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

Jesper Bagger†
Royal Holloway College
University of London

François Fontaine‡
Université de Lorraine
BETA and CREST

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

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

Revised: April 2013

*Earlier versions of this paper were circulated since 2006 under the title "A Feasible Equilibrium Search Model of Individual Wage Dynamics with Experience Accumulation". We are grateful to three anonymous referees for very detailed and constructive comments. We also wish to thank Joe Altonji, Henning Bunzel, Ken Burdett, Melvyn Coles and Rune Vejlin, 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. Career wage growth is decomposed into the contributions of human capital and job search, within and between jobs. Human capital accumulation is largest for highly educated workers, and both human capital accumulation and job search contribute to the observed concavity of wage-experience profiles. The contribution from job search to wage growth, both within- and between-job, declines over the first ten years of a career — the 'job-shopping' phase of a working life — after which workers settle into high-quality jobs and use outside offers to generate gradual wage increases, thus reaping the benefits from competition between employers.

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 working life. 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 Bunzel, Christensen, Kiefer, and Korsholm (1999), Rubinstein and Weiss (2006), Barlevy (2008), Burdett, Carrillo-Tudela, and Coles (2009), Yamaguchi (2010), and Veramendi (2011). Although these contributions have features not included in this paper, none simultaneously allows for worker and firm heterogeneity, idiosyncratic productivity shocks and human capital accumulation. Furthermore, none uses Matched Employer-Employee (MEE) data on both firm output and worker wages, which are required to ensure 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 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

[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.

(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 implausibly 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 produces 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 insignificant (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, 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

[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.

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 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 heterogeneity in productivity 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 induce the worker to accept a lower initial wage. In the interest 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 proven 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, Altonji, Smith, and Vidangos, 2009). The overwhelming majority of papers in that literature focus solely on wages and are silent about how

[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.

[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.

productivity shocks impact wages.[6] Our model offers a simple theoretical structure within which to think about the impact of firm-level productivity heterogeneity 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 explain 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 and the estimated model replicates conventional measures of labor market transitions, tenure and experience effects, and residual wage dynamics well. The decomposition of individual wage growth is qualitatively similar across education groups, but reveals that more educated workers have higher total wage growth. This reflects both more rapid human capital accumulation and, at the early stage of worker's career, higher returns to job search. Both human capital accumulation and job search contribute to the concavity of wage-experience profiles. The contribution from job search to wage growth, both within- and between-job, declines within the first ten years of a career, a period that we identify as the 'job-shopping' phase of a working life. After that initial period, workers settle into high-quality jobs and use outside offers to generate gradual wage increases, thus reaping the benefits from competition between employers. Indeed, the within-job component always dominates the between-job component, but especially so after ten years of labor market experience.

Additional model-based decompositions of means and variances of log wages, conditional on experience level, reveal that the wage-experience profile is almost entirely explained by human capital accumulation and an increasing mean employer productivity (due to 'job shopping'). Among highly educated workers, human capital accounts for about half of the accumulated growth at all experience levels. The weight of human capital is smaller among workers with medium or low education (a fifth to a quarter). Cross sectional log-wage variance is increasing with labor market experience at a declining rate (in the data and in the model). This is almost entirely driven by increased dispersion in employer productivity. The level of log-wage variance is explained for the most part

[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.

[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.

by dispersion in both employer productivity and piece rate contracts, the contributions of dispersion in worker ability and individual productivity shocks being comparatively small.

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 arises 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, 4 and 5 we present the data, the econometric model, and the estimation protocol. In sections 6 and 7 we discuss estimation results including the structural model's fit to the data, and in sections 8 and 9 we analyze decompositions of individual wage-experience and wage-tenure profiles. Section 10 concludes.

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

$$y_t = p + h_t, \qquad h_t = \alpha + g(t) + \varepsilon_t, \tag{1}$$

where $p \in [p_{\min}, p_{\max}]$ 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 the sum of three components: 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. At this point we do not attach any more specific

interpretation to the ε_t shock. It reflects stochastic changes in individual productivity that may come from preference or technological shocks, or from public learning about the worker's quality. This shock is worker-specific, and we only restrict it to follow a first-order Markov process. A useful benchmark may be to think of it as a linear $AR(1)$ process, possibly with a unit root.

Note that this specification implies that firm productivity p and human capital h_t are complementary in production. A central planner would thus want to reallocate workers as they accumulate human capital.

