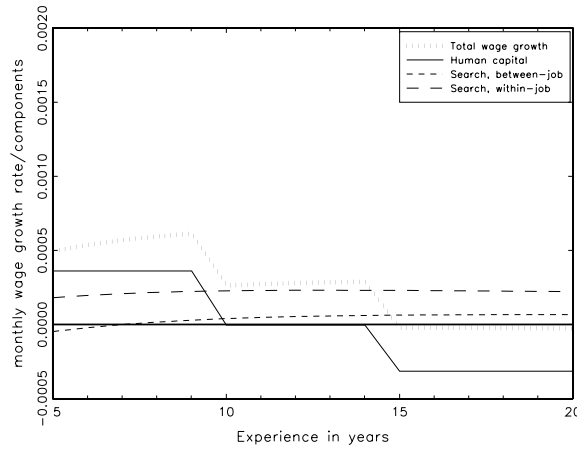
Figure 5 reveals that both job-search-related profiles are strikingly similar across worker categories. Both the within- and between-job spells components are essentially flat (indeed very slightly hump-shaped). The component reflecting between-job wage growth is everywhere below its within-job counterpart, which is a reflection of the option value of taking up a job at a higher- p firm that makes workers willing to take temporary wage cuts in exchange for better future career prospects. Both profiles are slightly higher for the high-educated group than for the other two groups—which reflects a combination of a higher offer arrival rate λ_1 , a lower overall job destruction rate $\delta + \mu$, and better job opportunities in the sense of a stochastically dominant sampling distribution of firm types, $F(\cdot)$ —but this is arguably the only noticeable difference between education groups in terms of the overall contribution of job search to expected wage growth.

Human capital accumulation profiles, on the other hand, are very different between worker groups. As seen before on Figure 4, low-educated workers do not accumulate any human capital past five years of experience; if anything, the human capital accumulation profile contributes slightly negatively to total wage growth throughout working life for this group. The rate of human capital accumulation varies much more across experience levels for the other two groups. Again as was

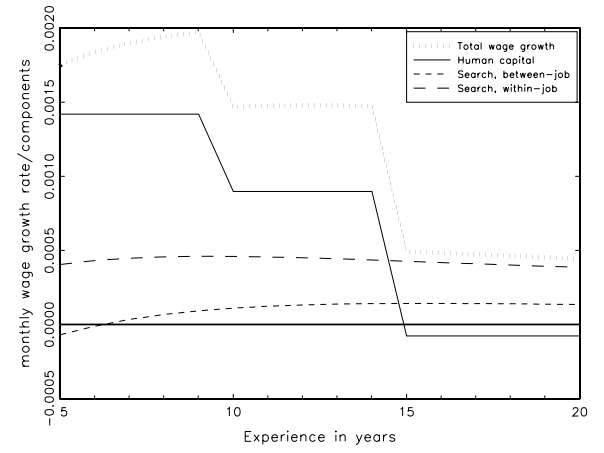
²²Obtaining a closed-form expression for the above decomposition is cumbersome, though feasible. Numerical integration is straightforward, however.



