Tenure, Experience, Human Capital and Wages: A Tractable Equilibrium Search Model of Wage Dynamics*

Jesper Bagger†
Royal Holloway College
University of London

François Fontaine‡
BETA - CNRS
Université Nancy 2

Fabien Postel-Vinay§
University of Bristol
Sciences Po, Paris

Jean-Marc Robin¶
Sciences Po, Paris
University College London

August 2011

*Earlier versions of this paper were circulated since 2006 under the title "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation". This paper took a long time to finalize partly due to the difficulty in constructing Matched Employer-Employee data including firm-level productivity data, an essential ingredient of this project. We are grateful to Joe Altonji, Henning Bunzel, Ken Burdett and Melvyn Coles, as well as participants in numerous seminars and conferences for useful comments at various stages of this project. The usual disclaimer applies.

†E-mail: `jesper.bagger@rhul.ac.uk`. Bagger gratefully acknowledges financial support from LMDG and CAP, Aarhus University (LMDG is a Dale T. Mortensen Visiting Niels Bohr professorship project funded by the Danish National Research Foundation; CAP is a research unit funded by the Danish Council for Independent Research).

‡E-mail: `francois.fontn@gmail.com`. Fontaine gratefully acknowledges financial support from LMDG and CAP, Aarhus University, and from the PRES-Université de Lorraine.

§Address: Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, United Kingdom. E-mail: `Fabien.Postel-Vinay@bristol.ac.uk`. Postel-Vinay gratefully acknowledges financial support from the ESRC under grant No. RES-063-27-0090.

¶E-mail: `jmarc.robin@gmail.com`. Robin gratefully acknowledges financial support from the Economic and Social Research Council through the ESRC Centre for Microdata Methods and Practice grant RES-589-28-0001, and from the European Research Council (ERC) grant ERC-2010-AdG-269693-WASP.

Abstract

We develop and estimate an equilibrium job search model of worker careers, allowing for human capital accumulation, employer heterogeneity and individual-level shocks. Monthly wage growth is decomposed into the contributions of human capital and job search, within and between jobs. Human capital accumulation is found to be the most important source of wage growth in early phases of workers' careers, but is soon surpassed by search-induced wage growth. Conventional measures of the returns to tenure hide substantial heterogeneity between different workers in the same firm and between similar workers in different firms.

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

JEL codes: J24, J31, J41, J62

1 Introduction

Our main objective in this paper is to quantify the relative importance of human capital accumulation and imperfect labor market competition in shaping individual labor earnings profiles over the life cycle. We contribute to the empirical literature on wage equations along three broad dimensions.

The first one relates to Mincer's (1974) original specification of log-earnings as a function of individual schooling and experience. In their review of Mincer's stylized facts about post-schooling wage growth in the U.S., Rubinstein and Weiss (2006) list human capital accumulation and job search as two of the main driving forces of observed earnings/experience profile.[1] As these authors note, the obvious differences between those two theories in terms of policy implications (concerning schooling and training on one hand and labor market mobility on the other) motivates a thorough quantitative assessment of their relative importance. Rubinstein's and Weiss's detailed review of the available U.S. evidence lends support to both lines of explanation, thus calling for the construction of a unified model. This paper offers such a model.

Existing combinations of job search and human capital accumulation include the models of Bunzel, Christensen, Kiefer, and Korsholm (1999), Rubinstein and Weiss (2006), Barlevy (2008), Burdett, Carrillo-Tudela, and Coles (2009) and Yamaguchi (2010). None of these models simultaneously allow for worker and firm heterogeneity, idiosyncratic productivity shocks and human capital accumulation. Furthermore, none use Matched Employer-Employee (MEE) data on both firm output and worker wages, which are required to make sure that inference on rent sharing mechanisms does not rely solely on the model's structure.

Introducing individual shocks into a sequential job search model with a wage setting mechanism that is both theoretically and descriptively appealing turns out to be a difficult undertaking, tractable only in special cases (see Postel-Vinay and Turon, 2010, Moscarini and Postel-Vinay, 2010, Robin, 2011). Barlevy (2008) chooses to sacrifice theoretical generality for

[1] Rubinstein and Weiss also point to learning about job, worker or match quality as a third potential determinant of life-cycle earnings profiles. Learning is formally absent from our structural model. It is difficult to tease out of wage data what is due to learning about unobserved productivity characteristics from true productivity dynamics.

a realistic process of individual productivity shocks. He restricts the set of available wage contracts to piece-rate contracts, stipulating what share of output is received by the worker in lieu of wage. In this paper, we follow Barlevy's lead and assume piece-rate contracts. However, our model and empirical analysis differ from Barlevy's in two main dimensions.

First, we use MEE data and put strong emphasis on both firm heterogeneity and individual productivity shocks, whereas Barlevy uses NLSY data and thus cannot separate between different sources of heterogeneity. Second, he follows the Burdett and Mortensen (1998) tradition and assumes that each firm posts a unique and constant piece rate.[2] We instead assume that piece rates are renegotiated as workers receive outside offers as in Postel-Vinay and Robin (2002), and the extensions in Dey and Flinn (2005) and Cahuc, Postel-Vinay, and Robin (2006). It is now understood that wage posting fails to describe the empirical relationship between wages and productivity because the relative mildness of between-employer competition toward the top of the productivity distribution inherent to wage posting models implies that those models require unrealistically long right tails for productivity distributions in order to match the long right tails of wage distributions (Mortensen, 2005, Bontemps, Robin, and van den Berg, 2000). By allowing firms to counter outside offers, the sequential auction framework of Postel-Vinay and Robin (2002) intensifies firm competition and yields a wage equation that fits well the empirical relationship between observed firm output and wages (see Cahuc, Postel-Vinay, and Robin, 2006 and the results therein).[3]

Our second contribution is to inform the debate on the effect of job tenure versus that of experience on wage growth. The available empirical evidence on that important question is mixed. Some papers find large and significant tenure effects while others estimate them small or insignif-

[2]He does not endogenize the distribution of piece rates. Burdett, Carrillo-Tudela, and Coles (2009) work out the full equilibrium version of the model but do not estimate it.

[3]Yamaguchi (2010) also uses a sequential auction framework augmented with bargaining. However, like Barlevy, he uses NLSY data to estimate his model. The lack of separate data on productivity and wages makes his bargaining power estimates depend on functional form assumptions. Another difference is that he allows for match-specific productivity shocks when we introduce a richer pattern of heterogeneity, with persistent worker-specific shocks to ability. Lastly, our model is considerably easier to simulate and estimate, thanks to the piece rate contract assumption.

icant (see Abraham and Farber, 1987, Altonji and Shakotko, 1987, Topel, 1991, Dustmann and Meghir, 2005, Beffy, Buchinsky, Fougère, Kamionka, and Kramarz, 2006, Buchinsky, Fougère, Kramarz, and Tchernis, 2010). This literature emphasizes the inconsistency of tenure effects estimated by OLS, owing to a composition bias: in a frictional labor market, jobs that are more productive in some unobserved way should both last longer and pay higher wages. Differences between papers then mainly come down to different choices of instruments. Those choices are based on sophisticated theoretical arguments which are often laid out without the help of a formal model, thus inevitably leaving scope for some loose ends in the reasoning. For example, with forward-looking agents, wage contracts should reflect expectations about firms' and workers' future outside options, which are not precisely defined outside of an equilibrium model. Moreover, estimation often relies on strong specification assumptions, such as Topel's assumed linearity of the relationship between log wages and match quality, on one side, and tenure and experience, on the other. Again, only formal theory can give us a handle on whether these assumptions are reasonable or not.

Search theory provides a powerful framework to understand why and how wages increase with firm tenure. Firms that face the basic moral hazard problem of workers being unable to commit not to accept attractive outside job offers have an incentive to backload wages in order to retain their workforce. Under full firm commitment (and with risk-averse workers), this backloading takes the form of wages increasing smoothly with tenure, as shown by Burdett and Coles (2003).[4] We instead follow Postel-Vinay and Robin (2002) and assume that firms do not commit over the indefinite future, but revisit the piece rate they pay a worker each time the worker receives an attractive outside offer, implying that the worker's piece rate also increases with tenure, albeit stochastically and in discrete steps, in response to competitors' attempts to poach the worker. The contract-posting model of Burdett and Coles has predictions that are very close (although not entirely identical) to ours. What makes us favor the offer-matching approach in this paper is mainly tractability and amenability to estimation: the Burdett and Coles model is very hard to solve in the presence of firm

[4]Burdett and Coles (2010) extend their earlier 2003 paper by allowing for human capital accumulation and piece rate contracts, as in Barlevy (2008) and here.

heterogeneity, whereas firm heterogeneity (a key feature of the data) is a natural ingredient of our model.

A related issue is whether we should explicitly distinguish between general and firm-specific human capital. In the empirical literature, firm-specific human capital is a somewhat elusive concept generally associated with positive returns to tenure. However, as pointed out by Lazear (2003), the truly firm-specific components of human capital[5] are unlikely to be as important as the general component. Lazear explains upward-sloping wage/tenure profiles and the occurrence of job-to-job mobility with wage cuts by an argument combining search frictions, firm heterogeneity, and multiple skills used in different combinations by different firms. However, multiple skills are not necessary to the argument. As already mentioned, a combination of search frictions and moral hazard explains upward-sloping wage/tenure profiles. Moreover, allowing for productive heterogeneity among firms makes voluntary job changes consistent with wage losses: if the poaching firm is sufficiently more productive than the incumbent one, the promise of higher future wages will lead the worker to accept a lower initial wage. For the sake of parsimony, we thus restrict our model to one single dimension of general human capital and test its capacity to replicate standard measures of tenure and experience effects.

The third body of empirical work related to the present paper is the voluminous 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 (see Hall and Mishkin, 1982, MaCurdy, 1982, Abowd and Card, 1989, Topel and Ward, 1992, Gottschalk and Moffitt, 2009, Browning, Ejrnaes, and Alvarez, 2010, Meghir and Pistaferri, 2004, Guiso, Pistaferri, and Schivardi, 2005). The overwhelming majority of papers in that literature focus solely on wages and are silent about how productivity shocks impact wages.[6] Our model offers a simple theoretical structure within which to think about the impact of firm productive hetero-

[5]Quoting Lazear: "knowing how to find the restrooms, learning who does what at the firm and to whom to go to get something done," etc.

[6]One notable exception is Guiso, Pistaferri, and Schivardi (2005), who take a reduced-form look at the extent to which firm-level shocks to value added are transmitted to wages in Italian MEE data.

geneity on within- and between-firm wage dynamics and the transmission of individual productivity shocks to wages.

Our model's main output is a structural wage equation similar to the standard 'Mincer-type' 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) individual productivity shocks that help explaining the frequent earnings cuts that we observe.[7] In addition, the model permits a decomposition of average monthly wage growth into the contributions of human capital accumulation and of job search, within and between jobs.

We estimate our structural model using indirect inference on separate MEE samples of Danish workers with low, medium and high levels of education, respectively. The model fit is good. The decomposition of individual wage growth reveals that human capital accumulation is the most important source of wage growth in early phases of workers' careers, but is soon overtaken by search-induced wage growth, with between-job wage growth dominating within-job wage growth. The decompositions are qualitatively similar for all levels of education. However, more educated workers have higher total wage growth and this reflects both higher human capital accumulation and higher returns to job search. We further find that conventional measures of returns to tenure (based on linear log-wage regressions) conceal substantial heterogeneity between different workers in the same firm and between similar workers in different firms. This heterogeneity arise because workers with different labor market histories differ in their ability to appropriate match surplus from a given employer, and because more productive employers can get away with offering lower starting wages (and higher subsequent wage growth) than less productive employers.

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 section 5 we discuss estimation results and in section 7 we analyze the decomposition of individual wage-experience profiles. Section 9 concludes.

[7] When we write wage, we mean annual earnings. Most data sets, and administrative data are no exception, generally do not distinguish between contractual wage and bonuses. A lot of the observed earnings cuts may in fact be cuts in bonuses.

2 The Model

We consider a labor market where a unit mass of workers face a continuum of firms producing a multi-purpose good, which they sell in a perfectly competitive market. Workers can either be unemployed or matched with a firm. Firms operate constant-return technologies and are modeled as a collection of job slots that can either be vacant and looking for a worker, or occupied and producing. Time is discrete and the economy is at a steady state.

2.1 Production and Timing of Events

Let t denote the number of periods that a worker has spent working since leaving school. Call it experience. Log-output per period, $y_t = \ln Y_t$, in a firm-worker match involving a worker with experience t is defined as

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

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

$$h_t = \alpha + g(t) + \varepsilon_t, \tag{2}$$

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 that only changes when the worker is employed. The latter shock is worker-specific, and we only restrict it to follow a first-order Markov process.[8] A useful benchmark may be to think of it as a linear $AR(1)$ process, possibly with a unit root.

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 unem-

[8] 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.

ployed workers do not accumulate experience, so that if a worker becomes unemployed at an experience level of $t-1$, her/his experience $t-1$ and productivity h_{t-1} stagnate for the duration of the ensuing spell of unemployment. In the first period of the next employment spell, experience increases to t and productivity changes to h_t.

At the end of the period any employed worker leaves the market for good with probability μ, or sees her/his match dissolved with probability δ, or receives an outside offer with probability λ_1 (with $\mu+\delta+\lambda_1 \leq 1$).[9] 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 $[p_{\min}, p_{\max}]$. We assume that unemployment is equivalent to employment in the least productive firm of type $p_{\min}$ (in a sense stated precisely below). The implication is that an unemployed worker accepts any job offer s/he receives.[10]

2.2 Wage Contracts

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 bargaining model of Dey and Flinn (2005) and Cahuc, Postel-Vinay, and Robin (2006). 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 $V(r, h_t, p)$, with experience t kept implicit in the state vector to

[9] Alternatively, one could define conditional probabilities $(\delta', \lambda_1') \in [0,1]^2$ and write the unconditional probabilities as $\delta = (1-\mu)\,\delta'$ and $\lambda_1 = (1-\mu)\,(1-\delta')\,\lambda_1'$.

[10] In an environment with search frictions, experience (i.e. human capital) accumulation and different arrival rates on- and off-the-job, the reservation strategy of an unemployed worker could depend on the worker's experience level. This complication would entail loss of analytical tractability. The assumption of a constant reservation productivity (equal to $p_{\min}$) is partly justified by empirical studies that concludes that the acceptance probability of an unemployed worker is close to unity (see e.g. van den Berg, 1990).

simplify the notation. 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 verification of that statement). Also note that a piece rate of $R = 1$ (or $r = 0$) allocates the entire (expected) match value to the worker and leaves the employer with zero expected profit from that particular match. Total match value thus equals $V(0, h_t, p)$.