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 productivity is updated from h_{t-1} to h_t as per equation (1). We assume that unemployed workers do not accumulate experience, so that if a worker becomes unemployed at an experience level of $t-1$, her 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 match dissolved with probability δ, or receives an outside offer with probability λ_1 (with $\mu+\delta+\lambda_1 \leq 1$). When a match is dissolved, we allow for the possibility that the worker finds a new employer right away, without an intervening unemployment period, an event that occurs with probability κ. This is a simple way of modeling the transition patterns observed in the data. With a slight abuse of terminology, we refer to κ as a reallocation probability. In reality, the (unconditional) probability of reallocation, an involuntary job-to-job transition, is $\delta\kappa$. It follows that the probability that a match is dissolved and the worker enters the state of unemployment is $\delta(1-\kappa)$. When unemployed, a worker finds a new match with probability λ_0 (such that $\mu+\lambda_0 \leq 1$). Upon receiving a job offer, any worker (regardless of her/his employment status or human capital) draws the type p of the firm from which the offer emanates from a continuous, unconditional sampling density $f(\cdot) = F'(\cdot)$, with support $[p_{\min}, p_{\max}]$.

2.2 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 adapted from 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 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 all of the match output to the worker and leaves the employer with zero profit from that particular match. The maximum value that a worker can hope to extract from a match 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 key 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 more—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

$$\mathbb{E}_t V(r', h_{t+1}, p') = \mathbb{E}_t \left\{ V(0, h_{t+1}, p) + \beta \left[V(0, h_{t+1}, p') - V(0, h_{t+1}, p) \right] \right\}, \tag{2}$$

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 worker rent brought about by the match with the type-p' firm. We refer to β as the worker's bargaining power.[8]

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'

[8]Strictly speaking, we cannot directly invoke the Nash bargaining solution to rationalize (2) because of our assumption that workers have logarithmic utility. Yet Cahuc, Postel-Vinay, and Robin (2006) rationalize (2) as the equilibrium of a strategic bargaining game adapted from Rubinstein (1982), which does directly apply to the case in this paper.

falls short of the threshold value $q \equiv q(r, h_t, p)$, defined by the indifference condition

$$\mathbb{E}_t V(r, h_{t+1}, p) = \mathbb{E}_t \left\{ V(0, h_{t+1}, q) + \beta \left[V(0, h_{t+1}, p) - V(0, h_{t+1}, q) \right] \right\}. \tag{3}$$

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 (who contact an employer with probability λ_0), and employed workers whose job is destroyed but who immediately contact a new employer (with probability κ), we consistently assume that they receive a share β of the expected match rent. The piece rate r_0 thus obtained by an unemployed worker with experience level t solves

$$\mathbb{E}_t V(r_0, h_{t+1}, p) = V_0(h_t) + \beta \mathbb{E}_t \left[V(0, h_{t+1}, p) - V_0(h_t) \right]. \tag{4}$$

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

We assume that unemployment is equivalent to employment in the least productive firm of type $p_{\min}$: $V_0(h_t) = \mathbb{E}_t V(0, h_{t+1}, p_{\min})$. The implication is that an unemployed worker accepts any job offer she receives. In an environment with search frictions, human capital accumulation, and different arrival rates on- and off-the-job, the reservation strategy of an unemployed worker would, in general, depend on the worker's experience level. This complication would cause a loss of analytical tractability. The assumption of a constant reservation productivity (equal to $p_{\min}$) is partly justified by empirical findings of the offer acceptance rate of unemployed workers being close to one (see e.g. van den Berg, 1990).

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 recursively as:

$$\begin{aligned} V(r, h_t, p) = w_t + \frac{\delta(1-\kappa)}{1+\rho} V_0(h_t) + \frac{\delta\kappa}{1+\rho} \int_{p_{\min}}^{p_{\max}} \mathbb{E}_t \left[(1-\beta) V_0(h_t) + \beta V(0, h_{t+1}, x) \right] dF(x) \\ + \frac{\lambda_1}{1+\rho} \int_{p}^{p_{\max}} \mathbb{E}_t \left[(1-\beta) V(0, h_{t+1}, p) + \beta V(0, h_{t+1}, x) \right] dF(x) \\ + \frac{\lambda_1}{1+\rho} \int_{q(r,h_t,p)}^{p} \mathbb{E}_t \left[(1-\beta) V(0, h_{t+1}, x) + \beta V(0, h_{t+1}, p) \right] dF(x) \\ + \frac{1}{1+\rho} \left[1 - \mu - \delta - \lambda_1 \overline{F}(q(r, h_t, p)) \right] \mathbb{E}_t V(r, h_{t+1}, p), \end{aligned} \tag{5}$$

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 (3).