Ed. 9-11

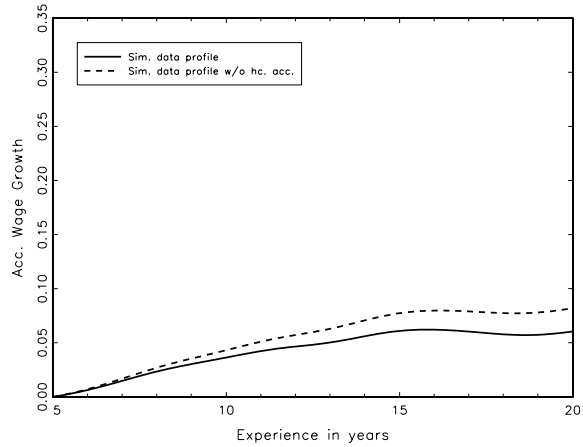


Ed. 12

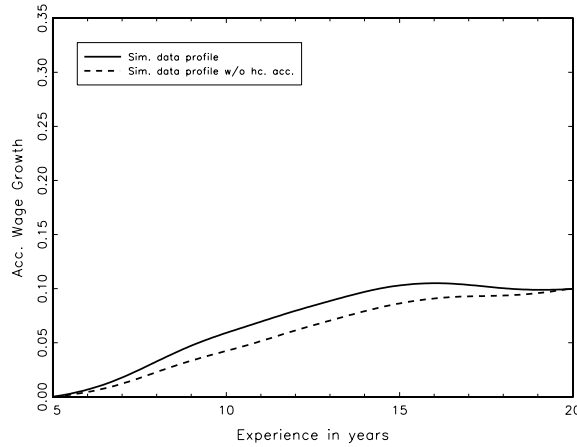


Ed. 13-18

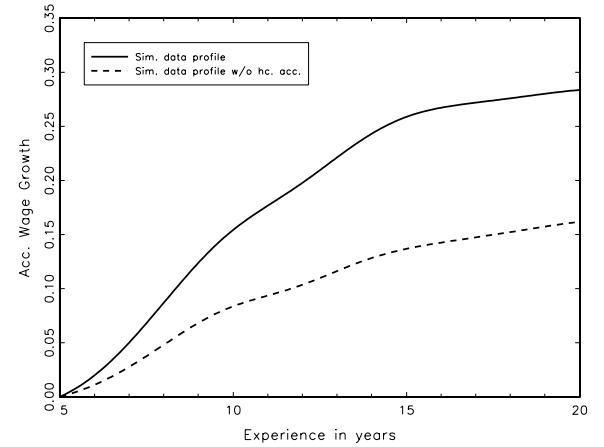
Figure 5: Decomposition of monthly wage growth



Ed. 9-11



Ed. 12



Ed. 13-18

Figure 6: Counterfactual wage-experience profiles—Turning off deterministic experience trends

evident from the concavity of $g(t)$ on Figure 4, human capital growth declines rather steeply over the working life for medium- and high-educated workers. These two worker categories, however, differ markedly in the relative importance of human capital (vis-à-vis job search) as an engine of wage growth. The contribution of human capital accumulation to total wage growth for high-educated workers is roughly 75% between five and ten years of experience, 50% between ten and fifteen years of experience, and zero thereafter. Turning to individuals with twelve years of education, human capital accumulation explains about a half of total wage growth for workers between five and ten years into their working lives. This contribution drops to zero over the following five years of experience and becomes negative (sufficiently so to cancel the positive impact of job search) from fifteen years of experience onwards.

Whatever the level of confidence one is prepared to place in these specific numbers, the following stylized patterns seem to arise for medium- and high-educated workers. The *absolute* contribution of job search to wage growth is roughly constant over the working life, whereas the *absolute* contribution of human capital accumulation declines sharply with experience. In *relative* terms, human capital is quantitatively most important in explaining wage growth at early stages of an individual's working life, and its contribution becomes nil or even negative as workers grow more experienced. It is also the case that the concavity of individual wage-experience profiles is mostly due to this decline in the the rate of human capital accumulation over the life cycle.

The picture of individual profiles of wage growth drawn by Figure 5—which plots monthly wage *growth rates* as a function of experience—can be usefully complemented by its “cumulative” counterpart, shown in Figure 6. This Figure shows a plot of a piecewise linear function of experience fitted to a sample of model-generated wages (solid line), together with a plot of a similar regression of counterfactual simulated wages where human capital accumulation was shut down, i.e. $g(t)$ was set identically equal to zero (dashed line). The differences between these two profiles provides a measure of the relative importance of human capital accumulation and between-firm labor market competition in explaining cumulated individual wage growth.

Comparison of the overall wage-experience profile and the counterfactual profile, reveals that cumulated wage growth for high-educated workers would on average be about a half of what it really is at any point between five years of experience and the end of their working life if their

stock of human capital stayed constant over time. Human capital accumulation is somewhat less important for workers with only twelve years of schooling as keeping their initial level of human capital constant would reduce their cumulated wage growth by roughly 20% between five and fifteen years of experience. Finally, as expected from results already shown, the low-educated group would, if anything, fare slightly *better* if they could maintain their stock of human capital at its initial level.

7 Conclusion

With the purpose of analyzing the causes of individual wage growth, we constructed a tractable equilibrium search model of individual worker careers allowing for human capital accumulation, employer heterogeneity and individual level shocks, which we estimate on Danish matched employer-employee data. The estimation procedure provides a platform of comparison between our structural model and commonly used reduced form models in three strands of the empirical labor literature, namely the “human capital” literature, the “wage dynamics” literature and the “job search” literature. Our structural model encompass all these models and does a good job of replicating related findings, up to some difficulties in fitting transitions between jobs.

The main analytical results of the paper is to provide a decomposition of individual wage growth (conditional on experience) into a term reflecting the contribution of human capital accumulation and two terms reflecting the impact of job search, within and between job spells. We find that the job-search-related wage growth is similar across education groups with within-job effects dominating between-job effects. Human capital’s role in generating wage growth differ markedly across education groups, and vary considerably across experience levels. For low-educated workers, human capital accumulation is found to contribute slightly negatively (if anything) to total wage growth. For medium- and high-educated workers, the absolute contribution of human capital accumulation to wage growth declines sharply with experience, with high-educated workers accumulating substantially more human capital than medium-educated workers. In relative terms, human capital accumulation is quantitatively more important for wage growth early in workers’ careers, and its quantitative importance increases in workers’ educational attainment. Indeed, human capital accumulation is the primary source for early career wage growth among high-educated workers.

References

- [1] Abowd, J. M. and D. Card (1989), “On the Covariance Structure of Earnings and Hours Changes,” *Econometrica*, 57, 2, 411-445.
- [2] Abowd, J. M., F. Kramarz, P. Lengeremann and S. Roux (2003), “Interindustry and Firm-Size Wage Differentials in the United States and France”, Cornell University working paper.
- [3] Abowd, J. M., F. Kramarz, and D. N. Margolis (1999), “High Wage Workers and High Wage Firms”, *Econometrica*, 67, 251-333.
- [4] Altonji, J. G. and R. Shakotko (1987), “Do Wages Rise With Job Seniority?”, *Review of Economic Studies*, LIV (3), 437-60.
- [5] Altonji, J. G. and N. Williams (2004), “Do Wages Rise With Job Seniority? A Reassessment”, *Industrial and Labor Relations Review*. Forthcoming.
- [6] Alvarez, J., M. Browning and M. Ejrnæs (2001), “Modelling Income Processes with Lots of Heterogeneity”, CAM discussion paper 2002-01.
- [7] Baker, M. (1997), “Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings”, *Journal of Labor Economics*, 15 (2), 338-75.
- [8] 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>).
- [9] 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>).
- [10] Bingley, P. and N. Westergaard-Nielsen (2003), “Returns to Tenure, Firm Specific Human Capital and Worker Heterogeneity”, *International Journal of Manpower*, 24 (7), 774-88.

- [11] 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.
- [12] 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.
- [13] Buhai, S., M. Portela, C. Teulings and A. van Vuuren (2006), “Returns to Tenure: Time or Rank?”, paper presented at SOLE 2006.
- [14] Bunzel, H., B. J. Christensen, N. M. Kiefer and L. Korsholm (2000), “Equilibrium Search with Human Capital Accumulation”, in H. Bunzel, B. J. Christensen, N. M. Kiefer and D. T. Mortensen, editors, *Panel Data and Structural Labour Market Models*, North-Holland Publishing.
- [15] Burdett, K. and M. Coles (2003), “Equilibrium Wage-Tenure Contracts”, *Econometrica*, 71(5), 1377-1404.
- [16] Burdett, K. and D. T. Mortensen (1998), “Wage Differentials, Employer Size and Unemployment”, *International Economic Review*, 39, 257-73.
- [17] Dey, M. S. and C. J. Flinn (2005), “An Equilibrium Model of Health Insurance Provision and Wage Determination,” *Econometrica*, 73(2), 571-627.
- [18] Dustmann, C. and C. Meghir (2005), “Wages, Experience and Seniority”, *Review of Economic Studies*, 72, 77-108.
- [19] Gouriéroux, C., A. Monfort and E. Renault (1993), “Indirect Inference”, *Journal of Applied Econometrics*, 8, 85-118.
- [20] Guiso, L., L. Pistaferri and F. Schivardi (2005), “Insurance Within the Firm”, forthcoming *Journal of Political Economy*.
- [21] Hall, R. E. and F. S. Mishkin (1982), “The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households,” *Econometrica*, 50(2), 461-81.

- [22] Harris, M. et B. Holmström (1982), “A Theory of Wage Dynamics,” *The Review of Economic Studies*, 49, 315-333.
- [23] MaCurdy, T. E. (1982), “The Use of Time-Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis,” *Journal of Econometrics*, 18, 83-114.
- [24] Meghir, C. and L. Pistaferri (2004), “Income Variance Dynamics and Heterogeneity”, *Econometrica*, 72, 1-32.
- [25] Mincer, J. (1974), *Schooling, Experience and Earnings*, New-York: Columbia University Press.
- [26] Mortensen, D. (2003), *Wage Dispersion: Why Are Similar Workers Paid Differently?*, MIT Press.
- [27] Postel-Vinay, F. and J.-M. Robin (2002), “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity”, *Econometrica*, 70 (6), 2295-350.
- [28] Postel-Vinay, F. and J.-M. Robin (2006), “Microeconomic 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.
- [29] Postel-Vinay, F. and H. Turon (2005), “On-the-job Search, Productivity Shocks, and the Individual Earnings Process”, CEPR Discussion Paper No. 5593.
- [30] Ridder, G. and G. J. van den Berg (2003) “Measuring Labor Market Frictions: A Cross-Country Comparison”, *Journal of the European Economic Association*, 1(1), 224-44.
- [31] Rosholm, M. and M. Svarer (2004) “Endogenous Wage Dispersion in a Search-Matching Model”, *Labour Economics*, 11, 623-45.
- [32] Rubinstein, Y. and Y. Weiss (2005), “Post-Schooling Wage Growth: Investment, search and Learning”, manuscript, Tel-Aviv University.
- [33] Stevens, M. (2004), “Wage-Tenure Contracts in a Frictional Labour Market: Firms’ Strategies for Recruitment and Retention,” *The Review of Economic Studies*, 71, 535-551.

- [34] Topel, R. (1991), “Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority”, *Journal of Political Economy*, 99 (1), 145-76.
- [35] Topel, R. and M. Ward (1992): “Job mobility and the careers of young men,” *Quarterly Journal of Economics*, 107, 439–79.
- [36] Yamaguchi, S (2006), “Job Search, Bargaining, and Wage Dynamics,” mimeo, McMaster University.

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 $(p_{j(i,t)}, q_{it})$. 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 :

$$u = \frac{\mu + \delta}{\mu + \delta + \lambda_0}. \quad (\text{A1})$$

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:

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

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

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:

$$(\mu + \lambda_0) u a_0(0) = \mu. \quad (\text{A4})$$

Jointly solving those three equations, one obtains:

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

The corresponding cdf is obtained by summation:

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

(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 $L(p|t)$ 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 $L(p|t = 1) = F(p)$. For $t > 1$ workers can come from both employment and unemployment and the flow-balance equation determining $L(p|t)$ is given by:

$$L(p|t) a_1(t) = (1 - \mu - \delta - \lambda_1 \bar{F}(p)) L(p|t-1) a_1(t-1) + (\mu + \delta) a_0(t-1) F(p). \quad (\text{A7})$$

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 $L(p|t) = \Lambda_1(p) L(p|t-1) + \Lambda_2 F(p)$, with:

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

This last equation solves as:

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

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

$$L(p) = \frac{(\mu + \delta) F(p)}{\mu + \delta + \lambda_1 \bar{F}(p)}. \quad (\text{A9})$$

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, $G(q|p, t)$, which we now derive. For $t > 1$, the flow-balance equation determining $G(q|p, t)$ is given by:

$$G(q|p, t) \ell(p|t) a_1(t) = (1 - \mu - \delta - \lambda_1 \bar{F}(q)) G(q|p, t-1) \ell(p|t-1) a_1(t-1) \\ + \lambda_1 L(q|t-1) a_1(t-1) f(p) + (\mu + \delta) a_0(t-1) f(p), \quad (\text{A10})$$

where $\ell(p|t) = L'(p|t)$ 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 $G(q|p, t) = 1$, yields the differential version of (A7):

$$\ell(p|t) a_1(t) = (1 - \mu - \delta - \lambda_1 \bar{F}(p)) \ell(p|t-1) a_1(t-1) \\ + \lambda_1 L(p|t-1) a_1(t-1) f(p) + (\mu + \delta) a_0(t-1) f(p). \quad (\text{A11})$$

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

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

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

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

B Indirect inference estimation procedure

We begin by introducing the following notation. Let θ denote the vector of structural parameters, the true value of which is θ_0 , and let \mathbf{Y}_N designate our estimation sample (the observed data). For a given value of θ , we further designate by $\text{DGP}(\theta)$ the structural model under consideration. 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 $\text{DGP}(\theta_0)$, which makes \mathbf{Y}_N a function of the structural parameter set at its true value, θ_0 . We further assume that $\text{DGP}(\theta)$ can be simulated for any given value of θ . Formally:

Assumption A1 *The DGP is parametric with $y_n = g(u_n, \theta_0)$ for $\theta_0 \in \Theta$, and $U_N = (u_1, \dots, u_N)$ is an i.i.d. sequence of random vectors of given, known distribution D . Write $Y_N = g(U_N, \theta_0)$.*

Indirect inference then works through the following steps. First, a number of statistics $\beta_N(\theta_0) \equiv b_N(g(U_N, \theta_0))$ are produced. These are either computed directly from the raw data or comes from a set of auxiliary models. The notation purposely emphasizes that $\beta_N(\theta_0)$ is a function of the structural parameter at the true value θ_0 , even though the auxiliary models are misspecified since they will generally differ from the original (true) structural model $\text{DGP}(\theta_0)$. Assumptions 2 and 3 state the regularity conditions we impose on the mapping from the structural parameter vector to the auxiliary statistics:

Assumption A2 *The function $\beta_N(\theta) = b_N(g(U_N, \theta)) : \Theta \mapsto B(U_N)$ is asymptotically continuous and differentiable.*

Assumption A3 *For any $\theta \in \Theta$ and any i.i.d. sequence $U_N = (u_1, \dots, u_N)$, there exists $\beta_\infty(\theta)$ such that $\text{plim}_{N \rightarrow \infty} \beta_N(\theta) = \beta_\infty(\theta)$ and $\sqrt{N}[\beta_N(\theta) - \beta_\infty(\theta)] \xrightarrow{d} N(0, \Sigma(\theta))$.*

The function $\beta_\infty(\theta)$ is termed the “binding function” by Gouriéroux, Monfort and Renault. It is pivotal for indirect inference since it provides a link between the data (summarized by β_∞) and the structural parameter θ .

Second, given a parameter value θ , the structural model $\text{DGP}(\theta)$ is simulated S times in order to produce S simulated data sets, on which the compute the same auxiliary statistics as with the actual data. Specifically, let $\mathbf{U}_N^s = (u_1^s, \dots, u_N^s)$, $s = 1, \dots, S$, be S independent samples drawn from $\mathcal{D}^{\otimes N}$. Let $\mathbf{Y}_N^s(\theta) = g(\mathbf{U}_N^s, \theta)$ define an simulated data sample and let $\beta_N^s(\theta) = b_N(g(\mathbf{U}_N^s, \theta))$ be the corresponding auxiliary statistics. From the sequence of simulated auxiliary statistics we consider the mean: $\bar{\beta}_N^S(\theta) = \frac{1}{S} \sum_{s=1}^S \beta_N^s(\theta)$. Since we perform the estimation by education, we have no covariates in our estimation procedure, and the $\beta_N^s(\theta)$ are independent across simulations.

Third and finally, we seek the value of θ that minimizes the distance between $\bar{\beta}_N^S(\theta)$ and $\beta_N(\theta_0)$. Formally, the indirect inference estimator θ_N is defined by

$$\theta_N = \arg \min_{\theta \in \Theta} Q_N(\theta) \equiv \left[\bar{\beta}_N^S(\theta) - \beta_N(\theta_0) \right]' \Omega \left[\bar{\beta}_N^S(\theta) - \beta_N(\theta_0) \right], \quad (\text{B1})$$

where Ω is a positive definite weighting matrix. The choice of an auxiliary model—an issue we will return to below—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.

Under Assumptions 1, 2 and 3 the indirect inference estimator is well-behaved; indeed, by the usual techniques it is possible to show that

$$\sqrt{N}(\theta_N - \theta_0) \xrightarrow{d} \mathcal{N}(0, W(S, \Omega, \theta_0)), \quad (\text{B2})$$

where the covariance matrix $W(S, \Omega, \theta_0)$ is given as

$$W(S, \Omega, \theta_0) = \left(1 + \frac{1}{S}\right) [H^S(\theta)' \Omega H^S(\theta)]^{-1} H^S(\theta_0)' \Omega \Sigma(\theta_0) H^S(\theta_0) [H^S(\theta)' \Omega H^S(\theta)]^{-1}, \quad (\text{B3})$$

with $H^S(\theta) = \text{plim}_{N \rightarrow \infty} \partial \bar{\beta}_N^S(\theta) / \partial \theta'$ being the Jacobian of the auxiliary statistics with respect to the structural parameter vector. We estimate the covariance matrix of the auxiliary statistics by re-sampling the real data, and denote the resulting estimate by Σ_N . The limiting Jacobian of the auxiliary statistic at the true value of the structural parameter $H^S(\theta_0)$ is estimated by numerical differentiation of $\bar{\beta}_N^S(\theta)$ evaluated at $\theta = \theta_N$. We denote the estimate of $H^S(\theta_0)$ by H_N . In our empirical implementation we take $\Omega = \Sigma_N^{-1}$, and we report standard errors of our estimates θ_N obtained from $W_N = (1 + 1/S) [H_N' \Sigma_N^{-1} H_N]^{-1} H_N' H_N [H_N' \Sigma_N^{-1} H_N]^{-1}$. Finally, we take $S = 1$.

C 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 assume that the labor market is in steady state and draw the initial cross-section of workers according to the steady state distributions derived in Appendix A. Recall that the initial cross section of workers in the real data have 21 or less years of experience. Hence, to mimic this aspect of the real data we draw the initial cross-section of the simulated data, conditional on experience $t \leq 21 \times 12 = 252$ periods²³.

We begin with a sample of N workers for which we draw individual (log) heterogeneity parameters α from $N(0, \sigma^2)$. Next, we assign labor market states (employed or unemployed) to workers according to (A1), and conditional on workers' labor market states we draw labor market experience t , conditional on $t \leq 252$, according to $A_1(t)$ and $A_0(t)$ defined by (A6). Given workers' labor market states and experience t we assign employer productivity. Unemployed workers are assigned productivity b independent of t while employed workers with experience t are assigned employer productivity p according to $L(p|t)$ defined by (A8). The productivities of the last firms

²³The real analysis sample is also conditioned on workers having at least 5 years of experience. However, to allow workers to obtain 5 years of experience during our T -period simulation and thus enter the analysis sample, we do not condition the initial stock of workers on a minimum level of experience. Instead, we impose the condition when selecting the relevant samples from the simulated data.

from which the workers were able to extract the whole surplus in the offer matching game—the q 's—are drawn (conditional on p and $t > 1$) from $G(q|p, t)$ defined by (A12). Unemployed workers and employed workers with experience $t = 1$ are assigned $q = b$. Finally, we draw the initial value of the idiosyncratic productivity shock process—the ε 's—from $N(0, \sigma_u^2/(1 - \eta^2))$.

We give the following tweak to the draws in the steady state distributions. Firm types p are theoretically distributed according to the continuous sampling distribution $F(p') = 1 - e^{-[\nu(p'-b)]^\omega}$. Because the theoretical $F(\cdot)$ is continuous, a rigorous implementation of this would invariably produce (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 (13) 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 $j = 1, \dots, J$ and assign corresponding productivity levels of $p_j = F^{-1}(j/J)$. Next, to assign the p_j 's to workers (conditional on experience), we draw in the usual way a N -vector $z = (z_1, \dots, z_N)$ of realizations of $U[0, 1]$, and determine worker i 's firm type as $p_{ij(i,t)} = \arg \min_{x \in \{p_1, \dots, p_J\}} |L(x|t) - z_i|$. Similarly, worker i 's q is assigned (conditional on $p = p_j$ and $t > 1$) as $q_{it} = q_{it}(p_j) = \arg \min_{x \in \{p_1, \dots, p_{j-1}\}} |G(x|p_j, t) - y_i|$, where y_i is a draw from $U[0, 1]$. The resulting cross-section of workers is used as the initial state of the labor market for our 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, the worker's duration of stay in the current job or unemployment spell, and if employed, the worker's employer type p and threshold value $q(\cdot)$ determining the worker's piece rate. With this information we can construct a simulated analysis sample containing the same information as the real analysis sample—namely unbalanced panels with information on earnings, the labor market states occupied and experience.

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 employer²⁴ (p'), an increase in experience and we set the worker's duration of stay in his current spell 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 worker 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 $H(\cdot) \equiv N(0, \sigma^2)$.

The simulated data sets, which have monthly wage observations (computed using (11) and the

²⁴With respect to the sampling of firm types, we let workers draw firm ranks j (and hence corresponding productivity levels of $p_j = F^{-1}(j/J)$) uniformly in the same J -vector of active firms that was used in the drawing of the initial cross-section of workers in the steady state distributions.

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). Also, worker the simulated individual labor market histories are trimmed at the minimal experience level of 5 years.