As described earlier, the worker contacts a potential alternative employer with probability λ_1 at the end of the current period. The alternative employer's type p' is drawn from the sampling distribution $F(\cdot)$. The central assumption is that the incumbent and outside employers bargain over the worker's services, based on the information available at the end of the current period. In particular, the idiosyncratic shock ε_{t+1}, determining human capital h_{t+1} for period $t + 1$, is not known when the new contract is negotiated. The outcome of the bargain is such that the firm that values the worker most—i.e. the firm with higher productivity—eventually hires (or retains, as the case may be) the worker.

Suppose for the time being that the dominant firm is the poacher (that is, suppose $p' > p$). Then the poacher wins the bargain by offering a piece rate r' defined as the solution to the equation

$$\mathbb{E}_t V(r', h_{t+1}, p') = \mathbb{E}_t \left\{ V(0, h_{t+1}, p) + \beta \left[V(0, h_{t+1}, p') - V(0, h_{t+1}, p) \right] \right\}, \quad (3)$$

where $\mathbb{E}_t$ designates the expectation operator conditional on the available information at experience t—here ε_{t+1} in h_{t+1} is the only random variable to integrate out conditional on ε_t —and where $\beta \in [0, 1]$ is a fixed, exogenous parameter. The dominant firm p' thus attracts the worker by offering, in expected terms, the value of the match with the dominated type-p firm plus a share β of the additional rent brought about by the match with the type-p' firm. Dey and Flinn (2005) view (3) as the solution of a Nash bargaining problem where the value of the dominated firm's best offer, $\mathbb{E}_t V(0, h_{t+1}, p)$, serves as the worker's threat point when bargaining with the winning firm. Cahuc, Postel-Vinay, and Robin (2006) rationalize (3) as the equilibrium of a strategic bargaining game adapted from Rubinstein (1982). By analogy

to the generalized Nash bargaining solution, we refer to β as the worker's bargaining power.

If $p' \leq p$ (the poacher is less productive than the incumbent), then the situation is a priori symmetric in that the incumbent employer is able to profitably retain the worker by offering a piece rate r' such that

$$\mathbb{E}_t V(r', h_{t+1}, p) = \mathbb{E}_t \left\{ V(0, h_{t+1}, p') + \beta \left[V(0, h_{t+1}, p) - V(0, h_{t+1}, p') \right] \right\}.$$

Note, however, that p' may be so low that this would not even entail a wage (or a piece rate) increase from the initial r. Such is indeed the case whenever the poacher's type p' falls short of the threshold value $q(r, h_t, p)$, defined by the indifference condition

$$\begin{aligned} \mathbb{E}_t V(r, h_{t+1}, p) = \mathbb{E}_t \{ & V(0, h_{t+1}, q(r, h_t, p)) \\ & + \beta \left[V(0, h_{t+1}, p) - V(0, h_{t+1}, q(r, h_t, p)) \right] \}. \end{aligned} \tag{4}$$

If $p' < q(r, h_t, p)$, the worker simply discards the outside offer from p'.

The above rules dictate the way in which the piece rate of an employed worker is revised over time. Concerning unemployed workers, we consistently assume that workers are able to secure a share β of the expected match surplus. The piece rate r_0 obtained by an unemployed worker with experience level t thus solves

$$\mathbb{E}_t V(r_0, h_{t+1}, p) = V_0(h_t) + \beta \mathbb{E}_t \left[V(0, h_{t+1}, p) - V_0(h_t) \right]. \tag{5}$$

where $V_0(h_t)$ is the lifetime value of unemployment at experience t.[11]

2.3 Solving the Model

We assume that the workers' flow utility function is logarithmic and that they are unable to transfer wealth across dates. Let ρ denote the discount rate. The typical employed worker's value function $V(r, h_t, p)$ is then defined

[11] We can now formally state our assumption regarding the assumed equivalence between unemployment and employment (with full surplus extraction) at the least productive firm type $p_{\min}$: $V_0(h_t) = \mathbb{E}_t V(0, h_{t+1}, p_{\min})$. As also stated above this assumption is motivated partly to retain tractability and partly to accommodate the empirical fact that unemployed workers typically accept the first job offer they receive.

recursively as:

$$
\begin{aligned}
V(r,h_t,p) = w_t &+ \frac{\delta}{1+\rho}V_0(h_t) \\
&+ \frac{1}{1+\rho}\mathbb{E}_t\Bigg\{\left[1-\mu-\delta-\lambda_1\overline{F}(q(r,h_t,p))\right]V(r,h_{t+1},p) \\
&+ \lambda_1\int_p^{p_{\max}}\left[(1-\beta)V(0,h_{t+1},p)+\beta V(0,h_{t+1},x)\right]dF(x) \\
&+\lambda_1\int_{q(r,h_t,p)}^{p}\left[(1-\beta)V(0,h_{t+1},x)+\beta V(0,h_{t+1},p)\right]dF(x)\Bigg\}, \quad (6)
\end{aligned}
$$

where $\overline{F} = 1 - F$ (the survivor function), $w_t = r + p + \alpha + g(t) + \varepsilon_t$ and the threshold $q(\cdot)$ is defined in (4).

The worker's value is the sum of current-period utility flow w_t and next-period continuation value, discounted with factor $1/(1+\rho)$. The continuation value has the following components: with probability δ, the worker becomes unemployed, a state that s/he values at $V_0(h_t)$. With probability μ, the worker leaves the labor force permanently and receives a value of 0. With probability λ_1, the worker receives an outside job offer emanating from a type-x firm and one of three scenarios emerges: the poaching employer may be more productive than the worker's current type-p employer ($x \geq p$), in which case the worker expects to come out of the bargain with value $\mathbb{E}_t\left[(1-\beta)V(0,h_{t+1},p)+\beta V(0,h_{t+1},x)\right]$. Alternatively, the poaching employer may be less productive than p but still worth using as leverage in the wage bargain ($p \geq x \geq q(r,h_t,p)$), in which case the worker expects to extract value $\mathbb{E}_t\left[(1-\beta)V(0,h_{t+1},x)+\beta V(0,h_{t+1},p)\right]$. Finally, the offer may not even be worth reporting ($x \leq q(r,h_t,p)$), in which case the worker stays with his/her initial contract with updated human capital, which has expected value $\mathbb{E}_t V(r,h_{t+1},p)$. Finally, with complementary probability $1-\delta-\mu-\lambda_1$, nothing happens and the worker carries on with her/his initial contract with updated human capital (expected value $\mathbb{E}_t V(r,h_{t+1},p)$).

Substitution of (6) into (4) produces the following implicit definition of

$q(\cdot)$ (see Appendix A for details):

$$r = -(1-\beta)\left[p - q(r, h_t, p)\right] - \int_{q(r,h_t,p)}^{p} \frac{\lambda_1 (1-\beta)^2 \overline{F}(x)\, dx}{\rho + \delta + \mu + \lambda_1 \beta \overline{F}(x)}$$
$$- \int \int_{q(r,h_t,p)}^{q(r,h_{t+1},p)} \frac{(1-\mu-\delta)(1-\beta)}{\rho + \delta + \mu + \lambda_1 \beta \overline{F}(x)}\, dx\, dM(h_{t+1}|h_t) \quad (7)$$

where $M(\cdot \mid h_t)$ is the law of motion of h_t which, up to the deterministic drift $g(t)$, is 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.

Conveniently, equation (7) has a simple, deterministic (indeed constant), consistent solution $q(r, p)$ implicitly defined by:

$$r = -(1-\beta)\left[p - q(r, p)\right] - \int_{q(r,p)}^{p} \frac{\lambda_1 (1-\beta)^2 \overline{F}(x)\, dx}{\rho + \delta + \mu + \lambda_1 \beta \overline{F}(x)}. \quad (8)$$

Now even though (7) implies no direct dependence of $q(\cdot)$ on t and h_t, other, nondeterministic solutions to (7) may still exist. We will ignore the possibility of more sophisticated expectational mechanisms in this paper, and concentrate on the deterministic solution (8).

2.4 The Empirical Wage Process

Under the deterministic solution (8), the (log) wage w_{it} earned by worker i hired at a firm with productivity p_{it} at time t is thus defined as follows:

$$w_{it} = \beta p_{it} + (1-\beta) q_{it} + \alpha_i + g(t) + \varepsilon_{it} - \int_{q_{it}}^{p_{it}} \frac{\lambda_1 (1-\beta)^2 \overline{F}(x)\, dx}{\rho + \delta + \mu + \lambda_1 \beta \overline{F}(x)}, \quad (9)$$

where q_{it} is the type of the last firm from which worker i was able to extract the whole surplus in the bargaining game. This wage equation implies a decomposition of individual wages into five components: an experience effect $g(t)$, a worker fixed effect α_i, an individual transitory productivity shock ε_{it}, an employer random effect p_{it}, and a random effect q_{it} relating to the most recent wage bargain.

The joint process governing the dynamics of $(p_{i,t+1}, q_{i,t+1})$ can be characterized as follows. If the worker is employed at time t then with probability μ she retires and with (exclusive) probability δ she becomes unemployed, in

which cases the value of $(p_{i,t+1}, q_{i,t+1})$ is irrelevant and can be set as missing; otherwise the worker may draw an outside offer with probability λ_1. Hence, given (p_{it}, q_{it}), $(p_{i,t+1}, q_{i,t+1})$ is determined by one of the following four regimes:

$$(p_{i,t+1}, q_{i,t+1}) = \begin{cases} (\cdot, \cdot), & \text{with probability } \mu + \delta \\ (p_{it}, q), \forall q \in (q_{it}, p_{it}], & \text{with density } \lambda_1 f(q) \\ (p, p_{it}), \forall p > p_{it}, & \text{with density } \lambda_1 f(p) \\ (p_{it}, q_{it}), & \text{with probability} \\ & 1 - \mu - \delta - \lambda_1 \overline{F}(q_{it}) \end{cases} \tag{10}$$

If the worker is unemployed in period t then

$$(p_{i,t+1}, q_{i,t+1}) = (p, p_{\min}), \forall p > p_{\min}, \quad \text{with density } \lambda_0 f(p).$$

Our model conveys an interpretation of the estimates of tenure and experience effects of the literature. With standard worker data on wages such as PSID data, all three stochastic components p_{it}, q_{it} and ε_{it} are unobservable, and p_{it} and q_{it} are correlated with tenure. Good matches, matches with high-p_{it} firms, indeed last longer. However, tenure is independent of the worker effect α_i because there is no sorting on unobservables.[12] Tenure is not a causal variable, but it can be used as a proxy for q_{it} because a longer tenure increases the chances of having drawn outside offers. It is likely to be very difficult to instrument, however, as one needs to correct for both the measurement error problem and the correlation with the unobservable firm effect p_{it}. Assuming that one could, if one believes in our model, current tenure "determines" within-firm wage dynamics because employers are forced to increase wages to retain their employees when they are

[12]The steady-state cross-sectional distribution of the pair (p_{it}, q_{it}) is derived in Appendix A.The other random components of wages appearing in (9) are exogenously distributed (α_i is just a fixed effect and ε_{it} follows an exogenous process of its own), and they are uncorrelated with p_{it} or q_{it}. In other words, the set of assumptions we have adopted implies that there is no assortative assignment of workers to firms based on those unobserved worker characteristics. As will become clear shortly, though, there will be assortative assignment based on experience. Characterization of this steady-state distribution is useful to simulate the model (see below in section 4 and Appendix C).

poached by competitors. Past tenure "determines" starting wages because any successful poacher had to compete with an incumbent employer to hire a worker. The next subsection develops an alternative way of empirically disentangling those two effects.

2.5 Wage Growth Decomposition

Making use of the wage equation (9) and the characterization of wage dynamics in (10), period-to-period wage growth $\Delta w_{i,t+1} = w_{i,t+1} - w_{it}$ goes as follows, for each one of the four regimes of equation (10),

$$\Delta w_{i,t+1} = \begin{cases} \text{missing}, \\ \Delta h_{i,t+1} + (1-\beta)(q - q_{it}) + \displaystyle\int_{q_{it}}^{q} \frac{\lambda_1 (1-\beta)^2 \overline{F}(x) dx}{\rho + \mu + \delta + \beta \lambda_1 \overline{F}(x)}, \\ \Delta h_{i,t+1} + (1-\beta)(p_{it} - q_{it}) + \beta(p - p_{it}) \\ \quad - \displaystyle\int_{p_{it}}^{p} \frac{\lambda_1 (1-\beta)^2 \overline{F}(x) dx}{\rho + \mu + \delta + \beta \lambda_1 \overline{F}(x)} + \int_{q_{it}}^{p_{it}} \frac{\lambda_1 (1-\beta)^2 \overline{F}(x) dx}{\rho + \mu + \delta + \beta \lambda_1 \overline{F}(x)}, \\ \Delta h_{i,t+1}. \end{cases}$$

Hence, conditional on staying employed between experience levels t and $t+1$, expected wage growth $\Delta w_{i,t+1}$ given p_{it}, q_{it} and experience t is the sum of three components:

- Human capital accumulation:

$$\mathbb{E}(\Delta h_{i,t+1} \mid t, p_{it}, q_{it}) = g(t+1) - g(t), \tag{11}$$

  which is a deterministic function of experience.

- Within-job wage mobility (always upward):

$$\frac{\lambda_1}{1-\mu-\delta} \int_{q_{it}}^{p_{it}} f(q) \Bigg[(1-\beta)(q - q_{it}) + \int_{q_{it}}^{q} \frac{\lambda_1 (1-\beta)^2 \overline{F}(x) dx}{\rho + \mu + \delta + \lambda_1 \beta \overline{F}(x)} \Bigg] dq. \tag{12}$$

This is the return to tenure reflected in the wage increases that employers have to grant workers to retain them in presence of poaching.

- Between-job wage mobility:

$$\frac{\lambda_1}{1-\mu-\delta}\int_{p_{it}}^{p_{\max}} f(p)\left[\beta(p-p_{it}) - \int_{p_{it}}^{p} \frac{\lambda_1(1-\beta)^2\overline{F}(x)dx}{\rho+\mu+\delta+\lambda_1\beta\overline{F}(x)}\right]dp$$
$$+\frac{\lambda_1\overline{F}(p_{it})}{1-\mu-\delta}\left[(1-\beta)(p_{it}-q_{it}) + \int_{q_{it}}^{p_{it}} \frac{\lambda_1(1-\beta)^2\overline{F}(x)dx}{\rho+\mu+\delta+\lambda_1\beta\overline{F}(x)}\right]. \quad (13)$$

This is the instantaneous return to job mobility. The negative component reflects the fact that workers are willing to give up a share of the surplus now in exchange for higher future earnings when moving from a less to a more productive employer.

Finally, the conditioning variables q_{it} and p_{it} can be integrated out using the conditional distributions derived in Appendix A. 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 and two terms reflecting the impact of interfirm competition for workers, both within and between job spells. Search frictions and sequential auctions generate wage/tenure profiles independently of human capital accumulation as employers raise wages in response to outside job offers as workers receive them.