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 $\delta(1-\kappa)$, the worker becomes unemployed, a state that s/he values at $V_0(h_t)$. With probability $\delta\kappa$, the worker's job is dissolved but she manages to immediately obtain a new offer from a type-x employer, drawn from the offer distribution $F(x)$. Bargaining with this employer, with unemployment as an outside option, results in value $\mathbb{E}_t\left[(1-\beta)V_0(h_t) + \beta V(0, h_{t+1}, x)\right]$. With probability λ_1, the worker receives an outside job offer emanating from a type-x firm drawn from $F(x)$, and one of three scenarios applies: 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 her initial contract with updated human capital, which has expected value $\mathbb{E}_t V(r, h_{t+1}, p)$. With probability μ, the worker leaves the labor force permanently and receives a value of 0, and with complementary probability $1 - \delta - \mu - \lambda_1$, nothing happens and the worker carries on with her initial contract with updated human capital (expected value $\mathbb{E}_t V(r, h_{t+1}, p)$).

In Appendix A we make use of equation (5) to show that equation (3) has a simple, deterministic (indeed constant), consistent solution $q(r, p)$ implicitly defined by:

$$r = -\int_{q(r,p)}^{p} \phi(x)\, dx, \qquad \phi(x) = (1-\beta)\frac{\rho + \delta + \mu + \lambda_1 \overline{F}(x)}{\rho + \delta + \mu + \lambda_1 \beta \overline{F}(x)}. \tag{6}$$

Now even though (6) implies no direct dependence of $q(\cdot)$ on t or h_t, other, nondeterministic solutions to (3) may still exist if agents expect future values of $q(\cdot)$ to depend on future values of h. We will ignore the possibility of such sophisticated expectational mechanisms in this paper, and concentrate on this deterministic solution.

2.4 The Empirical Wage Process

Under the deterministic solution (6), the (log) wage w_{it} earned by worker i hired at a firm with productivity p_{it} at time t is defined as follows:

$$w_{it} = \alpha_i + g(t) + \varepsilon_{it} + p_{it} - \int_{q_{it}}^{p_{it}} \phi(x)\, dx, \tag{7}$$

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 effect α_i, a transitory worker productivity shock ε_{it}, an employer effect p_{it}, and a random variable q_{it} relating to the most recent wage bargain. The worker's wage is equal to her marginal productivity if $q_{it} = p_{it}$, that is if she managed to force her employer to compete with an equally productive poacher.

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 probability $\delta(1-\kappa)$ she becomes unemployed, in which cases the value of $(p_{i,t+1}, q_{i,t+1})$ is set to missing; otherwise the worker may experience a reallocation with probability $\delta\kappa$ or draw an outside offer with probability λ_1. Hence, given (p_{it}, q_{it}), $(p_{i,t+1}, q_{i,t+1})$ is drawn from the following distribution:

$$(p_{i,t+1}, q_{i,t+1}) = \begin{cases} (\cdot,\cdot), & \text{with probability } \mu + \delta(1-\kappa) \\ (p_{it}, q), \forall q \in (q_{it}, p_{it}], & \text{with density } \lambda_1 f(q) \\ (p, p_{it}), \forall p > p_{it}, & \text{with density } \lambda_1 f(p) \\ (p, p_{\min}), & \text{with density } \delta\kappa f(p) \\ (p_{it}, q_{it}), & \text{with probability } 1 - \mu - \delta - \lambda_1 \overline{F}(q_{it}) \end{cases} \tag{8}$$

If the worker is unemployed in period t, then $(p_{i,t+1}, q_{i,t+1}) = (p, p_{\min})$ with density $\lambda_0 f(p)$.

Our model conveys natural interpretations of the wage returns to experience and tenure. Experience has both a direct causal impact on wages through human capital accumulation, reflected in the term $g(t)$ in (7), and an indirect effect through employed job search and the fact that, because of voluntary job-to-job transitions, more experienced workers tend to be higher up the job ladder — i.e., they tend to be in higher-p jobs. Tenure, on the other hand, has no direct causal impact on wages, but it has an indirect effect through the fact that, conditional on employer type p, workers with longer tenure tend to have received more outside job offers, and therefore to be on a higher piece rate based on a higher value of q.

All three stochastic components of wages (p_{it}, q_{it} and ε_{it}) are unobservable in standard worker panel data sets. Moreover, both q_{it} and p_{it} are correlated with tenure: workers are harder to poach out of matches with high-p firms, which therefore tend to last longer. According to our model, tenure will be positively correlated with wages in a cross section because employers are forced to increase wages to retain their employees when they are approached by competitors. Moreover, starting wages immediately following a voluntary

job-to-job mobility will be correlated with tenure in the previous job because any successful poacher had to compete with an incumbent employer to hire a worker: workers poached out of high-p firms tend to have both longer past tenure and higher starting wages after a job-to-job mobility. The latter statement is not true if the worker moves to another job following a match dissolution shock with no intervening unemployment. In this case the type of the previous employer is not relevant for the observed starting wage; see (4).

2.5 Steady-state Equilibrium

The bilateral wage determination process described above pins down wages as functions of four random variables, namely worker experience t, worker ability α_i, and two employer types p_{it} and q_{it}. Our final task is to characterize the equilibrium allocation, i.e. the joint distribution of $(t, \alpha_i, p_{it}, q_{it})$ that prevails in equilibrium. To that end, we follow the majority of the job search literature and assume that the economy is at a steady state, i.e. that the equilibrium allocation remains stable over time.

Under the steady-state assumption, equilibrium distributions can be derived from flow-balance equations reflecting the equality of flows in and out of various aggregate stocks. The steady-state cross-sectional distribution of $(t, \alpha_i, p_{it}, q_{it})$ is derived in Appendix A. Especially useful for what follows are the equilibrium marginal distribution of firm productivity p_{it} across employed workers, denoted as $L(p)$, and the equilibrium distribution of q_{it} given p_{it}, denoted $G(q|p)$. We establish in Appendix A that:

$$L(p) = \frac{(\mu+\delta)F(p)}{\mu+\delta+\lambda_1\overline{F}(p)}, \qquad \text{and} \qquad G(q|p) = \left[\frac{\mu+\delta+\lambda_1\overline{F}(p)}{\mu+\delta+\lambda_1\overline{F}(q)}\right]^2, \tag{9}$$

for $p > p_{\min}$ and $q \in [p_{\min}, p]$.

Let us derive the equation for $L(p)$ as an example. $L(p)(1-u)$ is the stock of all employees at firms with productivity less than p. The exit rate from that stock is $\mu + \delta(1-\kappa) + (\delta\kappa+\lambda_1)\overline{F}(p)$, where $\delta(1-\kappa)$ is the transition rate into unemployment and $\delta\kappa+\lambda_1$ is the overall job-to-job mobility rate, adding reallocation shocks to voluntary employer changes. So $[\mu + \delta(1-\kappa) + (\delta\kappa+\lambda_1)\overline{F}(p)]\, L(p)(1-u)$ is the outflow from the stock of employed workers at firms of productivity p or less. The inflow into that same stock has two component: first, $\lambda_0 u F(p)$ initially unemployed workers receive offers from firms with types less than p in the period. Second, $\delta\kappa[1-L(p)](1-u)F(p)$ workers initially employed at firms with types greater than p are reallocated to firms with types less than p. The total inflow into the stock $L(p)(1-u)$ of workers employed at firms with productivity p or less is thus $\lambda_0 u F(p) + \delta\kappa[1-L(p)](1-u)F(p)$. Further using the steady-state flow equation for unemployment, which implies that $\lambda_0 u F(p) = [\mu+\delta(1-\kappa)](1-u)$, we can rewrite that total inflow as $\{\mu+\delta(1-\kappa)+\delta\kappa[1-L(p)]\}(1-u)F(p)$. Finally equating

the inflow with the outflow yields the above expression for $L(p)$.

Another property of the equilibrium allocation is that worker ability α_i is uncorrelated with employer type. In other words, the model does not generate sorting on unobservable worker and firm types in equilibrium. This happens despite the existence of a complementarity between worker type α and firm type p in production (the mach output is multiplicative in α and p). Two important assumptions are driving this property. First, all matches contribute additively to total firm output, so that the productive type of an additional recruit does not affect output from existing matches. This is a very strong assumption, but adding complementarities between employees in production makes the determination of wages and the dynamics of the distribution of worker types within a firm extremely difficult to solve. Second, the value of a vacancy is zero, and the flow value of non-employment *and* the flow output of a match with any firm are both proportional to worker ability. This assumption makes the surplus of a match multiplicative in worker ability, and the decisions to leave unemployment (positive surplus) or to change employer (go for the higher surplus) become independent of worker ability.

Any departure from these two assumptions will generate endogenous sorting in equilibrium and will complicate the equilibrium solution tremendously. Yet, we can still obtain sorting without changing the model by making the job finding rates $(\lambda_0, \lambda_1, \kappa)$ functions of worker ability α_i. Then the distribution of p_{it} given α_i takes the same form as above with $\kappa_i = \kappa(\alpha_i)$ and with $\lambda_i = \lambda(\alpha_i)$. If, say, more able workers are also more efficient at job search, and therefore have higher job contact probabilities λ_i and κ_i, they will climb the job ladder faster and thus tend to be employed in more productive firms than less able workers. We will use this extension in the empirical analysis.

3 Data

We estimate our model using a comprehensive Danish Matched Employer-Employe (MEE) panel covering the period 1985-2003. The backbone of this data is a panel of individual labor market histories (the spell data), which 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, non participation and retirement.[9] The data is weekly and firm- (not establishment-) 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

[9]Non participation 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.

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.[10] The information on workers' labor market experience refers to the workers' actual (as opposed to potential) experience at the end of a calendar year. Experience is constructed from workers' mandatory pension payments ATP and goes back to January 1st, 1964.[11]

Finally, we use information on firms' accounts collected by Statistics Denmark in annual surveys in 1999-2003.[12] 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 their workforce size (see Appendix B for details on the sampling scheme).

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, indicating the high 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 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 11). 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 non participation) 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), 12-14 years (completion of high school, or vocational education) and 15-20 years (bachelor level or higher). The data set is structured such that there is one observation per worker per

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

[11] 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.

[12] 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.

year per spell.

We refer to the 19-year long (unbalanced) panels of individual labor market histories thus constructed as "Master Panels". The fact that we observe individual labor market histories over a long period of time, coupled with information on actual labor market experience, as well as employer ID and measures of employer productivity, makes this particular dataset ideal for identifying and estimating our model featuring on-the-job search, human capital accumulation, and two-sided heterogeneity with productivity shocks. Few other dataset would allow this type of exercise. Table 1 provides summary statistics on the Master Panels.

4 The Econometric Model

We shall estimate a different model for each education group that we treat as different labor markets with different workers and different employers. Education is assumed exogenous. There is no decision regarding education or human capital accumulation (except that workers take into account that unemployment freezes human capital accumulation). We shall thus estimate a different distribution of worker and firm heterogeneity for each skill-market. We now describe the specification of the structural parameters without indexing them on education, but it should be understood that all parameters are education- or skill-specific.

The matched employer-employee data provide log-wage observations w_{ic} for a sample of workers indexed by i and calendar time c, observations of log output (value added) per worker y_{jc} for a sample of firms indexed by j and time c, and a link function $j = \mathrm{J}(i, c)$ assigning an employer j to worker i at time c.

The model above provides a detailed description of the dynamics of individual wages. However, it says nothing about the dynamics of firm output. In the absence of a fully convincing model of the firm, we assume that log output per worker y_{jc} can be expressed

as a function of firm heterogeneity p_j via a simple linear relationship,

$$y_{jc} = \chi_0 + \chi_1 p_j + z_{jc}, \tag{10}$$

where the idiosyncratic output shock is assumed iid normal: $z_{jc} \sim \mathcal{N}(0, \sigma_p^2)$. We can afford this extra couple of parameters (χ_0, χ_1), as is that the sampling distribution of p_j, F, is partially identified by wages: first, the minimum value of p_j, $p_{\min}$, is related to reservation wages (the value of unemployment is by assumption the value of a job with type $p_{\min}$); second, all wages in a firm j, per unit of human capital, are bounded above by p_j. The sampling distribution of firm types is assumed Weibull: $F(p) = 1 - \exp(-[\nu_1(p - p_{\min})]^{\nu_2})$, where the location parameter $p_{\min}$, the scale parameter ν_1 and the shape parameter ν_2 are three parameters to be estimated. We shall estimate different parameters (χ_0, χ_1) as well as different distributions of p_j for each education group. We thus assume perfect substitutability between workers both between and within skill groups.

As for worker human capital, it is such that the distribution of permanent worker heterogeneity α is normal $\mathcal{N}(0, \sigma_\alpha^2)$. Individual-specific productivity shocks follow a Gaussian AR(1) process: $\varepsilon_{it} = \eta \varepsilon_{i,t-1} + u_{it}$, with $u_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. The deterministic trend is cubic: $g(t) = \gamma_1 t + \gamma_2 t^2 + \gamma_3 t^3$.

We allow for worker-level heterogeneity in transitions rates by specifying λ_0, κ and λ_1 as deterministic functions of α (and keep δ independent of worker ability, which is a normalization). Specifically, we assume that $\lambda_0 = (1 - \mu)\lambda_0'$, $\delta = (1 - \mu)\delta'$, and $\lambda_1 = (1 - \mu)(1 - \delta')\lambda_1'$ with

$$\lambda_0' = \frac{\exp(\lambda_{00} + \lambda_{01}\alpha)}{[1 + \exp(\lambda_{00} + \lambda_{01}\alpha)]}, \quad \lambda_1' = \frac{\exp(\lambda_{10} + \lambda_{11}\alpha)}{[1 + \exp(\lambda_{10} + \lambda_{11}\alpha)]}, \quad \kappa = \frac{\exp(\kappa_0 + \kappa_1\alpha)}{[1 + \exp(\kappa_0 + \kappa_1\alpha)]},$$

where $(\lambda_{00}, \lambda_{01}, \lambda_{10}, \lambda_{11}, \kappa_0, \kappa_1)'$ are structural parameters to be estimated.

The parameters thus introduced, plus workers' bargaining power β, constitute 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.

5 Estimation Procedure

We estimate the structural model by indirect inference (Gouriéroux, Monfort, and Renault, 1993), separately for three different education groups: 7-11 years of education, 12-14 and 15-20.[13] The indirect inference estimator is a simulated method of moments procedure where the 'moments' to be matched may include parameters from reduced form econometric models, or possibly mis-specified, but easy-to-estimate econometric models

[13] See Altonji, Smith, and Vidangos (2009) for another recent application of indirect inference to a different dynamic wage and job mobility model.

'resembling' the structural model. These are referred to as auxiliary models. The econometrician then seeks the structural parameter vector that minimizes the distance between the auxiliary models as estimated on real data and the same auxiliary models estimated on simulated data. The statistical properties of the indirect inference estimator is worked out in Gouriéroux, Monfort, and Renault (1993), to which we refer the reader for details.

Details of the algorithm used to simulate the structural model are given in Appendix C. It is worth noting that key steady state equilibrium outcomes, e.g. the structural wage equation, and the endogenous distributions of observables and unobservables, e.g. t, p and q, can be obtained analytically (see Appendix A). We can therefore impose equilibrium conditions directly in the simulation at each estimation step.

The choice of auxiliary models is a key and sometimes controversial step in indirect inference estimation. 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 four sets of moments.

5.1 Labor market mobility

We fit survivor functions for each type of spell and transition. Unemployment spells may end with a transition into a job, or be right-censored. Employment spells may end with a transition into unemployment, a transition into another job, or be right-censored. We account for competing risks and right-censoring by using Kaplan-Meier estimates of the survivor function for each spell type as follows. We stock-sample the Master Panel at given point in time, and record N residual durations. Then, taking job-to-job transitions as an example, let $\mathcal{P}_{EE}(\tau)$ be the set of spells at risk of ending in a job-to-job transition at duration τ. We have $|\mathcal{P}_{EE}(0)| = N$. Moreover, let $\mathcal{D}_{EE}(\tau)$ be the set of spells that do end in a job-to-job transition at duration τ. The Kaplan-Meier estimate of the survivor function relating to job-to-job transitions, denoted $\widehat{S}_{EE}(\tau)$, is given as

$$\widehat{S}_{EE}(\tau) = \prod_{s=1}^{\tau} \frac{|\mathcal{P}_{EE}(s)| - |\mathcal{D}_{EE}(s)|}{|\mathcal{P}_{EE}(s)|}. \tag{11}$$

Job-to-unemployment (EU) and unemployment-to-job (UE) survivor functions are obtained analogously. In the empirical implementation we stock-sample the Master Panel in the last week of November 1998, and include as moments to be matched $1 - \widehat{S}(12)$, $\widehat{S}(12) - \widehat{S}(24)$, and $\widehat{S}(24) - \widehat{S}(36)$ for each of the three types of transition we consider. For a given distribution of worker heterogeneity α, the labor market mobility moments identify the parameters in the transition probabilities λ_0, λ_1, δ, and κ. We argue below that the distribution of worker heterogeneity is identified from data on individual wages.

5.2 Mincer wage equations

The auxiliary Mincer wage regression is estimated on a panel of repeated annual (last-week-of-November) cross sections of employed workers extracted from the Master Panel. We only include spells that are not left censored (jobs starting after January 1st, 1985). Let i index individuals and let c index the annual cross sections. