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

Jesper Bagger[*]
University of Aarhus

Francois Fontaine[†]
Université de Strasbourg 3

Fabien Postel-Vinay[‡]
University of Bristol
and
Université de Paris 1

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

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:

---

[*]jbagger@econ.dk.

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

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

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

1 Introduction

We present a tractable equilibrium job search model of individual worker careers allowing for human capital accumulation, employer heterogeneity and individual-level shocks. We estimate our structural model on a panel of Danish matched employer-employee data and use it to analyze the determinants of individual wage dynamics. Our main motivation for doing this is to quantify the respective roles of human capital accumulation coming along with work experience and the forces of labor market competition activated by workers' job search behavior in shaping the profile of individual labor earnings over the life cycle. The model offers a decomposition of individual wage growth into human capital effects and job search effects. This latter component can be further decomposed into precise measures of within-job and between-job wage growth.

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

---

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

[2] Earlier combinations of job search and human capital accumulation include Bunzel, Christensen, Kiefer and Korsholm (2000), Rubinstein and Weiss (2005) and Barlevy (2005). These three models are based on a wage posting approach which has restrictive implications for individual earnings dynamics, particularly within job spells (see the discussion below). The model of Christensen et al. is restrictive in a number of other ways, with the human capital production function being linear, workers reaping all benefit from human capital accumulation, and complete depreciation of workers' human capital upon lay-offs. Moreover, the model, which is estimated on Danish data similar

The second body of empirical work to which the present paper relates is the (equally large) literature on individual earnings dynamics. The long tradition of fitting flexible stochastic decompositions to earnings data has proved very useful in documenting the statistical properties of individual earnings from a dynamic perspective (Alvarez, Browning and Ejrnæs, 2001; Meghir and Pistaferri, 2004; Guiso, Pistaferri and Schivardi, 2005). However, the specific economic mechanisms underlying these empirical models remain vague, making it difficult to interpret the empirical evidence that they provide, or even to agree on what the "true" earnings process should be. The comparative analysis of Baker (1997) provides a very convincing illustration of this point. Also, a famous example where the lack of explicit structural foundations may have added to the heat and confusion of the academic debate is the attempts at disentangling the 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: our explicit formalization of the joint impact of experience (through human capital accumulation) and labor market mobility (through interfirm competition) on individual wages offers well-defined (if specific) measures of human capital effects, and within- and between-job wage growth. Our model enable us to base our discussion of the returns to tenure and experience, as picked up from the data by a standard Mincer equation, on those well-defined theoretical concepts, for which we exhibit precise empirical measures.

---

to the data used for our analysis, do not account for intrinsic firm and worker heterogeneity, or productivity shocks. Rubinstein's and Weiss' analysis of their model is essentially qualitative, and Barlevy provides a structural estimation based on NLSY data. Our contribution is to use matched employer-employee data and to put strong emphasis on firm heterogeneity and individual productivity shocks.

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

---

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

shocks.

We estimate our structural model on Danish matched employer-employee data using indirect inference (Gouriéroux, Monfort and Renault, 1993).[4] 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 which they sell in a perfectly competitive market. Time is discrete and the economy is at a steady state. For reasons that will shortly become clear, we consider individual (as opposed to calendar) time, which we index by t. Workers can either be unemployed or matched with a firm. They will transit between employment and unemployment as well as from job to job following a search process to be defined momentarily. Firms operate constant-return technologies and are modeled as a collection of job slots which can either be vacant

---

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

and looking for a worker, or occupied and producing.

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

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

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

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

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

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

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

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

3. *Job mobility shocks*: At the end of the period any employed worker leaves the market for good with probability μ, or sees her/his match dissolved with probability δ, or receives an

---

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

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 $f(\cdot) = F'(\cdot)$, with support $[p_{\min}, p_{\max}]$.

2.2 The wage equation

Wage setting rules. Wages are defined as piece rate contracts. If a worker supplies h_t units of efficient labor and produces $y_t = p + h_t$ (always in log terms), s/he receives a wage $w_t = r + p + h_t$, where $R = e^r \leq 1$ is the endogenous contractual piece rate.

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

Formally, the outcome of the Bertrand game can be described as follows. First, if $p' > p$ (the

---

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

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

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

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

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

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

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

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

where $V_0\left(h_t\right)$ 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 $V\left(r, h_t, p\right)$ is then defined recursively as:

$$V\left(r, h_t, p\right) = w_t + \frac{\delta}{1 + \rho} V_0\left(h_t\right) + \frac{1}{1 + \rho} \mathbf{E}_t \left\{ \left[1 - \mu - \delta - \lambda_1 \overline{F}\left(q\left(r, h_t, p\right)\right)\right] \cdot V\left(r, h_{t+1}, p\right) \right.$$

$$\left. + \lambda_1 \overline{F}\left(p\right) \cdot V\left(0, h_{t+1}, p\right) + \lambda_1 \int_{q(r, h_t, p)}^{p} V\left(0, h_{t+1}, x\right) dF\left(x\right) \right\}, \tag{6}$$

8

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

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

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

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

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

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

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

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

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.

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

9

identifiers—is defined as follows:

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

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

$$\begin{pmatrix} p_{j(i,t_i+1)} \\ q_{i,t_i+1} \end{pmatrix} \bigg| \begin{pmatrix} p_{j(i,t_i)} \\ q_{it_i} \end{pmatrix} = \begin{cases} \begin{pmatrix} p_{j(i,t_i)} \\ q_{it_i} \end{pmatrix} & \text{with probability } 1 - \mu - \delta - \lambda_1 \overline{F}(q_{it_i}) \\[2ex] \begin{pmatrix} p_{j(i,t_i)} \\ p_{j(i,t_i)} > q' > q_{it_i} \end{pmatrix} & \text{with density } \lambda_1 f(q') \\[2ex] \begin{pmatrix} p' > p_{j(i,t_i)} \\ p_{j(i,t_i)} \end{pmatrix} & \text{with density } \lambda_1 f(p'). \end{cases}$$

$$\qquad (12)$$

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

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

3   Data

Background.   The data used in the empirical analysis consist of a ten percent random sample of workers from the Danish register-based matched employer-employee dataset IDA, merged with

---

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

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 covers the entire Danish population aged 16 to 69.[8] 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.[9] We treat temporary unemployment as employment and aggregate job spells that are interrupted by temporary unemployment into a single job spell of duration equal to the sum of durations of actual employment periods and of periods of temporary unemployment. 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.[10]

---

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

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

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

Besides information on labor market transitions and earnings, the most important piece of information for the purpose of this study is workers' labor market experience. This information is available on an annual basis (from IDA) and refers to the workers' experience at the end of a calendar year. 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, ATP.[11] 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

---

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

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 nonparticipation 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 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 indirect inference techniques. The technical details of indirect inference are spelled out in e.g. Gouriéroux, Monfort and Renault (1993; GMR), but for completeness and because we report the asymptotic standard errors of the estimated structural parameters, we briefly recap the asymptotic theory of our estimator. Readers familiar with indirect inference may want to skip directly to subsection 4.2 where we describe the specific details of our implementation.

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

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 $\mathrm{DGP}(\theta)$ 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 $\mathrm{DGP}(\theta_0)$, which makes $\mathbf{Y}_N$ a function of

14

the structural parameter set at its true value, θ_0. We further assume that $\mathrm{DGP}(\theta)$ can be simulated for any given value of θ. Formally:

Assumption 1 *The DGP is parametric with $y_n = g(u_n, \theta_0)$ for $\theta_0 \in \Theta$, and $\mathbf{U}_N = (u_1, ..., u_N)$ is an i.i.d. sequence of random vectors of given, known distribution $\mathcal{D}$. Write $\mathbf{Y}_N = g(\mathbf{U}_N, \theta_0)$.*

Indirect inference then works through the following steps. First, a number of statistics $\beta_N(\theta_0) \equiv b_N(g(\mathbf{U}_N, \theta_0))$ 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 $\mathrm{DGP}(\theta_0)$. Assumptions 2 and 3 state the regularity conditions we impose on the mapping from the structural parameter vector to the auxiliary statistics:

Assumption 2 *The function $\beta_N(\theta) = b_N(g(\mathbf{U}_N, \theta)) : \Theta \mapsto \mathcal{B}(\mathbf{U}_N)$ is asymptotically continuous and differentiable.*

Assumption 3 *For any $\theta \in \Theta$ and any i.i.d. sequence $\mathbf{U}_N = (u_1, ..., u_N)$, there exists $\beta_\infty(\theta)$ such that $\mathrm{plim}_{N \to \infty} \beta_N(\theta) = \beta_\infty(\theta)$ and $\sqrt{N}[\beta_N(\theta) - \beta_\infty(\theta)] \xrightarrow{d} \mathcal{N}(0, \Sigma(\theta))$.*

The function $\beta_\infty(\theta)$ is termed the "binding function" by GMR and 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 $\mathrm{DGP}(\theta)$ 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 $\mathbf{U}_N^s = (u_1^s, ..., u_N^s)$, $s = 1, ..., S$, be S independent samples drawn from $\mathcal{D} \otimes \mathcal{D} \otimes ... \otimes \mathcal{D}$ (N times). 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: $\overline{\beta}_N^S(\theta) = \frac{1}{S} \sum_{s=1}^S \beta_N^s(\theta)$. 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 $\overline{\beta}_N^S(\theta)$ and $\beta_N(\theta_0)$. Formally, the indirect inference estimator θ_N is defined by

$$\theta_N = \arg\min_{\theta \in \Theta} Q_N(\theta) \equiv \left[\overline{\beta}_N^S(\theta) - \beta_N(\theta_0)\right]' \Omega \left[\overline{\beta}_N^S(\theta) - \beta_N(\theta_0)\right], \tag{13}$$

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

$$\sqrt{N}(\theta_N - \theta_0) \xrightarrow{d} \mathcal{N}(0, W(S, \Omega, \theta_0)), \tag{14}$$

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

$$W(S, \Omega, \theta_0) = \left(1 + \frac{1}{S}\right) [H^S(\theta)'\Omega H^S(\theta)]^{-1} H^S(\theta_0)'\Omega\Sigma(\theta_0)H^S(\theta_0)[H^S(\theta)'\Omega H^S(\theta)]^{-1}, \tag{15}$$

with $H^S(\theta) = \text{plim}_{N\to\infty}\partial\overline{\beta}_N^S(\theta)/\partial\theta'$ 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 $\overline{\beta}_N^S(\theta)$ 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 $S = 1$.

4.2 Empirical specification

Indirect inference only requires that the structural model $\text{DGP}(\theta)$ can be simulated given a value of the parameter vector θ. To that end, functional form assumptions about the sampling distribution of firm productivity $F(\cdot)$, the distribution of worker heterogeneity $H(\cdot)$, 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.

16

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 $p \geq b$.[12] We next assume that log-worker effects α are normally distributed among workers. As the productivity of a match equals the sum of the firm and the worker effect $(p + \alpha)$ the two distributions are only identified up to a normalization. We normalize the mean worker effect to zero, hence assuming $H(\cdot) = \mathcal{N}(0, \sigma^2)$. The idiosyncratic productivity shocks are assumed to follow an $AR(1)$-model with autoregressive parameter η, and with innovations coming from a zero-mean normal distribution with variance σ_u^2, i.e. $\varepsilon_t = \eta \varepsilon_{t-1} + u_t$, $u_t \sim \mathcal{N}(0, \sigma_u^2)$. 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 experience[13], i.e. $g_t = \sum_{k=1}^{3} \gamma_k (t - t_k^*) \mathbf{1}(t > t_k^*)$, thus introducing an additional set of structural parameters $(\gamma_1, \gamma_2, \gamma_3)'$. The structural parameter vector can thus be spelled out as $\theta = (\mu, \delta, \lambda_0, \lambda_1, \omega, \nu, b, \sigma, \eta, \sigma_u, \gamma_1, \gamma_2, \gamma_3)'$.[14]

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 $\mathrm{DGP}(\theta)$ given a value of θ are given in Appendix B.

4.3 Auxiliary Models

Outline. The choice of an auxiliary model, i.e. the choice of "which moments to match" is a crucial step in the indirect inference approach. As we argued in the introduction, we relate our analysis to three different strands of the empirical labor literature, namely the "human capital" literature, the "wage dynamics" literature and the "job search" literature. Our selection of auxiliary models reflects this three-fold link in that it will borrow from the specifications commonly used in each of these strands of literature. Specifically, we will combine the following three auxiliary models: a duration model based on job search/wage posting theory, a Mincerian wage equation

---

[12]The 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 levels.

[13]Recall all workers in the sample have at least five years of experience.

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

with standard human capital variables included as regressors which we supplement with firm fixed-effects, and a first-differenced version of the Mincerian 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.

We now give details of our three auxiliary models in turn. We distinguish between the structural and auxiliary parameters by adding a superscript A to the auxiliary parameters wherever confusion might arise. The auxiliary models are estimated on different subset of the master panels described in section 3, and the presentation of the auxiliary models therefore includes a brief data description.

Spell durations: a job search/wage posting model. The auxiliary duration model is derived from a discrete time version of the job search/wage posting model of Burdett and Mortensen (1998). The basic environment of this latter model is much the same as the one of our structural model: 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. The difference with the model considered in this paper is that in the wage-posting model, firms "post wages", i.e. commit to paying a constant wage over the duration of the match and do not engage in offer-matching. As a consequence, from the workers' standpoint, the value of a job is uniquely characterized by the wage w it pays. Upon receiving a job offer, workers draw the associated wage from a continuous wage offer distribution F^A.[15]

Under this set of assumptions it can be shown that workers optimally follow a reservation wage strategy, whereby employed workers accept any offer higher than their current wage and unemployed workers accept any job offer (with the lower bound of the wage offer distribution being at least equal to the unemployed workers' common reservation wage). Hence, the hazard rate out

---

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

of an unemployment spell is simply λ_0^A and the job-to-job transition probability of a worker with current wage w is $\lambda_1^A \overline{F}^A(w)$. The job-to-unemployment hazard rate is δ^A and the hazard rate into nonparticipation is μ^A.

The transition pattern in the auxiliary model is very close to that of the structural model, with the only difference pertaining to job-to-job transitions which occur on receipt of a higher wage offer in the wage posting model, and on receipt of an offer from a more productive (higher p) firm in our structural offer-matching/piece-rate model.[16] The pattern of worker mobility in the wage-posting model implies the following useful steady-state relationship between the wage offer distribution F^A and the distribution of earned wages in a cross-section of workers, denoted by G^A:

$$F^A(w) = \frac{\left(\mu^A + \delta^A + \lambda_1^A\right) G^A(w)}{\mu^A + \delta^A + \lambda_1^A G^A(w)}. \tag{16}$$

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

The data used for estimation of the auxiliary duration model consist of an initial cross-section of I employed or unemployed workers whom we follow from 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.[17] Descriptive statistics for this sample are given in Table 1.

---

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

[17]We 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

< **Table 1 about here.** >

Let d denote spell durations, ι_μ (resp. ι_δ, ι_λ) an indicator for transition into non-participation (resp. unemployment, employment). Finally, let c be a censoring indicator, i.e. $c = 1 - \iota_\mu - \iota_\delta - \iota_\lambda$ for employed workers and $c = 1 - \iota_\mu - \iota_\lambda$ for unemployed workers. We follow the two-step semi-nonparametric estimation procedure developed in Bontemps, Robin and Van den Berg (2000). The first step consists of using (16) to back out from the (observed) cross-sectional distribution of wages $G^A(\cdot)$ an estimator of the unobserved distribution of wage offers $F^A(\cdot)$ as a function of the transition parameters:

$$\widehat{F}^A(w) = \frac{\left(\mu^A + \delta^A + \lambda_1^A\right)\widehat{G}^A(w)}{\mu^A + \delta^A + \lambda_1^A \widehat{G}^A(w)},\tag{17}$$

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

$$\ell_u^A\left(\mu^A, \lambda_0^A\right) = c\,d\ln\{1 - \mu^A - \lambda_0^A\} + (1-c)\left[\iota_\mu\ln\mu^A + \iota_\lambda\ln\lambda_0^A + (d-1)\ln\{1 - \mu^A - \lambda_0^A\}\right],\tag{18}$$

while that of an employed worker reads as:

$$\begin{aligned}\ell_e^A\left(\mu^A, \delta^A, \lambda_1^A\right) = {}& c\,d\ln\{1 - \mu^A - \delta^A - \lambda_1^A\widehat{\overline{F}}^A(w)\}\\ & + (1-c)\left[\iota_\mu\ln\mu^A + \iota_\delta\ln\delta^A + \iota_\lambda\ln\{\lambda_1^A\widehat{\overline{F}}^A(w)\} + (d-1)\ln\{1 - \mu^A - \delta^A - \lambda_1^A\widehat{\overline{F}}^A(w)\}\right].\end{aligned}\tag{19}$$

Conditional on the worker's employment status ($u = 1$ for unemployed workers), the log-likelihood contribution of a generic worker in the sampled stock thus writes as $\ell^A\left(\mu^A, \delta^A, \lambda_0^A, \lambda_1^A\right) = u \cdot \ell_u^A\left(\mu^A, \lambda_0^A\right) + (1-u)\cdot\ell_e^A\left(\mu^A, \delta^A, \lambda_1^A\right)$. From the auxiliary duration model we include the estimated transition parameters $(\widehat{\mu}^A, \widehat{\delta}^A, \widehat{\lambda}_0^A, \widehat{\lambda}_1^A)$ in the vector of moments used in the indirect inference procedure.[18]

---

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 to this particular recording mistake. 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 a "December month".

[18]As a final remark about the identification of the structural transition parameters μ, δ, λ_0 and λ_1, it should

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

The data used for the estimation of Mincer wage 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 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 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.[19] 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

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

[19]Intuitively, the auxiliary firm effects identifies the structural distribution of firm effects F. However, it is important to stress that identification of F does not hinge on the estimated distribution of auxiliary firm effects being a precise estimate of F. In fact, even if all auxiliary firm effects could be identified from the data, their distribution would be a biased estimate of F, due to the omission of the structural integral term in the auxiliary wage equation (20).

2 provides a brief statistical summary of the sample.

< **Table 2 about here.** >

Now, let $s_{i,j(i,t_i)}$ be individual i's observed seniority in the firm $j(i, t_i)$ he is currently working at, t_i the worker's experience, ψ_i a worker fixed-effect, ϕ_j a firm fixed-effect. The indicator $z_{i,j(i,t_i)}$ takes the value one if the observed seniority is left-censored, and the value zero otherwise. Then we consider:

$$w_{it_i} = \zeta_1 s_{ij(i,t_i)} z_{i,j(i,t_i)} + \zeta_2 s_{ij(i,t_i)}(1 - z_{i,j(i,t_i)}) + \zeta_3(t_i - 5) + \phi_{j(i,t_i)} + \psi_i + u_{it_i}, \qquad (20)$$

where u_{it_i} is a statistical residual. When implementing (11) 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 $p_{j(i,t_i)}$, a fixed worker heterogeneity component α_i, a human capital component g_{t_i}, idiosyncratic productivity shocks ε_{it_i}, and labor market frictions through the last integral term in (11). The auxiliary wage equation (20) predicates a similar decomposition, up to the difference that the type of the last employer from which the worker was last able to extract surplus (q_{it_i} in the notation of 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 $s_{i,j(i,t_i)}$ 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[20], we can estimate (20) by first regressing log-earnings on the piecewise linear functions in seniority and experience as well as a full set of firm dummies.[21] Worker effects are subsequently recovered from the residuals. Specifically, write (20) in matrix form:

$$\mathbf{w} = \mathbf{X}\zeta + \mathbf{F}\phi + \mathbf{D}\psi + \mathbf{u}, \qquad (21)$$

where $\mathbf{w}$ is the $N \times 1$ vector of log-wages, $\mathbf{X}$ is the $N \times 7$ matrix of regressors, $\mathbf{F}$ is the $N \times J$

---

[20]This restriction is also imposed in the structural model, where fixed worker heterogeneity is orthogonal to all other stochastic components in the model. Our treatment of worker specific effects as (uncorrelated) fixed effects in the auxiliary wage equation (20) is therefore consistent with the structural model.

[21]We run the full regression in one step using the "SparseSolve" routine in GAUSS, but the estimator is equivalent to the within-firm estimator of (20).

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

$$(\widehat{\zeta}', \widehat{\phi}')' = ([\mathbf{X} : \mathbf{F}]' [\mathbf{X} : \mathbf{F}])^{-1} [\mathbf{X} : \mathbf{F}]' \mathbf{w}, \tag{22}$$

and the residuals as $\widehat{\mathbf{e}} = \mathbf{M}_{[\mathbf{X}:\mathbf{F}]}\mathbf{w}$, where $\mathbf{M_A} = \mathbf{I} - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$ is the usual least squares projection matrix. We can only recover worker effects from the subset of workers observed more than once in the data. Hence, if we let $\widetilde{\mathbf{u}}$ be the subset of residuals from $\widehat{\mathbf{e}}$ adhering to workers observed more than once, and let $\widetilde{\mathbf{D}}$ be the corresponding design matrix of worker indicators, the worker effects are recovered as the within-worker average residuals:

$$\widehat{\psi} = (\widetilde{\mathbf{D}}'\widetilde{\mathbf{D}})^{-1}\widetilde{\mathbf{D}}'\widetilde{\mathbf{u}}, \tag{23}$$

and the "real" residuals are given as $\widehat{\mathbf{u}} = \mathbf{M}_{\widetilde{\mathbf{D}}}\widetilde{\mathbf{u}}$. From the auxiliary wage equation we include the estimated slope parameters $\widehat{\zeta}$, 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.

Within-job wage growth equation. Using the auxiliary wage equation (20) 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_i and t_i+1. Taking first differences in equation (20) 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):

$$\Delta w_{it_i} = \xi_0 + \xi_1 \Delta(t_i - 10)\mathbf{1}(t_i > 10) + \xi_2 \Delta(t_i - 15)\mathbf{1}(t_i > 15) + \Delta u_{it_i}. \tag{24}$$

First-differencing eliminates the firm and worker fixed heterogeneity components. Moreover we only include experience in the r.h.s. of (24) as, within a job spell, experience and seniority are undistinguishable.

We estimate (24) directly rather than using the estimated residuals $\widehat{u}_{it}$ from (20) for two reasons. First, contrary to $\widehat{u}_{it}$, the residuals from (24) are not affected by estimation errors on the firm and worker effects. Second, the estimation of (24) provides us with additional slope parameters $\widehat{\xi} = (\widehat{\xi}_0, \widehat{\xi}_1, \widehat{\xi}_2)'$ which convey information and can be incorporated into the set of moments to match. Note that we thus do *not* impose consistency of coefficient estimates between the auxiliary log-wage equations in levels (20) and growth rates (24). According to our structural model, this pair of equation is a misspecified representation of the individual earnings process and one should therefore not expect it to be consistent in any particular way.

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[22]. More consecutive observations will be needed when we later compute residual autocovariances from (24). 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.

Under the maintained assumption that Δu_{it_i} is independent of t_i, (24) can be estimated by OLS. Doing this and using the resulting vector of estimated residuals $\widehat{\Delta \mathbf{u}}$ we next compute the first three within-job autocovariances for the residuals of each of the J jobs in the analysis sample as $\widehat{\Delta \mathbf{u}}_j' \mathbf{L}^k \widehat{\Delta \mathbf{u}}_j / N_j$ for $k = 0, 1, 2, 3$, where $\mathbf{L}$ is the lag operator, $\widehat{\Delta \mathbf{u}}_j$ is the vector of residuals pertaining to a particular job $j = 1, ..., J$, and $N_j \geq 4$ is the number of residuals of job j. The sample autocovariances are then obtained as the average job-specific autocovariances taken over the J jobs in the analysis sample. From the auxiliary wage growth model (24) we include the estimated slope parameters $\widehat{\xi'}$ and residual autocovariances up to order three.

---

[22]This 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.

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 $\left(\widehat{\mu}^A, \widehat{\delta}^A, \widehat{\lambda}_0^A, \widehat{\lambda}_1^A\right)$, the seven slope parameters of the wage equation (20) $\left(\widehat{\zeta}_1, \cdots, \widehat{\zeta}_7\right)$, 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 $\left(\widehat{\xi}_0, \widehat{\xi}_1, \widehat{\xi}_2\right)$ and the three autocovariances of residuals from the wage growth equation (24).

<center>< **Table 4 about here.** ></center>

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 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 job search/wage posting 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

<center>25</center>

structural model.

<center>< **Table 5 about here.** ></center>

Estimates on the real data basically reflect the average spell durations and transition rates shown in the descriptive Table 1 (only adding a correction for right-censoring). 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.[23]

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 wage posting 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 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.

---

[23]These average durations and probabilities are derived as follows. In any job search model where workers rank jobs according to some scalar index y (specifically, y is the wage in the case of the wage posting model and the firm's productivity p in the case of our structural model), the average job spell duration conditional on a value of y is $d(y) = 1/\left(\mu + \delta + \lambda_1 \overline{F}(y)\right)$, where F is the sampling distribution of y. It follows from the steady-state assumption that the cross-section distribution of y values among employed workers is $L(y) = (\mu + \delta) F(y) / \left(\mu + \delta + \lambda_1 \overline{F}(y)\right)$ (see equation (A9) in appendix A for a derivation). The average job spell duration is then obtained by integrating y out of the conditional duration: $\overline{d} = \int d(y) \, dL(y) = (\mu + \delta + \lambda_1/2) / \left[(\mu + \delta)(\mu + \delta + \lambda_1)\right]$. 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 $\int \lambda_1 \overline{F}(y) \, dL(y) = (\mu + \delta)(\mu + \delta + \lambda_1) \ln(1 + \lambda_1/(\mu + \delta))/\lambda_1 - \mu - \delta$. 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 wage posting model and for the structural model.