3 Data

We estimate our model using a comprehensive Danish MEE panel covering the period 1985-2003. The backbone of this data is a panel of individual labor market histories recorded on a weekly basis (the spell data). The spell data combines information from a range of public administrative registers and effectively covers the entire Danish labor force in 1985-2003. Spells are initially categorized into one of five labor market states: employment, self-employment, unemployment, nonparticipation and retirement.[13] Start

[13]Nonparticipation is a residual state which in addition to out-of-the-labor-force spells captures imperfect take-up rates of public transfers, reception of transfers not used to construct the spell data and erroneous start and end dates.

and end dates of spells are measured in weeks and the unit of observation in the spell data is a given worker in a given spell in a given year. We identify employers (and hence jobs) at the firm-level.

We supplement the spell data with background information on workers, firms and jobs from IDA, an annual population-wide (age 15-70) Danish MEE panel constructed and maintained by Statistics Denmark from several administrative registers. IDA provides us with a measure of the average hourly wage for jobs that are active in the last week of November, and a worker's age, gender, education including graduation date from highest completed education, labor market experience and ownership code of the employing establishment.[14] The information on workers' labor market experience is central to this study and refers to the workers' actual (as opposed to potential) experience at the end of a calendar year. Experience information is constructed from workers' mandatory pension payments ATP and dates back to January 1st, 1964.[15]

Finally, we use information on firms' financial statements (accounting data) collected by Statistics Denmark in annual surveys in 1999-2003.[16] The accounting data essentially contain the sampled firms' balance sheets, along with information on the number of worker hours used by the firm, from which we can compute value added. The survey covers approximately 9,000 firms which are selected based on the size of their November workforce (see Table 7 further below).

These three sources of information are linked via individual, firm and establishment identifiers. Even though the datasets are of large scale and complexity, matching rates are high which we take as a confirmation of the unique quality and reliability of our data. On average, a last-week-of-November cross section contains 3.6 million workers and 130,000 firms (of which, on average, 8,700 have accounting data information in 1999-2003).

To weed out invalid or inconsistent observations, reduce unmodeled heterogeneity and to select a population for which our model can be taken as

[14]Ownership allows us to identify private sector establishments.

[15]ATP is a mandatory pension scheme for all salaried workers aged 16-66 who work more than eight hours per week that was introduced in 1964. ATP-savings are optional for the self-employed. ATP effectively covers the entire Danish labor force.

[16]The survey was initiated in 1995 for a few industries and was gradually expanded until its 1999 coverage included most industries with a few exceptions such as agriculture, public services and parts of the financial sector (source: Statistics Denmark).

Table 1: Summary statistics on Master Panels

	Ed. 7-11	Ed. 12-14	Ed. 15-20
N observations	2,534,203	4,344,288	663,362
N workers	168,649	320,638	66,155
N firms	66,787	113,813	24,792
N firms w/ accounting data	9,874	16,361	6,570
N employment spells	405,171	958,676	142,194
N unemployment spells	536,722	502,418	59,423
N nonparticipation spells	475,814	443,458	57,378

Source: Matched Employer-Employee data obtained from Statistics Denmark.

a reasonable approximation to actual labor market behavior we impose a number of sample selection criteria on the data (see Appendix B for details). We try to steer clear of labor supply issues by focusing on males that are at least two years past graduation. Then, we discard workers born before January 1st, 1948, as those workers may have accumulated experience prior to the period for which we can measure experience (from 1964 onwards, see footnote 14). The maximum age in the data thus increases from 37 in 1980 to 55 in 2003. Conveniently, this also makes our sample immune from retirement-related issues. We further combine the five labor market states listed above into three (employment, unemployment and nonparticipation) by truncating individual labor market histories at entry into retirement, self-employment, the public sector, or any industry for which we lack firm-level value added data. Finally, we stratify the sample into three levels of schooling, based on the number of years spent in education: 7-11 years (completion of primary school or equivalent), 12-14 years (completion of high school or equivalent, or completion of vocational education) and 15-20 years (completion of at least bachelor level or equivalent). We will refer to these 19 year long (unbalanced) panels of labor market histories as the "Master Panels". Table 1 provides summary statistics.

4 Estimation

We estimate the structural model by indirect inference (Gouriéroux, Monfort, and Renault, 1993). The principle of indirect inference is to find values

of the structural parameters that minimize the distance between a chosen set of moments of the real data and the same moments calculated on artificial data obtained by simulation of the structural model. The set of moments that are matched in this fashion are either standard raw data moments or parameters of auxiliary models, simpler to estimate than the original structural model.

We specify the model parametrically as follows. The sampling distribution of log firm types p is assumed to be three-parameter Weibull. The distribution of worker types α is zero-mean normal. The individual specific productivity shocks ε_t are assumed to follow an $AR(1)$ process with zero-mean normal innovations. Finally, the experience accumulation function $g(t)$ is assumed to be piecewise linear with knots at 120 and 240 months. The parameters thus introduced, and the transition parameters λ_0, δ and λ_1, workers' bargaining power β as well as a few additional parameters that are more conveniently introduced at a later stage, constitutes the structural parameter vector to be estimated. We do not estimate the discount rate ρ or the attrition rate μ, which we instead calibrate prior to estimation. Details of the algorithm used to simulate the structural model are given in Appendix C.

The choice of an auxiliary model is a key and sometimes controversial step in the indirect inference approach. Our selection of auxiliary models partly reflects the link between our structural analysis and the empirical labor literature on wage equations. Specifically, we combine the following three sets of moments:

1. Transition rates between the various labor market states. The theoretical counterparts of those transition rates can be obtained as functions of the structural probabilities δ, λ_0 and λ_1 only. We therefore estimate these structural parameters in a separate first step by matching those empirical transition probabilities, and conduct the rest of the estimation procedure conditional on those first-step estimates (see Appendix D for details).

2. A Mincer wage equation with worker and firm fixed-effects and a first-differenced version of the Mincer equation as a model of within-job wage dynamics. Because these auxiliary models are fairly standard

reduced-forms used for the analysis of labor market transitions and earnings dispersion/dynamics, our indirect inference procedure has the additional benefit of explicitly linking our structural approach to well-known results from the reduced-form literature.

3. Moments of the distribution of firm-level average value-added per worker, as observed in the accounting data. Those afford a direct measure of labor productivity and will therefore help assess the success of our model at capturing the link between wages and productivity.

We discuss these auxiliary models in detail in Appendix E.

5 Model Fit

We begin the analysis of estimation results with a look at the results pertaining to our three auxiliary models and our structural model's capacity to replicate those results. We will then turn to structural parameter estimates, and comment on what we can learn from the structural model about individual earnings dynamics within and between job spells.

5.1 Auxiliary Transition Probabilities

Table 2 reports estimates of the auxiliary transition probabilities obtained on the real data and those generated by the estimated structural model. As explained above, λ_0, δ and λ_1 are estimated in a separate first step using the empirical transition probabilities in Table 2 only. In this just-identified first step we obtain a perfect fit.

Estimates on the real data reflect average observed spell durations. Moreover, as noted above, the auxiliary UE and EU transition probabilities are in fact equal to the structural parameters λ_0 and δ while the EE transition probabilities are functions of λ_0, δ and λ_1. Based on the numbers in Table 2, the predicted average unemployment spell durations are 23.0, 17.1 and 26.5 months for the low, medium and high-education groups, respectively while the corresponding average employment spell durations are 4.1, 7.8 and 19.4 years with monthly probabilities of job-to-job transition being 0.0078, 0.0084 and 0.0074. Hence, while our data implies that highly

Table 2: Auxiliary monthly transition probabilities (simulated and real)

	Ed. 7-11		Ed. 12-14		Ed. 15-20	
	Sim.	Real	Sim.	Real	Sim.	Real
P^{UE}	0.0434	0.0434 (.0008)	0.0586	0.0586 (0.0009)	0.0377	0.0377 (0.0021)
P^{EU}	0.0203	0.0203 (.0002)	0.0107	0.0107 (0.0001)	0.0043	0.0043 (0.0001)
P^{EE}	0.0078	0.0078 (.0001)	0.0084	0.0084 (0.0001)	0.0074	0.0074 (0.0001)
P^{UE}/P^{EU}		2.14		5.48		8.77
P^{EE}/P^{EU}		0.38		0.79		1.72

Note: Standard errors in parentheses. Standard errors computed by bootstrapping the variance-covariance matrix of the real moments (10,000 replications).

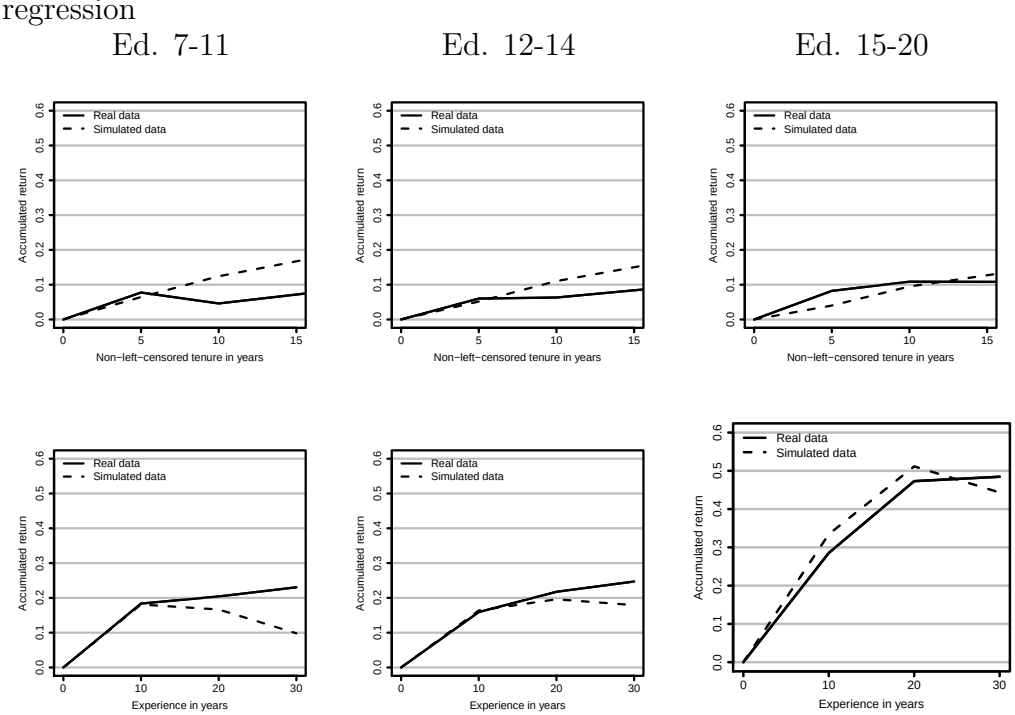
educated workers face longer unemployment spells on average, that group also face considerably lower risk of becoming unemployed. The ratios of unemployment duration to employment duration are 2.1, 5.5 and 8.8 for the low, medium and high-educated groups, respectively. The estimated transition probabilities are somewhat sensitive to the sample selection procedure described in Section 3 which explains both the relatively long unemployment durations and the ranking of unemployment durations between education groups.

5.2 Auxiliary Wage Regression

Figure 1 shows the experience and tenure profiles of individual wages as estimated from the Mincer-type auxiliary wage regression, equation (E1) of Appendix E. In Figure 1, 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 fixed effect distributions—again based on the auxiliary wage regression—are reported in Table 3.

We first review estimates based on the real data. The auxiliary wage regression indicates positive returns to experience in all three subsamples (second row in Figure 1). These are quantitatively rather modest for the low-educated group (who benefit from a 18 percent wage increase as they go from 0 to 10 years of experience, followed by 3 percent as they go from

Figure 1: Cumulative tenure and experience profiles from auxiliary wage regression



Note: Estimated and simulated left-censored tenure profiles are available on request.

10 to 20 years of experience and a further 3 percent as they go from 20 to 30 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 30 percent between 0 and 10 years of experience, and then by another 20 percent over the following 10 years of their careers, at which point wages only rises very modestly, if at all, with experience).

As can be seen from the first row in Figure 1, the auxiliary wage regression predicts moderate returns to tenure in all three subsamples.[17] Workers typically enjoy a 6-8 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. These profiles are correctly picked up by our structural model, albeit with a slight tendency to understate experience effects and overstate tenure effects at high tenure and experience levels. Inherent to the structure of all search models is the fact that arrival probability of outside job offers λ_1 is a common determinant of job to job transitions and wage profiles: in our case λ_1 governs the frequency of outside job offers which directly impacts the average frequency of wage increases. Now in the estimation, λ_1 is estimated separately in a first step from job transition data alone, while the auxiliary wage regression (as all other auxiliary models) is only used in the second, over-identified, step of our two-step procedure, conducted conditional on first-step estimates. As it turns out, fitting labor market transitions requires a λ_1 that tends to be slightly "too high" in the sense that it induces wage increases that are too frequent, and distorts the balance between tenure and experience effects. However, given the mildness of the discrepancy suggested by Figure 1, we conclude that the data lend support to the model's structure.

Next turning to Table 3, comparison of the firm and worker effect distributions across education groups hints at some degree of positive sorting 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

[17] A technical point: as explained in Appendix E, we include separate controls for left-censored and non-left-censored tenure in the auxiliary wage regression (E1) although we do not report or comment on the left-censored tenure profiles. The estimated left-censored wage profiles (and the structural model's fit to these profiles) are available on request.

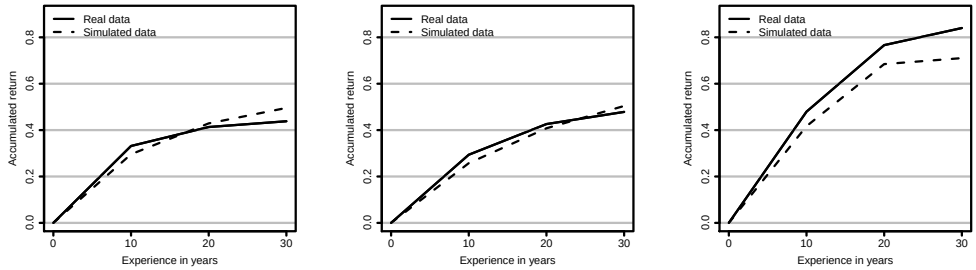
Table 3: Auxiliary wage regression (simulated and real)

	Ed. 7-11		Ed. 12-14		Ed. 15-20	
	Sim.	Real	Sim.	Real	Sim.	Real
Firm effects						
Mean	4.9524	4.9369 (0.0016)	5.0279	5.0221 (0.0013)	5.2066	5.2277 (0.0025)
Std. dev.	0.1640	0.1578 (0.0006)	0.1585	0.1478 (0.0003)	0.1740	0.1779 (0.0010)
Skewness	0.6990	−0.0361 (0.0151)	0.9365	0.3266 (0.0091)	0.7781	−0.2484 (0.0257)
Kurtosis	2.6721	4.9479 (0.0444)	2.9900	4.1128 (0.0270)	2.5599	5.2694 (0.0999)
Worker effects						
Std. dev.	0.1147	0.1180 (0.0005)	0.1283	0.1356 (0.0003)	0.1633	0.1665 (0.0009)
Skewness	0.0819	0.2357 (0.0234)	0.0771	0.6755 (0.0095)	−0.0124	0.4425 (0.0294)
Kurtosis	3.1890	7.2277 (0.1050)	3.3654	5.0941 (0.0362)	5.5706	5.8034 (0.1381)
Residuals						
Std. dev.	0.1237	0.1372 (0.0003)	0.1273	0.1354 (0.0001)	0.1434	0.1471 (0.0005)
Skewness	0.1197	−0.1235 (0.0096)	0.0264	0.1358 (0.0044)	−0.1212	−0.1754 (0.0304)
Kurtosis	3.9806	6.1422 (0.0529)	3.8340	5.0594 (0.0192)	4.5021	9.5946 (0.2985)
Within-job residual autocovariance						
Order 1	0.0003	0.0027 (0.00003)	0.0011	0.0028 (0.00002)	0.0027	0.0031 (0.0001)
Order 2	−0.0006	0.0002 (0.00002)	−0.0005	0.0004 (0.00001)	0.0003	0.0004 (0.00004)
Order 3	−0.0011	−0.0010 (0.00002)	−0.0010	−0.0007 (0.00001)	−0.0007	−0.0010 (0.00004)
Order 4	−0.0011	−0.0014 (0.00002)	−0.0011	−0.0012 (0.00001)	−0.0012	−0.0015 (0.00003)
Within-job residual autocorrelation						
Order 1	0.0231	0.1440 (0.0016)	0.0738	0.1516 (0.0009)	0.1302	0.1452 (0.0025)
Order 2	−0.0404	0.0098 (0.0012)	−0.0309	0.0227 (0.0007)	0.0134	0.0185 (0.0017)
Order 3	−0.0705	−0.0528 (0.0011)	−0.0628	−0.0409 (0.0006)	−0.0359	−0.0484 (0.0017)
Order 4	−0.0705	−0.0737 (0.0009)	−0.0679	−0.0663 (0.0005)	−0.0568	−0.0695 (0.0014)

Note: Standard errors in parentheses. Standard errors computed by bootstrapping the variance-covariance matrix of the real moments (10,000 replications). The estimated slope coefficients on left-censored tenure, non-left-censored tenure and experience are available on request. We report autocorrelations and autocovariances although the structural estimation was based on autocovariances only. Standard errors of autocorrelations computed using the delta method.

Figure 2: Cumulative experience profiles from auxiliary wage growth regression

Ed. 7-11 Ed. 12-14 Ed. 15-20



effect at zero in all samples.) Moreover, dispersion of worker- (and, to a smaller extent, firm-) effects tends to be slightly higher among more educated groups. Except the first order residual wage autcorrelations among workers with 7-11 years and 12-14 years of education, all numbers in Table 3 are accurately replicated by the structural model, especially so for the lower order moments. The failure to fit first order autocorrelations among low- and medium educated workers reflects a poor fit to the first order autocovariance. Bearing these caveats in mind, we conclude that our model captures wage dynamics as measured by the within-job residual autocorrelations in the wage regression (E1) reasonably well.

5.3 Auxiliary Wage Growth Equation

Results from the auxiliary wage growth equation, equation (E2) of Appendix E, are reported in Figure 2, which plots the wage-experience profiles estimated from that equation both on real (solid line) and simulated (dashed line) data, and Table 4 which reports the parameter estimates of the wage profiles, the moments of the residual distribution and the autocovariance structure of wage growth residuals.

The profiles in Figure 2 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, 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, although the structural model slightly underestimates the experience profiles in the high education group.

In Table 4, the second moment of the distribution of residual wage growth is well captured by the structural model (first moment normalized at zero). The model has difficulty fitting the higher order moments. Residual autocorrelations 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.

5.4 Auxiliary Value Added Equation

We finally turn to value added data. When thinking about the contribution of value added data to the identification of our structural model, it should be kept in mind that the relevant individual productivity parameter in the model is p, which is firm- and education-level specific. Yet we do not have a direct measure of the contribution to observed output of the different education groups in the data. Rather, we only observe firm-level productivity (as measured by hourly value added), which mixes all education levels. We circumvent this problem by positing a group-specific relationship between average output per worker and labor productivity p: see equation (E4) in Appendix E. A related problem when using the empirical distribution of firm-level value added per worker to identify our model is that we cannot plausibly assume that the underlying distribution of firm types (in the population of firms) is exactly identical to the sampling distribution faced by workers. Instead, we match employment-weighted firm specific moments. Again, see Appendix E for details.

Results from the auxiliary equation linking firm productivity and value added data are displayed in Table 5 which reports moments of the employment weighted distributions of log hourly value added, individual wages, as well as the standard deviation of within-job annual growth in log value added per FTE worker.

We start by considering the moments based on real data. Overall, our data exhibits a considerable amount of dispersion in average log labor pro-

Table 4: Auxiliary wage growth regression (simulated and real)

	Ed. 7-11		Ed. 12-14		Ed. 15-20	
	Sim.	Real	Sim.	Real	Sim.	Real
Residuals						
Std. dev.	0.1277	0.1345 (0.0003)	0.1221	0.1273 (0.0002)	0.1251	0.1354 (0.0008)
Skewness	0.2104	0.0557 (0.0178)	0.2685	0.0206 (0.0078)	0.4466	−0.1082 (0.0567)
Kurtosis	4.0641	9.2285 (0.1483)	3.8058	8.0027 (0.0662)	4.6556	24.4069 (1.2105)
Within-job residual autocovariance						
Order 1	−0.0051	−0.0045 (0.00005)	−0.0047	−0.0042 (0.00003)	−0.0048	−0.0037 (0.00008)
Order 2	−0.0004	−0.0008 (0.00003)	−0.0008	−0.0008 (0.00002)	−0.0009	−0.0008 (0.00004)
Order 3	−0.0005	−0.0006 (0.00003)	−0.0004	−0.0005 (0.00001)	−0.0004	−0.0006 (0.00004)
Order 4	−0.0005	−0.0006 (0.00002)	−0.0004	−0.0004 (0.00001)	−0.0004	−0.0005 (0.00003)
Within-job residual autocorrelation						
Order 1	−0.3109	−0.2485 (0.0022)	−0.3173	−0.2600 (0.0013)	−0.3062	−0.2026 (0.0042)
Order 2	−0.0242	−0.0463 (0.0016)	−0.0568	−0.0483 (0.0009)	−0.0604	−0.0438 (0.0021)
Order 3	−0.0297	−0.0319 (0.0014)	−0.0268	−0.0316 (0.0009)	−0.0273	−0.0312 (0.0021)
Order 4	−0.0306	−0.0321 (0.0012)	−0.0265	−0.0277 (0.0007)	−0.0253	−0.0284 (0.0017)

Note: Standard errors in parentheses. Standard errors computed by bootstrapping the variance-covariance matrix of the real moments (10,000 replications). The estimated slope coefficients on experience are available on request. We report autocorrelations and autocovariances although the structural estimation was based on autocovariances only. Standard errors of autocorrelations computed using the delta method.

Table 5: Auxiliary value added equation (simulated and real)

	Ed. 7-11		Ed. 12-14		Ed. 15-20	
	Sim.	Real	Sim.	Real	Sim.	Real
Employment weighted log value added per FTE worker (y)						
Mean	5.3289	5.3447 (0.0017)	5.3304	5.3461 (0.0010)	5.5630	5.5532 (0.0023)
Std. dev.	0.3104	0.3216 (0.0013)	0.3237	0.3478 (0.0008)	0.3679	0.4000 (0.0018)
Skewness	0.4636	0.3248 (0.0147)	0.6384	0.4511 (0.0085)	0.4887	0.3069 (0.0153)
Kurtosis	2.9949	3.8453 (0.0272)	3.1726	3.9026 (0.0176)	2.8219	3.8019 (0.0276)
Log wages, individual level (w)						
Mean	5.1199	5.1707 (0.0014)	5.2333	5.2809 (0.0008)	5.6572	5.6541 (0.0020)
Std. dev.	0.2488	0.2641 (0.0012)	0.2643	0.2574 (0.0006)	0.3483	0.3363 (0.0017)
Skewness	0.6611	0.0284 (0.0205)	0.7805	0.6240 (0.0063)	0.3936	0.3404 (0.0302)
Kurtosis	3.5093	4.1689 (0.0412)	3.7199	3.4026 (0.0148)	3.0508	4.1503 (0.1301)
Correlations and innovations						
$\text{Corr}(y, w)$	0.6125	0.2651 (0.0050)	0.5943	0.2631 (0.0028)	0.5743	0.2052 (0.0057)
Std. dev. within-job Δy	0.2055	0.2036 (0.0010)	0.2185	0.2124 (0.0006)	0.2510	0.2494 (0.0014)

Note: Standard errors in parentheses. Standard errors computed by bootstrapping the variance-covariance matrix of the real moments (10,000 replications). The estimation was based on $\text{Cov}(y, w)$. Standard errors of $\text{Corr}(y, w)$ computed using the delta method. Fit to $\text{Cov}(y, w)$ available on request.

ductivity. As one would expect based on the estimated wage regressions presented above, the education specific log wage distributions are also clearly ranked in terms of mean and dispersion with higher educated workers having higher average wages and higher dispersion as well. The fact that average log wages exceed average log value added among the high educated workers is an artifact of not observing the relevant productivity parameter p. Note that the simple relationship between structural labor productivity p and value added is sufficiently flexible to capture this feature of the data. The fit to the marginal distributions of log value added and wages is overall good, even for higher order moments. The fit to the standard deviation of within-job changes in log hourly value added is also good. This moment pins down the stochastic shock to the proposed relationship between structural labor productivity p and log hourly value added.

Finally, wages and value added are positively correlated. The structural model does captures the sign of the correlation but overestimates its magnitude considerably.

6 Structural Parameter Estimates

Estimates of the structural parameters are reported in Table 6. Recall that the monthly discount rate $\rho = 0.0050$ and attrition probability $\mu = 0.0018$ were both fixed prior to estimation.

6.1 Job Mobility

Parameters relating to labor market mobility (i.e. offer arrival and layoff probabilities) are reported in the top panel of Table 6. By construction the job offer arrival probability for unemployed workers λ_0 and the layoff probability δ differ from the estimated job finding and job destruction probabilities p^{UE} and p^{EU} (Table 2) only by simulation noise.[18] Parameter

[18]The estimated transition parameters implies that the steady state monthly job-to-job transition probabilities are 0.0079, 0.0084 and 0.0070 for the low, medium and high-educated groups, respectively. These probabilities differ slightly from the empirical job-to-job transition probabilities reported in Table 2 because the former are computed using the steady state distribution of firm productivity, not conditional on experience, while the latter are computed from a population which by construction ages over the sampling period (see section 3 for details).

Table 6: Structural parameter estimates

	Ed. 7-11	Ed. 12-14	Ed. 15-20
Parameters estimated in 1st step			
Job mobility			
λ_0 (monthly)	0.0435	0.0590	0.0380
	(0.0009)	(0.0008)	(0.0029)
δ (monthly)	0.0203	0.0107	0.0042
	(0.0002)	(0.0001)	(0.0001)
λ_1 (monthly)	0.0200	0.0265	0.0304
	(0.0005)	(0.0005)	(0.0011)
Parameters estimated in 2nd step (conditional on 1st step)			
Workers' bargaining power			
β	0.4141	0.3160	0.2475
	(0.0019)	(0.0009)	(0.0024)
Sampling distribution $F(p) = 1 - \exp\{-[\chi_1(p-\chi_0)]^{\chi_2}\}$			
$\chi_0 = p_{\min}$ (location)	4.7194	4.9230	5.1365
	(0.0038)	(0.0012)	(0.0029)
χ_1 (scale)	2.4373	4.1405	4.7878
	(0.0216)	(0.0283)	(0.0849)
χ_2 (shape)	1.4010	0.9324	0.8231
	(0.0131)	(0.0039)	(0.0086)
Worker type distribution $H(\alpha) = \mathcal{N}(0, \sigma_\alpha^2)$			
σ_α	0.0873	0.0940	0.1015
	(0.0006)	(0.0004)	(0.0016)
Productivity shocks $\varepsilon_{it} = \eta\varepsilon_{it-1} + u_{it}$ with $u_{it} \sim \mathcal{N}(0, \sigma_u^2)$			
η	−0.3973	0.5868	0.5986
	(0.0364)	(0.0062)	(0.0173)
σ_u	0.3217	0.1269	0.1252
	(0.0024)	(0.0014)	(0.0041)
Human capital $g(t) = \frac{1}{12}\sum_{k=1}^{3}\gamma_k(t-\tau_k)\mathbf{1}_{\{t>\tau_k\}}$			
γ_1 (knot $\tau_1 = 0$)	0.0143	0.0112	0.0239
	(0.0002)	(0.0001)	(0.0003)
γ_2 (knot $\tau_2 = 10 \times 12$ months)	−0.0161	−0.0088	−0.0100
	(0.0003)	(0.0002)	(0.0004)
γ_3 (knot $\tau_3 = 20 \times 12$ months)	−0.0058	−0.0046	−0.0230
	(0.0002)	(0.0002)	(0.0005)
Value added equation $y_{jt} = \kappa_0 + \kappa_1 p_j + z_{jt}, z_{jt} \sim \mathcal{N}(0, \sigma_p^2)$			
κ_0	0.3696	0.5737	0.6555
	(0.0351)	(0.0192)	(0.0435)
κ_1	0.9558	0.8907	0.8675
	(0.0069)	(0.0036)	(0.0078)
σ_v^2	0.1445	0.1543	0.1777
	(0.0007)	(0.0004)	(0.0010)

Note: Standard errors in parentheses. See Appendix D for details on the computation of the standard errors. The discount rate ρ is fixed at a monthly value of 0.0050 and the attrition rate μ at a monthly value of 0.0018.

estimates indicated that education protects from the risk of unemployment and increases the occurrence of job offers while employed.

6.2 Worker Bargaining Power

Our estimates of worker bargaining power β are reported in the third panel of Table 6. The parameter estimates indicates that a worker's bargaining power declines with education from a value of 0.41 for low educated workers down to 0.31 for medium educated workers and to 0.25 for high educated workers. These estimates differ substantially from those obtained by Cahuc, Postel-Vinay, and Robin (2006) who estimate a similar model with no experience accumulation on French data and find that workers in low skilled occupations have virtually no bargaining power with bargaining power increasing from less to more skilled occupations. No doubt the discrepancy is partly explained by our use of Danish (rather than French) data, stratification on education (rather than occupations), different model specification (in particular our inclusion of human capital accumulation), and different way of including firm level output data in the estimation.[19] Note, however, that a worker's steady state share of match output, i.e. the piece rate, depends on several structural parameters in addition to β, most notably the probability that the worker obtains an outside offer λ_1 (see (8)). Since λ_1 is higher for high educated workers, this may offset the lower β for this group. Indeed, the average steady state piece rates are 0.82, 0.82 and 0.79 for workers with low, medium and high education, respectively.[20] Hence, differences in on-the-job search behavior across education groups all but offset the differences in estimated bargaining power.

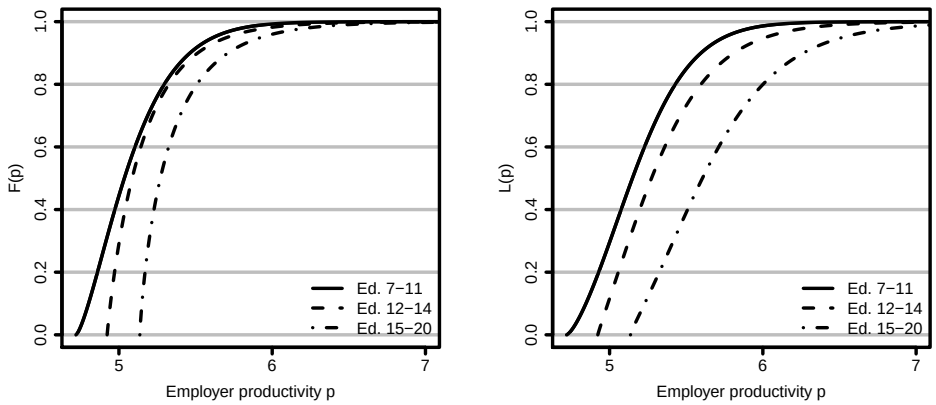
[19]The inclusion of experience accumulation prevents us from easily constructing labor productivity from structural firm-level production function estimates as is done by Cahuc, Postel-Vinay, and Robin (2006).

[20]The piece rate in a match between a worker with outside option q and employer of type p is $R(p,q) = e^{r(p,q)}$ where $r(p,q)$ is given by (8). Workers' average share of match output in cross section n of our data is thus

$$\int_b^\infty \int_b^p R(p,q) dG(q \mid p, t \leq T_n) dL(p \mid t \leq T_n)$$

where T_n is the maximum experience level in cross section n (see Appendix A for derivation of the various distributions). The reported average piece rates refer to the simple average across nineteen annual cross sections.

Figure 3: Sampling (left panel) and employer (right panel) distributions



6.3 The Sampling Distribution of Firm Productivity

Estimates of the parameters of $F(\cdot)$ are reported in the bottom panel of Table 6. Perhaps more directly informative are the implied mean and variances of the relating sampling distributions. The mean sampled productivity is 5.0933 for workers with 7-11 years of schooling, 5.1725 for workers with 12-14 years of schooling and 5.3685 for workers with 15-20 years of schooling (all in log terms). The corresponding standard deviations are 0.2704, 0.2677 and 0.2837. Finally, the lower support of $F(\cdot)$ is the parameter $p_{\min}$, which is directly available from Table 6.

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. This ranking extends to a ranking in terms of first-order stochastic dominance (see left panel of Figure 3). A similar plot in the right panel of Figure 3 of the corresponding cross-sectional distributions of employer types $L(p)$, which are deduced from the estimated sampling distributions $F(p)$ and transition parameters μ, δ and λ_1 using equation (A13) of Appendix A, shows that the same FOSD-ordering holds for these cross-sectional distributions, thus confirming the presence of positive sorting by education.

6.4 Worker Heterogeneity

The bottom panel five of Table 6 also contains the estimated standard deviation of the distribution of worker fixed, innate ability, α. These do not differ much between education groups although within-group dispersion in ability is increasing from low to high educated workers. Interestingly, the structural model estimates a much lower variance of the person-effect than the auxiliary Mincer equation. This is likely due to the fact that the person effect in the auxiliary equation captures the persistence generated by the $AR(1)$ idiosyncratic shock ε_{it}.

6.5 The Stochastic Component of Individual Productivity

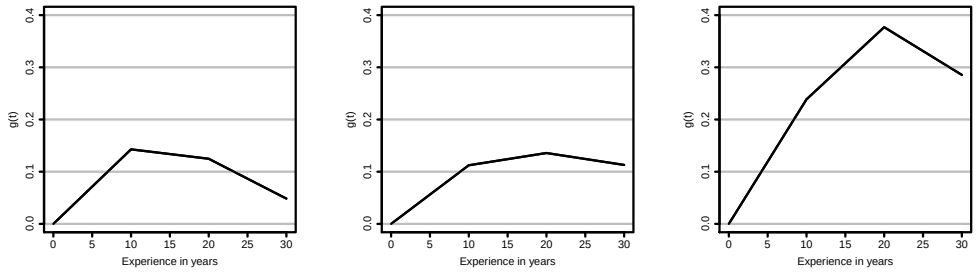
The first thing to notice about our estimates of the parameters of the assumed monthly $AR(1)$ process followed by ε_{it} (also in Table 6) is a clear tendency toward less dispersed innovations among more educated workers. The standard deviation of innovations is almost three times as high for workers with 7-11 years of education compared to that of workers with 15-20 years of education (or 12-14 years of education). In terms of persistence, low educated workers face a $AR(1)$ process with a negative autoregressive coefficient of -0.40 while medium and highly educated workers face positive autocorrelated $AR(1)$ processes with AR coefficients 0.59 and 0.60. We may characterize the risk faced by the different groups of workers by comparing the standard deviation of the stationary distributions of ε_{it}. These are, respectively, 0.35, 0.16 and 0.16 for the low, medium and high education groups. Low educated workers face the most risk. Note that the reported AR coefficients are based on a period length of one month, and translate into much smaller annual coefficients of -0.0509, 0.0361 and 0.0369 for the low, medium and high-education groups, respectively.[21]

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

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

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

Ed. 7-11 Ed. 12-14 Ed. 15-20



6.6 Human Capital Accumulation

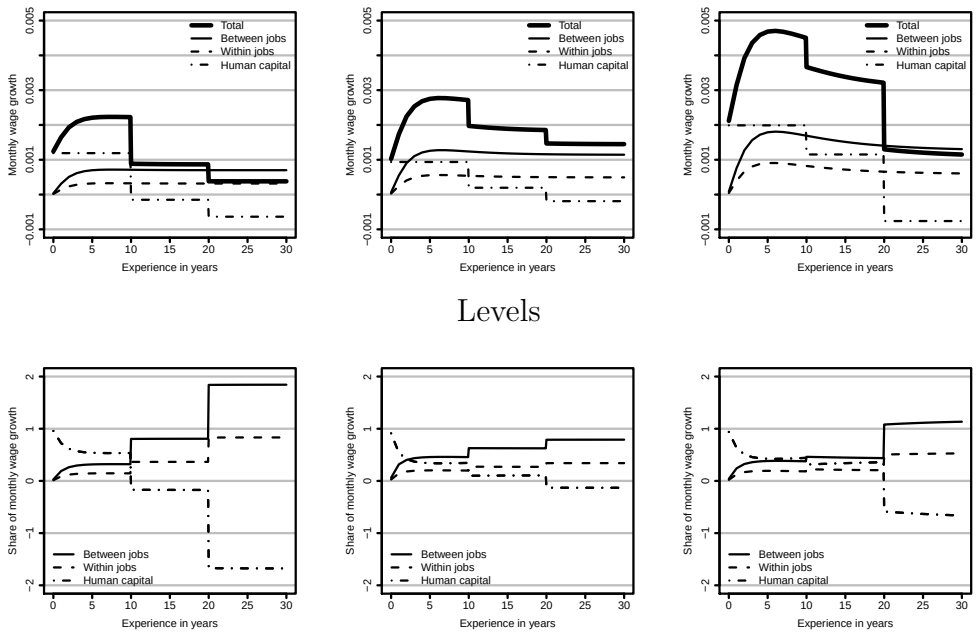
Table 6 further reports estimates of the deterministic trend in individual human capital accumulation, $g(t)$. For added legibility, those trends are also plotted in Figure 4.

There are qualitative similarities between education categories in human capital accumulation patterns. For all education categories, the pace of human capital accumulation is fastest in the first ten years of the labor market career, after which accumulation slows down, giving human capital profiles an overall concave shape.

The quantitative differences between education categories in terms of human capital accumulation patterns are striking. Low-educated workers accumulate some human capital in their first 10 years, raising their productivity by a total of 14 percent, but this initial gain in productive skills is offset by a subsequent gradual *loss* of productivity, which one may wish to interpret as fatigue or obsolescence. The rate at which that loss occur increases with experience. At 30 years of experience, cumulated productivity growth for low educated workers stands at a meagre 5 percent. At the other extreme, workers with more than 15 years of schooling grow about 24 percent more productive over the first 10 years of their careers. The human capital profile then tapers off (and even declines) for these high-educated workers towards the end of their working lives. At 20 years and 30 years of experience, the accumulated productivity growth amounts to around 37 and 28 percent, respectively. Workers in the intermediate education group have a similar profile to the low educated workers in the first 10 years of

Figure 5: Decomposition of monthly wage growth

Ed. 7-11 Ed. 12-14 Ed. 15-20



Levels

Shares

experience, but do not experience any productivity loss as their careers progress beyond 10 years of experience. In the next section we look at the implications of these productivity profiles for post-schooling wage growth.

7 Wage Growth Decomposition

The structural decomposition of monthly wage growth described in Subsection 2.5 is rendered graphically as a function of work experience in Figure 5. The three components and the total monthly wage growth by experience are plotted in the first row, together with total wage growth ($\mathbb{E}\left(\Delta w_{t+1} \mid t\right)$). The second row shows these components as a share of total wage growth. In all six plots, the solid/dashed/dash-dotted lines represent the contributions of between- and within-job wage mobility due to employers' competition for employees, and human capital accumulation, respectively, as described in Subsection 2.5.

The experience profile of wage growth is similar across education groups:

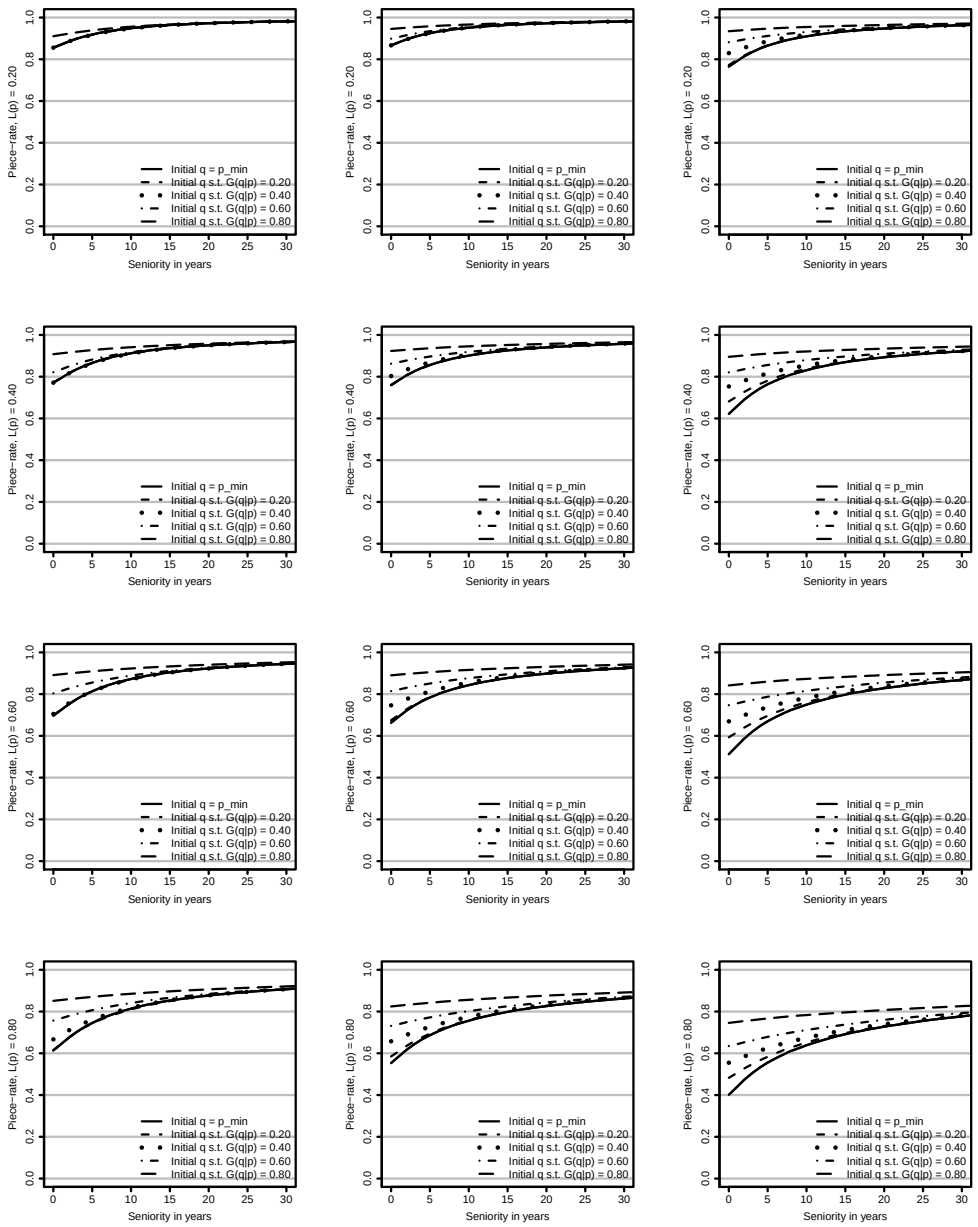
wages tend to increase at an accelerating rate during the first five years, after which growth gradually tapers off. However this overall effect of experience reflects very different evolutions of the different components of wage growth. The contribution of human capital accumulation, reflecting our estimates of $g(t)$ (Figure 4), is largest for highly educated workers and decreasing with experience for all three education groups. For all workers, general human capital contributes negatively to wage growth past 20 years of experience. As for the contribution of job search (reflecting competition between employers), both the within- and between-job spells components are slightly hump-shaped, with a steep increasing segment between 0-5 years. The between-job component dominates the within-job component at all experience levels in all three education groups. As a share of total wage growth, human capital accumulation is the mirror image of between-job wage mobility. The former steadily decreases while the latter steadily increases over the life-cycle: less experienced workers gain less in terms of wages from changing jobs. Finally, the share of within-job wage mobility, which one may view as reflecting the return to firm-specific human capital, increases with experience, but (about 50 percent) less than the between-job component.

Putting these elements together and whatever the level of confidence one is prepared to place in those specific numbers, the following stylized patterns emerge from these results. First, human capital is an important component of wage growth at the beginning of a worker's career. However, even during that early stage, job search and competition explain a large part of the wage profiles and drive most of the wage dynamics afterwards. The higher wage growth between 0 and 20 years of experience for high educated worker comes both from differences in job search frictions (especially the between-job related profiles) and more rapid human capital accumulation. It is also the case that the concavity of individual wage-experience profiles is mostly due to this decline in the rate of human capital accumulation over the life cycle.

8 Returns to Tenure

Positive "returns to tenure" arise in our structural model because the piece rates are gradually revised upward within a job spell as workers receive

Figure 6: Piece rate profiles



outside job offers. The contribution of that mechanism to average wage growth over the life cycle is measured by the "within-job wage growth" component plotted on Figure 5. This average profile, however, hides a great deal of heterogeneity. First, returns to tenure are firm-specific: one expects more productive employers to offer steeper piece rate profiles as there is more scope for upward wage renegotiation at a highly productive firm. Second, returns to tenure are not constant: they depend on the point on the firm-specific salary scale at which they are evaluated. For example, a worker just hired from unemployment tends to receive a relatively low piece rate with a lot of scope for future raises, while another worker in the same firm may have already negotiated a piece rate close to 100 percent and have very little chance of benefitting from further raises within that firm.

To illustrate and quantify both dimensions of heterogeneity, we simulate piece-rate/tenure profiles for different firm types p and renegotiation thresholds q. We select quintiles of the $L(p)$ distribution as our set of firm types p. Then, for each of those p, we consider five different piece rates corresponding to quintiles of $G(q|p)$, the distribution of renegotiation thresholds within a type-p firm (the lowest value of q for which those piece rates are calculated is therefore $q = b$, which yields the piece rate obtained by workers just hired from unemployment). For each of those (p, q) pairs we then simulate trajectories for $10,000$ workers over 30 years, switching off job-to-job transitions by assuming that outside offers are drawn from $F(\cdot)$ truncated from above at p. We finally plot average piece rate profiles for each (p, q) pair by averaging over those workers. Results are displayed in Figure 6.

As expected, more productive firms tend to offer lower starting piece rates and steeper subsequent tenure profiles. Those differences are more pronounced in higher education groups. Furthermore, returns to tenure also depend on worker history: workers with lower starting piece rates (lower initial values of q) face higher returns to tenure. Differences in initial piece rates are also persistent: in most cases, it takes over 20 years for piece rates to converge, by which point most workers have left their employer to take up a job at a more productive firm (or to become unemployed).

9 Conclusion

With the purpose of analyzing the sources of individual wage growth, we have 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 permits an in-depth comparison of our structural model to 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.

The main output of the paper is to provide a theoretically founded quantitative decomposition of individual wage growth into two terms reflecting the respective contributions of human capital accumulation and job search, the latter term being further split into a between- and a within-job spell component. The decompositions are qualitatively similar for workers with low, medium and high levels of education and reveals that human capital accumulation is the most important source of wage growth in early phases of workers' careers. However it is soon surpassed by search-induced wage growth. The wage-growth returns to job search are relatively stable after around five years of experience, with between-job wage growth dominating within-job wage growth. By implication, the well-documented concavity of individual wage-experience profiles is primarily due to the decline in the rate of human capital accumulation over the life cycle. In terms of quantitative differences between education groups, we find that more educated workers have higher total wage growth and that this reflects both more rapid human capital accumulation and higher returns to job search.

Finally, our structural model implies that that conventional log wage regression-based measures of returns to tenure conceal substantial heterogeneity.

References

Abowd, J. M., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–45.

Abraham, K. G., and H. S. Farber (1987): "Job Duration, Seniority, and Earnings," *American Economic Review*, 77(3), 278–97.

Altonji, J. G., and R. A. Shakotko (1987): "Do Wages Rise with Job Seniority?," *Review of Economic Studies*, 54(3), 437–59.

Barlevy, G. (2008): "Identification of Search Models using Record Statistics," *Review of Economic Studies*, 75(1), 29–64.

Beffy, M., M. Buchinsky, D. Fougère, T. Kamionka, and F. Kramarz (2006): "The Returns to Seniority in France (and Why are They Lower than in the United States?)," CEPR Discussion Papers 5486, C.E.P.R. Discussion Papers.

Bontemps, C., J.-M. Robin, and G. J. van den Berg (2000): "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation," *International Economic Review*, 41(2), 305–58.

Browning, M., M. Ejrnaes, and J. Alvarez (2010): "Modelling Income Processes with Lots of Heterogeneity," *Review of Economic Studies*, 77(4), 1353–1381.

Buchinsky, M., D. Fougère, F. Kramarz, and R. Tchernis (2010): "Interfirm Mobility, Wages and the Returns to Seniority and Experience in the United States," *Review of Economic Studies*, 77(3), 972–1001.

Bunzel, H., B. Christensen, N. Kiefer, and L. Korsholm (1999): "Equilibrium Search with Human Capital Accumulation," Papers 99-11, Centre for Labour Market and Social Research, Danmark-.

Burdett, K., C. Carrillo-Tudela, and M. Coles (2009): "Human Capital Accumulation and Labour Market Equilibrium," IZA Discussion Papers 4215, Institute for the Study of Labor (IZA).

Burdett, K., and M. Coles (2003): "Equilibrium Wage-Tenure Contracts," *Econometrica*, 71(5), 1377–1404.

——— (2010): "Tenure and Experience Effects on Wages: A Theory," CESifo Working Paper Series 3203, CESifo Group Munich.

Burdett, K., and D. T. Mortensen (1998): "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 39(2), 257–73.

Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): "Wage Bargaining with On-the-Job Search: Theory and Evidence," *Econometrica*, 74(2), 323–364.

Dey, M. S., and C. J. Flinn (2005): "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica*, 73(2), 571–627.

Dustmann, C., and C. Meghir (2005): "Wages, experience and seniority," *The Review of Economic Studies*, 72(1), 77–108.

Gottschalk, P., and R. Moffitt (2009): "The Rising Instability of U.S. Earnings," *Journal of Economic Perspectives*, 23(4), 3–24.

Gouriéroux, C., A. Monfort, and E. Renault (1993): "Indirect Inference," *Journal of Applied Econometrics*, 8(S), S85–118.

Guiso, L., L. Pistaferri, and F. Schivardi (2005): "Insurance within the firm," *Journal of Political Economy*, 113(5), 1054–1087.

Hall, R. E., and F. S. Mishkin (1982): "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, 50(2), 461–81.

Lazear, E. P. (2003): "Firm-Specific Human Capital: A Skill-Weights Approach," NBER Working Papers 9679, National Bureau of Economic Research, Inc.

MaCurdy, T. E. (1982): "The use of time series processes to model the error structure of earnings in a longitudinal data analysis," *Journal of Econometrics*, 18(1), 83–114.

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, pp. 1–32.

Mincer, J. (1974): *Schooling, experience, and earnings*. National Bureau of Economic Research; distributed by Columbia University Press, New York.

Mortensen, D. T. (2005): *Wage Dispersion. Why Are Similar Workers Paid Differently?* The MIT Press.

Moscarini, G., and F. Postel-Vinay (2010): "Stochastic Search Equilibrium," Cowles Foundation Discussion Papers 1754, Cowles Foundation for Research in Economics, Yale University.

Newey, W. K., and D. McFadden (1994): "Large sample estimation and hypothesis testing," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. McFadden, vol. 4 of *Handbook of Econometrics*, chap. 36, pp. 2111–2245. Elsevier.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70(6), 2295–2350.

Postel-Vinay, F., and H. Turon (2010): "On-The-Job Search, Productivity Shocks, And The Individual Earnings Process," *International Economic Review*, 51(3), 599–629.

Robin, J.-M. (2011): "On the Dynamics of Unemployment and Wage Distributions," Discussion paper, Sciences Po, Paris, forthcoming in Econometrica.

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50(1), 97–109.

Rubinstein, Y., and Y. Weiss (2006): "Post Schooling Wage Growth: Investment, Search and Learning," in *Handbook of the Economics of Education*, ed. by E. Hanushek, and F. Welch, vol. 1 of *Handbook of the Economics of Education*, chap. 1, pp. 1–67. Elsevier.

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

Topel, R. H., and M. P. Ward (1992): "Job Mobility and the Careers of Young Men," *The Quarterly Journal of Economics*, 107(2), 439–79.

van den Berg (1990): "Search Behaviour, Transitions to Non-participation and the Duration of Unemployment," *The Economic Journal*, 100(402), 842–865.

van den Berg, G., and G. Ridder (2003): "Measuring Labor Market Frictions: A Cross-Country Comparison," *Journal of the European Economic Association*, 1(1), 224–244.

Yamaguchi, S. (2010): "Job Search, Bargaining, and Wage Dynamics," *Journal of Labor Economics*, 28(3), 595–631.

APPENDIX

A Details of some theoretical results

A.1 Value function derivation

Consider (6) and integrate by parts in the r.h.s. to obtain:

$$V(r,h_t,p) = w_t + \frac{\delta}{1+\rho}V_0(h_t) + \frac{1}{1+\rho}\mathbb{E}_t\Bigg\{(1-\mu-\delta)V(r,h_{t+1},p) + \lambda_1\beta\int_p^{p_{\max}}\frac{\partial V}{\partial p}(0,h_{t+1},x)\overline{F}(x)dx + \lambda_1(1-\beta)\int_{q(r,h_t,p)}^{p}\frac{\partial V}{\partial p}(0,h_{t+1},x)\overline{F}(x)dx\Bigg\}. \quad \text{(A1)}$$

Because the maximum profitable piece rate is $r=0$, it follows that $q(0,h_t,p)=p$. Applying (A1) with $r=0$ thus yields:

$$V(0,h_t,p) = p + h_t + \frac{\delta}{1+\rho}V_0(h_t) + \frac{1}{1+\rho}\mathbb{E}_t\Bigg\{(1-\mu-\delta)V(0,h_{t+1},p) + \lambda_1\beta\int_p^{p_{\max}}\frac{\partial V}{\partial p}(0,h_{t+1},x)\overline{F}(x)dx\Bigg\}. \quad \text{(A2)}$$

Differentiating w.r.t. p:

$$\frac{\partial V}{\partial p}(0,h_t,p) = 1 + \frac{1-\mu-\delta-\lambda_1\beta\overline{F}(p)}{1+\rho}\mathbb{E}_t\frac{\partial V}{\partial p}(0,h_{t+1},p),$$

which solves as:

$$\frac{\partial V}{\partial p}(0,h_t,p) = \frac{1+\rho}{\rho+\mu+\delta+\lambda_1\beta\overline{F}(p)}. \quad \text{(A3)}$$

Substituting into (A1) yields

$$V(r,h_t,p) = w_t + \frac{\delta}{1+\rho}V_0(h_t) + \frac{1}{1+\rho}\mathbb{E}_t\Bigg\{(1-\mu-\delta)V(r,h_{t+1},p) + \lambda_1\beta\int_p^{p_{\max}}\frac{(1+\rho)\overline{F}(x)dx}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)} + \lambda_1(1-\beta)\int_{q(r,h_t,p)}^{p}\frac{(1+\rho)\overline{F}(x)dx}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)}\Bigg\}. \quad \text{(A4)}$$

A.2 Derivation of the mobility piece rate

Substitution of (A4) into (4) yields (after rearranging terms):

$$r = -(1-\beta)\left[p - q(r,h_t,p)\right] - \lambda_1(1-\beta)^2 \int_{q(r,h_t,p)}^{p} \frac{(1+\rho)\overline{F}(x)dx}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)}$$
$$+ \frac{1-\mu-\delta}{1+\rho}\mathbb{E}_t\left[(1-\beta)V(0,h_{t+2},q(r,h_t,p)) + \beta V(0,h_{t+2},p) - V(r,h_{t+2},p)\right].$$

Using the law of iterated expectations, and substituting (4) again within the expectation term in the latter equation, we obtain:

$$\begin{aligned}
r &= -(1-\beta)\left[p - q(r,h_t,p)\right] - \lambda_1(1-\beta)^2 \int_{q(r,h_t,p)}^{p} \frac{\overline{F}(x)dx}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)} \\
&\quad + \frac{(1-\mu-\delta)(1-\beta)}{1+\rho}\mathbb{E}_t\left[V(0,h_{t+2},q(r,h_t,p)) - V(0,h_{t+2},q(r,h_{t+1},p))\right] \\
&= -(1-\beta)\left[p - q(r,h_t,p)\right] - \lambda_1(1-\beta)^2 \int_{q(r,h_t,p)}^{p} \frac{\overline{F}(x)dx}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)} \\
&\quad - \frac{(1-\mu-\delta)(1-\beta)}{1+\rho}\mathbb{E}_t \int_{q(r,h_t,p)}^{q(r,h_{t+1})} \frac{\partial V}{\partial p}(0,h_{t+2},x)dx \\
&= -(1-\beta)\left[p - q(r,h_t,p)\right] - \lambda_1(1-\beta)^2 \int_{q(r,h_t,p)}^{p} \frac{\overline{F}(x)dx}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)} \\
&\quad - (1-\mu-\delta)(1-\beta)\mathbb{E}_t \int_{q(r,h_t,p)}^{q(r,h_{t+1})} \frac{dx}{\rho+\mu+\delta+\lambda_1\beta\overline{F}(p)},
\end{aligned}$$

where the last equality uses (A3). This is exactly (7).

A.3 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 (9), namely (p_{it}, q_{it}). This derivation is 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:

$$u = \frac{\mu+\delta}{\mu+\delta+\lambda_0}. \tag{A5}$$

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:

$$(\lambda_0 + \mu)ua_0(t) = \delta(1-u)a_1(t) \tag{A6}$$

$$(1-u)a_1(t) = (1-\mu-\delta)(1-u)a_1(t-1) + \lambda_0 ua_0(t-1) \tag{A7}$$

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:

$$(\mu + \lambda_0)ua_0(0) = \mu. \tag{A8}$$

Jointly solving those three equations, one obtains:

$$a_1(t) = (\mu + \frac{\mu\delta}{\mu+\lambda_0})(1-\mu-\frac{\mu\delta}{\mu+\lambda_0})^{t-1}. \tag{A9}$$

The corresponding cdf is obtained by summation:

$$A_1(t) = \sum_{\tau=1}^{t} a_1(\tau) = 1 - (1-\mu-\frac{\mu\delta}{\mu+\lambda_0})^t. \tag{A10}$$

(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 (A6): $A_0(t) = \frac{\mu(\mu+\delta+\lambda_0)}{(\mu+\delta)(\mu+\lambda_0)} + \frac{\delta\lambda_0}{(\mu+\delta)(\mu+\lambda_0)}A_1(t)$.

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:

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

Using (A6), and since (A9) implies:

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

one can rewrite (A11) as $L(p \mid t) = \Lambda_1(p)L(p \mid t-1) + \Lambda_2 F(p)$, with:

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

This last equation solves as:

$$L(p \mid t) = \left[\Lambda_1(p)^{t-1} + \Lambda_2 \frac{1 - \Lambda_1(p)^{t-1}}{1 - \Lambda_1(p)}\right] F(p). \tag{A12}$$

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

$$L(p) = \frac{(\mu + \delta)F(p)}{\mu + \delta + \lambda_1 \overline{F}(p)}. \tag{A13}$$

Conditional distribution of piece rates. Equation (8) states that piece rates are of the form $r = r(q, p)$. Thus the conditional distribution of piece rates within a type-p firm is fully characterized by the distribution of threshold values q in a type-p firm, $G(q \mid p, t)$, which we now derive. For $t > 1$, the flow-balance equation determining $G(q \mid p, t)$ is given by:

$$\begin{aligned} G(q \mid p, t)\ell(p \mid t)a_1(t) = {} & (1 - \mu - \delta - \lambda_1 \overline{F}(q))G(q \mid p, t-1)\ell(p \mid t-1)a_1(t-1) \\ & + \lambda_1 L(q \mid t-1)a_1(t-1)f(p) + (\mu + \delta)a_0(t-1)f(p), \end{aligned} \tag{A14}$$

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 (A12). Rewriting this last equation in the case $q = p$, so that $G(q \mid p, t) = 1$, yields the differential version of (A11):

$$\begin{aligned} \ell(p \mid t)a_1(t) = {} & (1 - \mu - \delta - \lambda_1 \overline{F}(p))\ell(p \mid t-1)a_1(t-1) \\ & + \lambda_1 L(p \mid t-1)a_1(t-1)f(p) + (\mu + \delta)a_0(t-1)f(p). \end{aligned} \tag{A15}$$

Dividing (A14) and (A15) by $f(p)$ throughout shows that $\frac{G(q|p,t)\ell(p|t)a_1(t)}{f(p)}$ and $\frac{\ell(q|t)a_1(t)}{f(q)}$ solve the same equation. Hence:

$$G(q \mid p, t) = \frac{\ell(q \mid t)/f(q)}{\ell(p \mid t)/f(p)} \qquad \text{for} q \in [p_{\min}, p]\,,\ t > 1. \tag{A16}$$

The unconditional version, (A17), obtains by similar reasoning:

$$G(q \mid p) = \frac{\ell(q)/f(q)}{\ell(p)/f(p)} = (\frac{\mu + \delta + \lambda_1 \overline{F}(p)}{\mu + \delta + \lambda_1 \overline{F}(q)})^2, \qquad \text{for } q \in [p_{\min}, p]. \tag{A17}$$

B Details of the sample selection criteria

Starting from the full MEE, we apply the following selection rules.

- We discard observations on firms with missing firm IDs, missing ownership structure information or missing industry information (1,141,393 observations deleted).
- The raw spell data does contain workers with gaps in their observed labor

market histories. The deletion of observations on firms with missing IDs, ownership or industry information exacerbates this problem. We remove all workers with gaps in their observed labor market histories (23,742,568 observations deleted).

- We define a temporary unemployment spell to be a short (viz. 13 weeks or shorter) non-employment, non-retirement spells (i.e combined unemployment and nonparticipation spells) in-between job spells with the same employer. Temporary unemployment spells are recoded as employment. The recoding renders some observations redundant. Furthermore, we define job spells at the level of the firm (and not the establishment). However, IDA information on employers is recorded at the establishment level and we thus aggregate establishment specific IDA information to the firm level. In doing so we assume that the industry and ownership structure of the firm are those of its largest establishment in terms of remaining workers in the analysis data. The establishment-to-firm level aggregation creates additional redundant observations. Removing these reduces the analysis data with 6,780,594 observations.

- We only keep men in the sample (41,789,290 observations deleted).

- As explained in the main text, we discard workers born before January 1st 1948, as these cohorts might have accumulated experience prior to the introduction of ATP (14,296,072 observations deleted).

- Workers are only included into our analysis sample two years after the date of graduation from their highest completed education. If a worker is ever observed in education or if the worker's education ever change after the inclusion date, all observations on that worker are removed from the dataset. At this point we also discard workers with missing or invalid education data. Using information on type of the highest completed education we compute education length (in years). As a consistency check on the education data we discard any worker who is ever observed with years of education exceeding the worker's age minus seven years. In total, we discard 9,793,603 observations at this step.

- Labor market experience is available on an annual basis and refers to the workers' experience at the end of the calender year. Experience from 1964-1979 and experience from 1980 and onwards are measured in two distinct variables. Pre-1980 experience is measured in years and post-1980 experience in 1/1000th of a year's full-time work. We impose the following consistency requirements on the experience data: First, pre-1980 labor market experience cannot change during our sample period 1985-2003. Second, workers cannot lose experience or obtain more than two years of experience during one calender year. Finally, total experience can at no time exceed the worker's age minus fifteen years. If these requirements are not met the worker is discarded (55,387 observations deleted).

- We truncate individual labor market histories at entry into retirement (546,039 observations and 10,800 workers deleted), a public sector job (4,411,620 observations and 107,057 workers deleted), self-employment (1,425,075 observations and 36,573 workers deleted), or a job in an industry for which we have no accounting data (1,161,373 observations and 46,732 workers deleted). Our data thus cover three labor market states: (private sector) employment, unemployment and nonparticipation.[22]

- Annual value added/FTE observations are transformed into hourly measures by scaling annual value added/FTE by 12×166.33 hours[23] and the strata-specific distributions are trimmed by recoding the top and bottom 1% to missing. We trend nominal variables (wages and value added) to the 2003 level using Statistic Denmark's CPI.

C Details of the simulation procedure

This Appendix describes the procedure that we implement in order to simulate a panel of I workers over T periods given values of the structural model's parameters. In practice, we have used $I = 20,000$ and $T = 228$ months (nineteen 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. To mimic that the distribution of experience in the initial cross section is capped at 21 years we draw the initial cross-section of the simulated data, conditional on experience $t \leq 21 \times 12 = 252$ months.

We begin with a sample of I workers for which we draw individual (log) heterogeneity parameters α from $\mathcal{N}(0, \sigma_\alpha^2)$. Next, we assign labor market states (employed or unemployed) to workers according to (A5), and conditional on workers' labor market states we draw labor market experience t, conditional on $t \leq 252$, according to $A_1(t)$ and $A_0(t)$ defined by (A10). 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 $L(p \mid t)$ defined by (A12). 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 $G(q \mid p, t)$ defined by (A16). Unemployed workers and employed workers with experience $t = 1$ are assigned $q = b$. Finally,

[22]Nonparticipation is a residual state (see above) and is not a rare occurrence in our panel: 47 percent of the workers in our data experience at least one nonparticipation spell, and, on average, 5.5 percent of the last-week-of-November spells are nonparticipation spells. For this reason we do not truncate labor market histories at entry into nonparticipation. However, treating nonparticipation spells as genuine unemployment spells is likely to bias our estimates of the job finding rates. Instead, we base our estimation of unemployed workers' job finding rate on genuine unemployment spells only. Job destruction rates are computed using transitions into unemployment or nonparticipation.

[23]166.33 hours being the monthly norm for a full time job.

we draw the value of the idiosyncratic productivity shock process—the ε's—conditional on labor market experience t from $\mathcal{N}(0, \sigma_u^2(1-\eta^{2t})/(1-\eta^2))$.

We give the following tweak to the draws in the steady state distributions. Firm types p are theoretically distributed according to the continuous sampling distribution $F(p)$ (Weibull as explained in the main text). Because the theoretical $F(\cdot)$ is continuous, a rigorous implementation of this would invariably produce (finite) samples with at most one worker observation per simulated firm type. To get round this problem, we discretize $F(\cdot)$ by taking a fixed number J of firm types (in practice we take $J = 100$), give each of them a rank $j = 1, \cdots, J$ and assign corresponding productivity levels of $p_j = F^{-1}(j/(J+1))$.[24] Next, to assign the p_j's to workers (conditional on experience), we draw in the usual way a I-vector $(u_1, ..., u_I)$ of realizations of $\mathcal{U}[0,1]$, and determine worker i's firm type as $p_{j(i,t)} = \arg\min_{x \in \{p_1, ..., p_J\}} \mid L(x \mid t) - u_i \mid$. Similarly, worker i's q is assigned (conditional on $p = p_j$ and $t > 1$) as $q_{it} = q_{it}(p_j) = \arg\min_{x \in \{p_1, ..., p_{j-1}\}} \mid G(x \mid p_j, t) - v_i \mid$, where v_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.

The simulation of the labor market careers of the initial cross section of workers is conducted in the following way. 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 value of the worker's productivity shock, the worker's duration of stay in the current job or unemployment spell, and if employed, the worker's employer type p and threshold value $q(\cdot)$ determining the worker's piece rate. Furthermore, in accordance with the stipulated relationship between firm types and observed value added (see equation (E4)) we draw and record an idiosyncratic disturbance z from $\mathcal{N}(0, \sigma_p^2)$ for every firm type in every period. 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[25] (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

[24] Experimenting with the value of J in the estimation revealed that our results are insensitive to different (reasonable) values of J.

[25] With respect to the sampling of firm types, we let workers draw firm ranks j (and hence corresponding productivity levels of $p_j = F^{-1}(j/(J+1))$) 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.

p', and also increment the worker's tenure 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 $\mathcal{N}(0, \sigma_\alpha^2)$.

The simulated data sets, which have monthly wage observations (computed using (9) and the information recorded for each worker), are remodeled to replicate the structure of the actual data set (which only has annual within-job average wage observations for the active job spell at the end of November—see section 3).

D Indirect inference estimation procedure

Let θ denote the vector of structural parameters, the true value of which is θ_0. We assume that the data generating process DGP(θ), our model, can be simulated for any given value of θ. Let $\beta_N(\theta_0)$ be a vector of auxiliary parameters computed on the real data. Similarly, let $(\beta_N^s(\theta), s = 1, ..., S)$ be a sequence of the same vector of auxiliary parameters, computed on each of S simulated datasets for a given value θ of the structural parameter vector. From the sequence of simulated auxiliary statistics we consider the mean: $\overline{\beta}_N^S(\theta) = \frac{1}{S}\sum_{s=1}^S \beta_N^s(\theta)$.[26] The indirect inference estimator $\widehat{\theta}_N$ minimizes the distance between $\beta_N(\theta_0)$ and $\overline{\beta}_N^S(\theta)$ in a metric defined by Ω, a positive definite matrix. Formally,

$$\widehat{\theta}_N = \arg\min_{\theta\in\Theta} Q_N(\theta) \equiv \left[\overline{\beta}_N^S(\theta) - \beta_N(\theta_0)\right]' \Omega \left[\overline{\beta}_N^S(\theta) - \beta_N(\theta_0)\right]. \tag{D1}$$

Under a set of regularity conditions on the DGP and the asymptotic behavior of $\beta_N(\theta_0)$, Gouriéroux, Monfort, and Renault (1993) show that $\sqrt{N}(\widehat{\theta}_N - \theta_0) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{W}(S, \Omega, \theta_0))$ where the covariance matrix $\mathbf{W}(S, \Omega, \theta_0)$ is given as

$$\begin{aligned}\mathbf{W}(S, \Omega, \theta_0) = (1 + \frac{1}{S})\times \\ [\mathbf{H}^S(\theta_0)'\Omega\mathbf{H}^S(\theta_0)]^{-1}\mathbf{H}^S(\theta_0)'\Omega\Sigma(\theta_0)\Omega'\mathbf{H}^S(\theta_0)[\mathbf{H}^S(\theta_0)'\Omega\mathbf{H}^S(\theta_0)]^{-1},\end{aligned} \tag{D2}$$

with $\mathbf{H}^S(\theta_0) = \text{plim}_{N\to\infty}\partial\overline{\beta}_N^S(\theta_0)/\partial\theta'$. We estimate the covariance matrix of the auxiliary statistics by re-sampling the real data, and denote the estimate by Σ_N. $\mathbf{H}^S(\theta_0)$ is estimated by numerical differentiation of $\overline{\beta}_N^S(\theta)$ evaluated at $\theta = \widehat{\theta}_N$ and is denoted $\mathbf{H}_N$.

A slight complication arises in our case because estimation is conducted in two steps. The two estimation steps utilize disjoint sets of auxiliary statistics. Our indirect estimator is thus characterized by the weighting matrix

$$\Omega = \begin{bmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 \end{bmatrix} \tag{D3}$$

[26] 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.

In our empirical implementation we take $\Omega_1 = \Sigma_{11,N}^{-1}$ and $\Omega_2 = \Sigma_{22,N}^{-1}$, where $\Sigma_{11,N}$ and $\Sigma_{22,N}$ are the covariance matrices of the sets of auxiliary statistics used in the first step and the second step, respectively.[27] The covariance matrix of the first step parameters is naturally unaffected by the two-step procedure and we report standard errors of our first step estimates $\widehat{\theta}_{1,N}$ obtained from

$$\mathbf{W}_{1,N} = (1 + \frac{1}{S})\times \\ [\mathbf{H}'_{1,N}\Omega_1\mathbf{H}_{1,N}]^{-1}\mathbf{H}'_{1,N}\Omega_1\Sigma_{11,N}\Omega'_1\mathbf{H}_{1,N}[\mathbf{H}'_{1,N}\Omega_1\mathbf{H}_{1,N}]^{-1}. \quad \text{(D4)}$$

where $\mathbf{H}_{1,N}$ is the Jacobian of the first step vector of auxiliary statistics with respect to θ_1.[28]

The covariance matrix of the second step estimator is affected by the two-step procedure (Newey and McFadden, 1994). We derive the second step covariance matrix by considering the artificial estimation problem based on stacking the first and second step vectors of auxiliary statistics into $\beta_N = (\beta'_{1,N}, \beta'_{2,N})'$ with covariance matrix Σ_N and the weighting matrix Ω defined above. Let $\mathbf{H}_N$ be the Jacobian of $\beta_N = (\beta'_{1,N}, \beta'_{2,N})'$ with respect to $\theta' = (\theta'_1, \theta'_2)$. By (D2), the covariance matrix of this artificial indirect inference estimator of θ is given as

$$\widetilde{\mathbf{W}}_N = (1 + \frac{1}{S})[\mathbf{H}'_N\Omega\mathbf{H}_N]^{-1}\mathbf{H}'_N\Omega\Sigma_N\Omega'\mathbf{H}_N[\mathbf{H}'_N\Omega\mathbf{H}_N]^{-1}, \quad \text{(D6)}$$

and since $\theta' = (\theta'_1, \theta'_2)$, $\widetilde{\mathbf{W}}_N$ is naturally partitioned as

$$\widetilde{\mathbf{W}}_N = \begin{bmatrix} \widetilde{\mathbf{W}}_{11,N} & \widetilde{\mathbf{W}}_{12,N} \\ \widetilde{\mathbf{W}}_{12,N} & \widetilde{\mathbf{W}}_{22,N} \end{bmatrix}.$$

where $\mathbf{W}_{2,N} = \widetilde{\mathbf{W}}_{22,N}$ is the covariance matrix of the second step estimator $\widehat{\theta}_{2,N}$. All computations are carried out with $S = 20$.

[27] That is, if we organize the auxiliary statistics such that $\beta = (\beta'_1, \beta'_2)'$, we get the corresponding partition of the covariance matrix Σ:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{bmatrix}.$$

[28] Since $\Omega_1 = \Sigma_{11}^{-1}$ we in fact obtain

$$\mathbf{W}_{1,N} = (1 + \frac{1}{S})[\mathbf{H}'_{1,N}\Sigma_{11}^{-1}\mathbf{H}_{1,N}]^{-1}\mathbf{H}'_{1,N}\Sigma_{11}^{-1}\mathbf{H}_{1,N}[\mathbf{H}'_{1,N}\Sigma_{11}^{-1}\mathbf{H}_{1,N}]^{-1}. \quad \text{(D5)}$$

Note however that the first step estimator is just identified and the choice of weight matrix is immaterial.

E Auxiliary Models

E.1 Labor Market Transitions

Let $c = 1, \cdots, C$ index annual last-week-of-November cross sections in the Master Panels and let P_c^{UE} (P_c^{EU}) be the probability that an unemployed (employed) worker drawn at random from cross section c workers finds a job (loses his job) within a month of being sampled. Also, let P_c^{EE} be the probability that an employed worker drawn at random makes a job-to-job transition within a month of being sampled.

As unemployed workers accept all job offers and job destruction shocks hit all employed workers with equal probability $P_c^{UE} = \lambda_0$ and $P_c^{EU} = \delta$. Hence, the empirical counterparts of the transition probabilities P^{UE} and P^{EU} identify the structural transition probabilities λ_0 and δ. Estimates of P_c^{UE} and P_c^{EU} are easily obtained from the Master Panel and we base our inference on the average of each of these transition probabilities over the C cross sections.[29]

Identification of the job offer arrival probability λ_1 is complicated by the fact that we only observe accepted job offers. Indeed, the unconditional probability job-to-job transition probability is $P_c^{EE} = \int_{\underline{p}}^{\overline{p}} \lambda_1 \overline{F}(p)\, dL(p|t \leq T_c)$ where $L(p|t \leq T)$ is the cross section distribution of firm types among employed workers with experience T or less and T_c is the maximum experience level in cross section c. It is easy to show that the integral is a function of transition parameters only, and that, given λ_0, δ and μ (the latter of which we shall not attempt to estimate), it represents a one-to-one mapping between the observable P_c^{EE} and the unobservable λ_1.[30]

E.2 Wages and Labor Market Experience

E.2.1 Log Wage Regression

The estimated coefficients on experience and tenure from Mincer-type wage regressions, along with summary statistics of worker and firm heterogeneity in the wage data, are central components in our vector of auxiliary statistics. These statistics facilitate identification of both the structural human capital accumulation function $g(\cdot)$ and the sampling distribution $F(\cdot)$ of firm types faced by searching workers. Moreover, the regressions provide a direct link between our structural approach and the literature on wage-experience and wage-tenure profiles.

The auxiliary wage regression is estimated on a panel of repeated annual (last-week-of-November) cross sections of employed workers extracted from the Master

[29] We retain only employment- and unemployment spells that are either censored or end in a transition to employment or unemployment. Recall that nonparticipation is a residual state that may not correspond to nonparticipation in the usual economic sense.

[30] To see this, notice from the derivation in Appendix A that $L(p|t \leq T)$ only depends on p through $F(p)$, and use the change of variables $x = F(p)$ in the integral (van den Berg and Ridder, 2003).

Panel.[31] Let $i = 1, 2, \cdots, I$ index individuals and, as before, let $c = 1, 2, \cdots, C$ index the annual cross sections. Let $j = 1, 2, \cdots, J$ index firms and let $\mathrm{J}(i, c)$ be the firm-ID of worker i in cross section c. Hence, the data used for estimation of the wage regressions can be represented by $\{w_{ic}, s_{ic}, t_{ic}, d_{ic}, i, \mathrm{J}(i, c)\}_{i=1,c=1}^{I,C}$, where w_{ic}, t_{ic} and s_{ic} are worker i's wage, experience and job tenure in cross section c, respectively. A worker's job tenure is time elapsed since the start of the job (if non-censored) or January 1st 1985 (if left-censored).[32] d_{ic} is a binary indicator for left-censoring of worker i's job in cross-section c. We let seniority and experience enter the wage regression via linear spline functions:

$$\begin{aligned}\ln w_{ic} = \sum_{k=1}^{3} \xi_{1k}(s_{ic} - s_{1k}^*)\mathbf{1}_{\{s_{ic} \geq s_{1k}^*\}} d_{ic} \\ + \sum_{k=1}^{3} \xi_{2k}(s_{ic} - s_{2k}^*)\mathbf{1}_{\{s_{ic} \geq s_{2k}^*\}}(1 - d_{ic}) \\ + \sum_{k=1}^{3} \xi_{3k}(t_{ic} - t_k^*)\mathbf{1}_{\{t_{ic} \geq t_k^*\}} + \psi_i + \phi_{\mathrm{J}(i,c)} + u_{ic}, \quad \text{(E1)}\end{aligned}$$

where $\{s_{1k}^*, s_{2k}^*, t_k^*\}_{k=1}^3$ are the knots of the spline functions (not estimated), ψ_i and ϕ_j are time-invariant unobserved worker and firm effects and u_{ic} is the residual. We assume that the worker effects are orthogonal to all other components in the auxiliary wage regression.

The auxiliary wage regression (E1) predicates a wage decomposition similar to the structural wage equation (9), except for the fact that, the type of the last employer from which the worker was able to extract surplus (q in the notation of the structural model) being unobserved in the data, tenure s_{ic} proxies for this factor.

We estimate the parameters relating to tenure and experience by applying within-firm OLS to (E1). Firm and worker effects are subsequently recovered from the resulting residuals in two steps (first firm effects, then the worker effects). Finally, to further describe wage dynamics in our data we select sequences of consecutive within-job log wage residuals containing at least five observations and compute the average residual autocovariances of order up to 4 over these observations.

E.2.2 Log Wage Growth Regression

Using the auxiliary wage equation (E1) we can consider the autocorrelation structure of within-job wage growth, which is what the estimation of statistical models

[31] Recall that wage information is only available for jobs that active in the last week of November.

[32] We only observe labor market histories from January 1st, 1985. All jobs that are ongoing at this time will have left-censored tenure. Rather than discarding these observations we separately control for tenure and left-censored tenure in the wage regression. Our indirect estimation procedure allow us to handle this somewhat unusual specification of the wage regression.

of earnings dynamics is typically based on (see e.g. Browning, Ejrnaes, and Alvarez 2010). For simplicity, we condition the analysis on worker i staying in the same firm between experience levels t and $t+1$. Taking first differences in equation (E1) under this restriction yields the following auxiliary model for within-job wage growth:

$$\Delta \ln w_{ic} = \zeta_1 + \zeta_2 \Delta(t_{ic} - t_2^*)\mathbf{1}_{\{t_{ic} \geq t_2^*\}} + \zeta_3 \Delta(t_{ic} - t_3^*)\mathbf{1}_{\{t_{ic} \geq t_3^*\}} + \Delta u_{ic}. \qquad \text{(E2)}$$

First-differencing eliminates the firm and worker fixed heterogeneity components. Moreover we only include experience in the r.h.s. of (E2) as, within a job spell, experience and tenure are indistinguishable.

We estimate (E2) directly rather than using the estimated residuals u_{ic} from (E1) for two reasons. First, contrary to u_{ic}, the residuals from (E2) are not affected by estimation errors on the firm and worker effects. Second, the estimation of (E2) provides us with additional slope parameters $\zeta = (\zeta_1, \zeta_2, \zeta_3)'$ which convey information and can be incorporated into the set of moments to match.[33]

The auxiliary wage growth model (E2) is estimated by OLS. We include the estimated slope parameters ζ', the residual autocovariances as well as the standard deviation on the residual in the set of moments to match.[34]

E.3 Firm Productivity

The Master Panels contains data on firms' annual value added which we convert into an hourly measure. We face three problems using this data. First, our structural model assumes that a firm is a collection of jobs with the same productivity p. Hence, we cannot directly use firm-specific moments without making the implausible assumption that the underlying distribution of firm types is exactly identical to the sampling distribution faced by workers. Instead, we match employment-weighted firm specific moments. Second, the value added data originates in a survey which Statistics Denmark constructs using a known sampling scheme, which we must take into account by appropriately re-weighting moments of the value added distribution. Third, we are not able to attribute output to individual workers. Hence, we do not have a direct measure of the contribution to observed output of the different education groups. We circumvent this problem by positing a strata-specific relationship between average output per worker and labor productivity p (to be specified soon).

In terms of data we select the 1999-2003 last-week-of-November cross sections and, from here, we keep only observations on jobs with wage information. We split

[33]Note that we do not impose consistency of coefficient estimates between the auxiliary log-wage equations in levels (E1) and growth rates (E2). According to our structural model, this pair of equations is a misspecified representation of the individual earnings process and one should therefore not expect it to be consistent in any particular way.

[34]The auxiliary wage growth equation is estimated on the subsample of job spells with at least two consecutive annual wage observations, so as to make first differencing possible. More consecutive observations will be needed when we later compute residual autocovariances from (E2). We report autocovariances up to order 4, so that autocovariances are computed from the subset of jobs with at least six observations.

Table 7: Sampling scheme for accounting data

	Statistics Denmark		Empirical sampling probabilities		
Labor force size	P	Years in/out	Ed. 7-11	Ed. 12-14	Ed. 15-20
0-9 (ω_1)	0.10	1/9	0.07	0.07	0.09
10-19 (ω_2)	0.20	2/8	0.29	0.29	0.30
20-49 (ω_3)	0.50	3/3	0.62	0.60	0.61
$>$ 49 (ω_4)	1.00	-	0.95	0.93	0.91

Note: P is the theoretical sampling probability. The empirical sampling probabilities computed from the pooled 1999-2003 last-week-of-November cross sections. Statistics Denmark also include firms with revenue exceeding DKK 100 mill. (in Wholesale DKK 200 mill.). Statistics Denmark in fact sample 10% of firms with 5-9 employees and no firms with 0-4 employees (unless revenue is sufficiently high). See footnote 22 for a rational for lumping firms with 0-4 employees together with firms with 5-9 employees.

Source: Matched Employer-Employee data obtained from Statistics Denmark.

the Master Panel observations on employees into four bins depending on the size of the employer's workforce (measured in the raw data, before the selection of the Master Panel): 0-9 employees, 10-19 employees, 20-49 employees and more than 49 employees. Within each bin a random sample of employers have been selected to submit their accounting data.[35] Rather than using the sampling probabilities used by Statistics Denmark, we compute the actual fractions of employees in our Master Panel with value added information in each bin. These fractions are denoted ω_1, ω_2, ω_3 and ω_4 and are tabulated in Table 7.[36] Using these weights we compute employment weighted mean, variance, skewness and kurtosis of log value added.

We index the individual bin b observations by $n = 1, 2, \cdots, N_b$ and compute moments of the distribution of log labor productivity y by weighting the bin-specific observations by the empirical sampling probabilities. That is, we estimate the employment weighted mean log labor productivity as $\widehat{E(y)} = \frac{1}{N}\sum_{b=1}^{4}\sum_{n=1}^{N_b}\omega_b^{-1}y_{nb}$ where $N = \sum_{b=1}^{4}\omega_b^{-1}N_b$ is the total number of observations in the four cross sections (with or without accounting data). Variance, skewness and kurtosis are computed in an analogous fashion. We further include moments of the raw log wage distribution and the covariance between employer-specific average log output and individual log wages.[37]

[35]Statistics Denmark does not sample employers with less than 5 employees in the last week of November. However, the data does contain such small firms with accounting data and we include the 0-4 employees employers bin in the 5-9 employees employer bin.

[36]The fact that the empirical sampling probabilities in our Master Panel are relatively close to the sampling probabilities applied by Statistics Denmark confirms our expectation that the selection criteria imposed to extract the Master Panel from the raw data does not seriously hamper our ability to generalize our findings to the full labor market.

[37]Individual wages are not subject to sampling, and so, $\widehat{E(w)}$ is straightforward to com-

As we mentioned before, we only observe firm-level average output per worker, denoted y_j in log terms, while the firm-specific productivity parameter that is relevant in the model is p_j, a measure of the marginal productivity of labor for a given education category. To get around that problem we assume that y_j can be expressed as a function of the p_j's of the various education groups, of which we take linear approximations (one for each education group). Specifically, we impose the following reduced-form relationship between observed firm-specific average output $y_{\mathrm{J}(i,c)}$ and education-group specific labor productivity $p_{\mathrm{J}(i,c)}$:

$$y_{\mathrm{J}(i,c)c} = \kappa_0 + \kappa_1 p_{\mathrm{J}(i,c)} + z_{\mathrm{J}(i,c)c} \tag{E4}$$

where $z_{\mathrm{J}(i,c)c} \sim \mathcal{N}(0, \sigma_p^2)$ is i.i.d. across firms and time, and where $\mathrm{J}(i,c) = j$ if and only if worker i is employed by firm j in cross section c. The parameter σ_p is easily estimated from the dispersion of within-job year-to-year differences in observed labor productivity. Consistent with the specification of $z_{\mathrm{J}(i,c)}$ as idiosyncratic measurement and/or specification errors we compute an estimate of σ_p as the non-weighted empirical standard error of of $\Delta z_{\mathrm{J}(i,c)}$ where the difference operator is applied to observation (i,c) if and only if $\mathrm{J}(i,c) = \mathrm{J}(i,c-1)$.

Hence, the reduced form model (E4) introduces three additional parameters, but allow for an additional six moments related to the distribution of p to be included in the estimation and thus adds to the identification of our structural parameters.

pute. The covariance between log value added y and log individual wage w is computed as

$$\widehat{\mathrm{Cov}(y,w)} = \frac{1}{N}\sum_{b=1}^{4}\sum_{n=1}^{N_b}\omega_b^{-1}(y_{nb} - \widehat{E(y)})(w_{nb} - \widehat{E(w)}). \tag{E3}$$