<center>26</center>

The auxiliary wage equation (20). Figures 1 and 2 show the experience and seniority profiles of individual wages as estimated from the auxiliary wage equation (20). 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 (20)—are also reported in Table 6.

<center>**< Figures 1 and 2, Table 6 about here. >**</center>

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 (20) 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 Westeraard-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 (20), which is a mere auxiliary 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.

<center>27</center>

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 (20) finally allows for a simple cross-sectional variance decomposition of log wages. The share of total log wage variance explained by (20) 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 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 (20) to US and French data.

The auxiliary wage growth equation (24). Finally, results from the auxiliary wage growth equation (24) are reported in Figure 3, which plots the wage-experience profiles estimated from (24) both on real (solid line) and simulated (dashed line) data, and Table 7 which reports the autocovariance structure of unexplained wage growth based on (24).

< **Figure 3, Table 7 about here.** >

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 (20),

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 (24). 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 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. Figure 4 shows a plot of the estimated deterministic human capital-experience profile, g_t.

<div align="center">< Table 8 about here. ></div>

As a preliminary remark, we should point out that all values contained in Table 8 are very precisely estimated.

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.

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, yet the estimated difference (in variances) is less than 3 percent of the point estimates.

The stochastic component of individual productivity, ε_t. Parameters of the assumed monthly AR(1) process followed by ε_t are reported in rows 9 and 10 of Table 8. These parameters suggest that the dynamics of ε_t 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 $\eta^{12} \simeq 0.07$. Most of the year-to-year persistence in wage shocks thus seems to stem from the stochastic dynamics of the piece rate r_t rather than from the built-in persistence of individual productivity shocks.

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, 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:

$$w_{t+1} - w_t =$$

$$
\begin{cases}
h_{t+1} - h_t & \text{prob } 1 - \mu - \delta - \lambda_1 \overline{F}(q_t) \\[2ex]
h_{t+1} - h_t + \int_{q_t}^{q'} \left(1 + \frac{\lambda_1 \overline{F}(x)}{\rho + \delta + \mu} \right) dx & \text{density } \lambda_1 f(q'), \ q' \in (q_t, p_t) \\[2ex]
h_{t+1} - h_t + p' - p_t - \int_{p_t}^{p'} \left(1 + \frac{\lambda_1 \overline{F}(x)}{\rho + \delta + \mu} \right) dx + \int_{q_t}^{p_t} \left(1 + \frac{\lambda_1 \overline{F}(x)}{\rho + \delta + \mu} \right) dx & \text{density } \lambda_1 f(p'), \ p' > p_t \\[2ex]
[\text{unobserved}] & \text{prob } \mu + \delta.
\end{cases}
$$

Hence, conditional on staying employed between experience levels t and $t+1$, the following holds true (where $\widetilde{\lambda}_1 = \frac{\lambda_1}{1 - \mu - \delta}$ includes this latter conditioning):

$$
\mathbf{E}\left[w_{t+1} - w_t \mid p_t, q_t, t \right] = \underbrace{\mathbf{E}\left(h_{t+1} - h_t \mid t \right)}_{\text{human capital}}
$$

$$
\underbrace{-\widetilde{\lambda}_1 \int_{p_t}^{+\infty} f(x) \int_{p_t}^{x} \frac{\lambda_1 \overline{F}(z)}{\rho + \delta + \mu} dz \, dx + \widetilde{\lambda}_1 \overline{F}(p_t) \int_{q_t}^{p_t} \left(1 + \frac{\lambda_1 \overline{F}(x)}{\rho + \delta + \mu} \right) dx}_{\text{search, between-job spells}}
$$

$$
\underbrace{-\widetilde{\lambda}_1 \overline{F}(p_t) \int_{q_t}^{p_t} \left(1 + \frac{\lambda_1 \overline{F}(x)}{\rho + \delta + \mu} \right) dx + \widetilde{\lambda}_1 \int_{q_t}^{p_t} \overline{F}(x) \left(1 + \frac{\lambda_1 \overline{F}(x)}{\rho + \delta + \mu} \right) dx}_{\text{search, within-job spell}}. \tag{25}
$$

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):

$$
\mathbf{E}\left(w_{t+1} - w_t \mid t \right) = \int_{p_{\min}}^{+\infty} \int_{p_{\min}}^{p} \mathbf{E}\left[w_{t+1} - w_t \mid p, q, t \right] dG(q \mid p, t) \, dL(p \mid t). \tag{26}
$$

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.[24]

---

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

Results. The decomposition of monthly wage growth given by (25) and (26) 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, $\mathbf{E}\left(w_{t+1} - w_t \mid t\right)$, 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 (25).

<center>< **Figure 5 about here.** ></center>

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, $F\left(\cdot\right)$—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 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

<center>33</center>

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.

<div align="center">< Figure 6 about here. ></div>

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 these models rather neatly, albeit we face some difficulties in fitting transitions between jobs.

The main analytical results of the paper relates to 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., F. Kramarz, P. Lengermann and S. Roux (2003),"Interindustry and Firm-Size Wage Differentials in the United States and France", Cornell University working paper.

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

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

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

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

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

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

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

[9] 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.

[10] 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.

[11] 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.

[12] Buhai, S., M. Portela, C. Teulings and A. van Vuuren (2006), "Returns to Tenure: Time or Rank?", paper presented at SOLE 2006.

[13] Bunzel, H., B. J. Christensen, N. M. Kiefer and L. Korsholm (2000), "Equilibrium Search with Human Capital Accumulation", *in* H. Bunzel, B. J. Chistensen, N. M. Kiefer and D. T. Mortensen, editors, *Panel Data and Structural Labour Market Models*, North-Holland Publishing.

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

[15] Dustmann, C. and C. Meghir (2005), "Wages, Experience and Seniority", *Review of Economic Studies*, 72, 77-108.

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

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

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

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

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

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

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

[23] 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.

[24] Rosholm, M. and M. Svarer (2004) "Endogenous Wage Dispersion in a Search-Matching Model", *Labour Economics*, 11, 623-45.

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

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

APPENDIX

A Derivation of steady-state distributions

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

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

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

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

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

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

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

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

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

Jointly solving those three equations, one obtains:

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

The corresponding cdf is obtained by summation:

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

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

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

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

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

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

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

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

This last equation solves as:

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

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

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

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

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

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

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

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

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

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

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

B   Details of the simulation procedure

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

We 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 $t \le 21 \times 12 = 252$ periods[25]. We begin with a sample of N workers for which we draw individual (log) heterogeneity parameters α from $\mathcal{N}\left(0, \sigma^2\right)$. 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 $t \le 252$, 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 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 $\mathcal{N}\left(0, \sigma_u^2/(1-\eta^2)\right)$.

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 (20) impossible. To get round this problem, we discretize $F(\cdot)$ in the following way. We take a fixed number J of active firms (e.g. the number of firm observations in the actual sample), give each of them a rank $j = 1, \cdots, J$ and assign corresponding productivity levels of $p_j = F^{-1}(j/J)$. Next, to assign the p_j's to workers (conditional on experience), we draw in the usual way a N-vector $z = (z_1, ..., z_N)$ of realizations of $\mathcal{U}[0, 1]$, and determine worker i's firm type as $p_{ij(i,t)} = \arg\min_{x \in \{p_1, ..., p_J\}} |L(x|t_i) - z_i|$. Similarly, worker i's q is assigned (conditional on $p = p_j$ and $t > 1$) as $q_{it} = q_{it}(p_j) = \arg\min_{x \in \{p_1, ..., p_{j-1}\}} |G(x|p_j, t_i) - y_i|$, where y_i is a draw from $\mathcal{U}[0, 1]$.

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

---

[25]The 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.

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 employer[26] (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 does not change firms. However we need to update the worker's productivity threshold $q(\cdot)$ to p', and also increment the worker's seniority and experience. Finally, workers who leave the sample (probability μ are automatically (i.e. deterministically) replaced by newborn unemployed workers with zero experience and new values of α drawn from $H(\cdot) \equiv \mathcal{N}\left(0, \sigma^2\right)$.

The simulated data sets, which have monthly wage observations (computed using (11) and the information recorded for each worker), are remodeled to replicate the structure of the actual data set (which only has yearly wage observations for the active job spell at the end of November—see section 3). Also, worker the simulated individual labor market histories are trimmed at the minimal experience level of 5 years.

---

[26]With 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.

	Ed. 9-11	Ed. 12	Ed. 13-18
Number of spells	7,068	12,334	2,396
U-spells (in months):			
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 (*S.D.*) noncens. res. duration./mts.	7.69 (*9.03*)	7.35 (*7.76*)	9.87 (*13.37*)
J-spells (in months):			
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 (*S.D.*) noncens. res. duration/mts.	32.27 (*26.01*)	31.73 (*26.21*)	32.71 (*25.43*)
Cross-Section Log-Wages:			
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

	# Workers	Number of employers per worker						Average
		1	2	3	4	5	6+	
Ed. 9-11	13,310	7,442	3,585	1,444	594	175	70	1.70
Ed. 12	23934	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

	# Workers	Number of observations per worker						Average
		1	2	3	4	5-10	11-14	
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

	# Firms	Observed firm size						Average
		2-5	6-10	11-20	21-50	51-100	100+	
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

	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

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

Table 4: Monte-Carlo study of the estimation procedure

	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 wage posting model

	Ed. 9-11		Ed. 12		Ed. 13-18	
	Real	Sim.	Real	Sim.	Real	Sim.
Firm effects						
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
Worker effects						
Std. dev.	0.1186	0.1265	0.1251	0.1244	0.1699	0.1576
Residual						
Std. dev.	0.1120	0.1167	0.1170	0.1164	0.1159	0.1234
Variance decomposition						
$V(w)$	0.0453		0.0486		0.0700	
$V(X\zeta)$	0.0012		0.0023		0.0095	
$V(\phi)$	0.0192		0.0193		0.0230	
$V(\psi)$	0.0120		0.0133		0.0253	
$\mathrm{Cov}\,(X\zeta,\phi) \times 10^4$	0.5112		-1.2593		-7.7588	
$\mathrm{Corr}\,(X\zeta,\phi)$	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 (20): moments of heterogeneity distributions

	Ed. 9-11		Ed. 12		Ed. 13-18	
	Real	Sim.	Real	Sim.	Real	Sim.
Autocovariances						
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 (24): autocovariance structure of residuals

	Ed. 9-11	Ed. 12	Ed. 13-18
Transition parameters			
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)
Firm productivity, $F(p) = 1 - e^{-[\nu(p-b)]^\omega}$			
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)
Worker productivity, $H(\alpha) = \mathcal{N}(0, \sigma^2)$			
σ	0.0690 (0.0032)	0.0807 (0.0014)	0.0737 (0.0053)
Idiosyncratic shocks, $\varepsilon_t = \eta\varepsilon_{t-1} + u_t, u_t \sim \mathcal{N}(0, \sigma_u^2)$			
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)
Experience, $g_t = \sum_{k=1}^3 \gamma_k(t - \tau_k)\mathbf{1}(t > \tau_k)$			
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

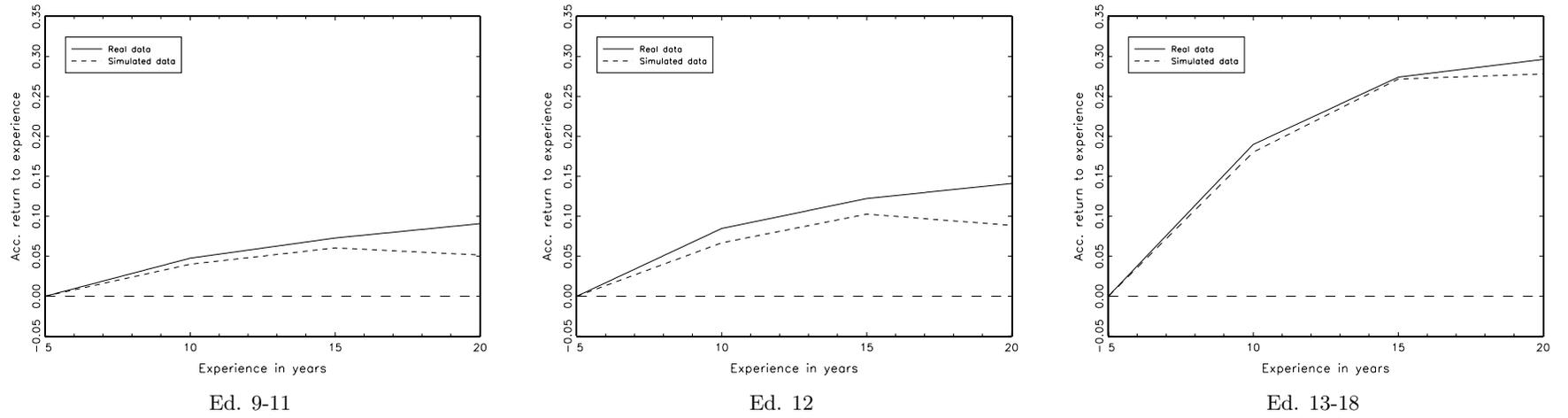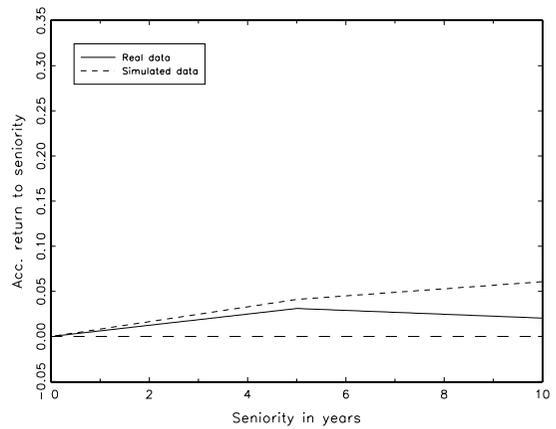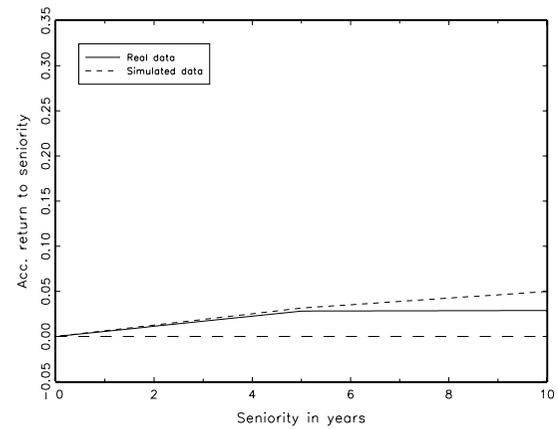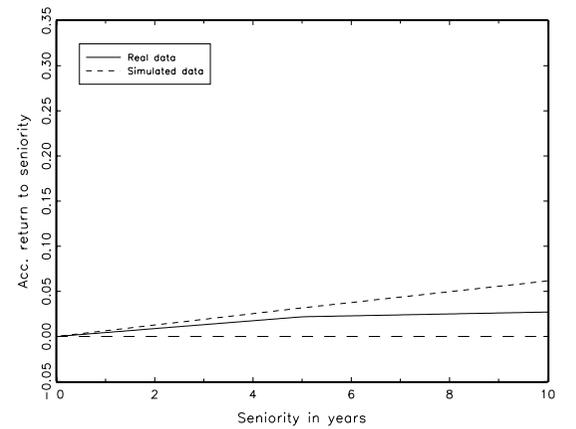Figure 1: Model fit—Mincer equation experience profiles

Ed. 9-11   Ed. 12   Ed. 13-18
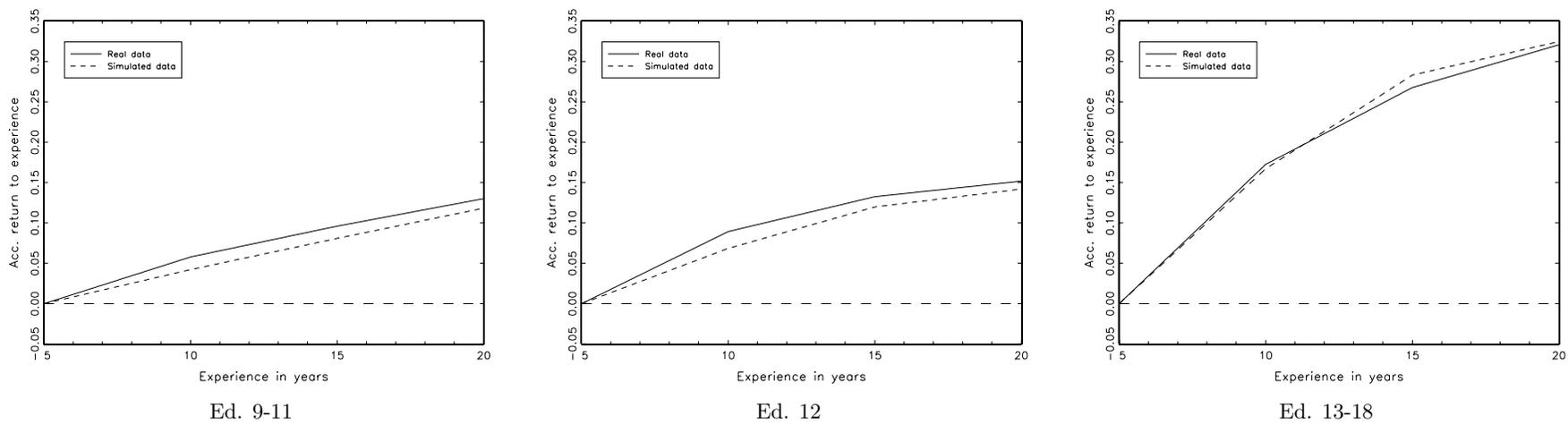
Figure 2: Model fit—Mincer equation tenure profiles
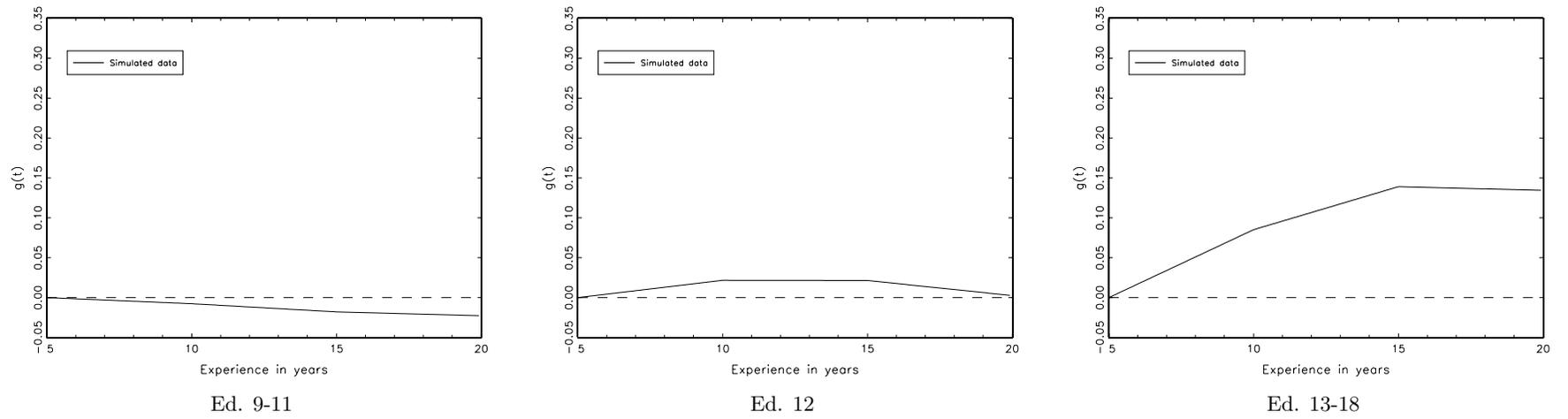
Figure 3: Model fit—Wage growth equation experience profiles

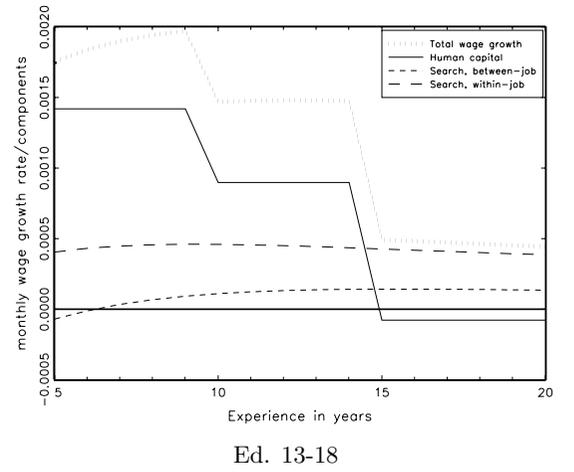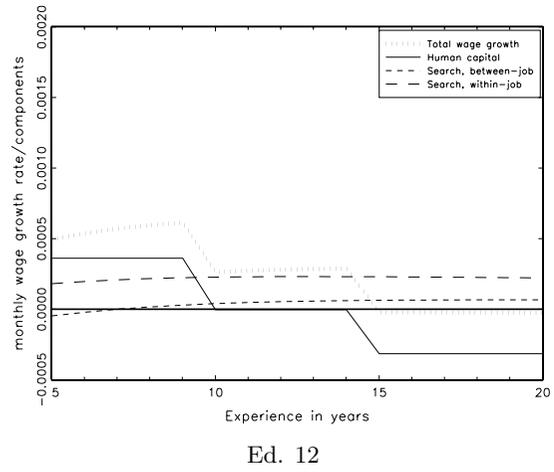Figure 4: Structural human capital-experience profile (g_t)

Ed. 9-11   Ed. 12   Ed. 13-18
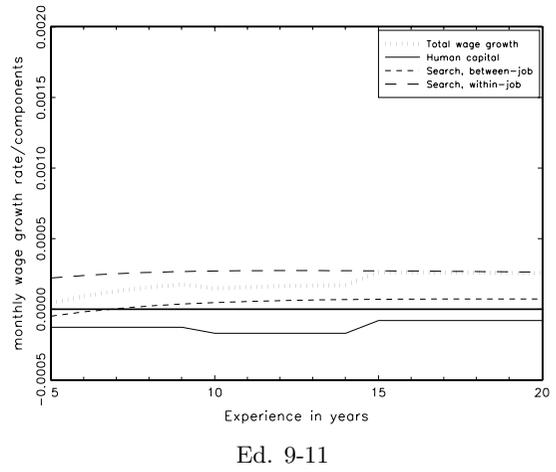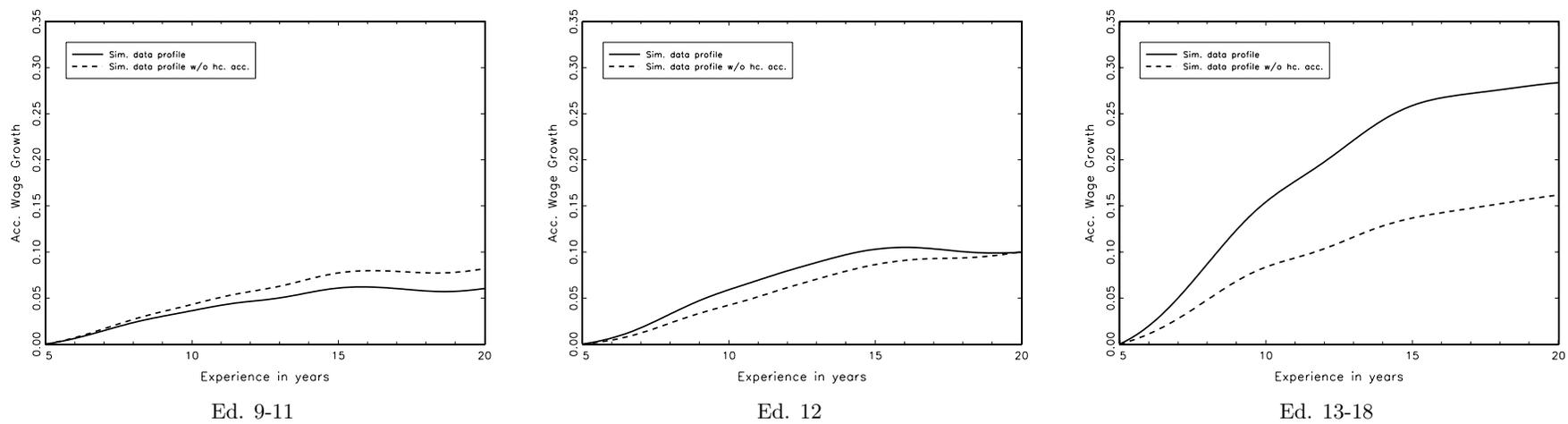
Figure 5: Decomposition of monthly wage growth

Figure 6: Countefactual wage-experience profiles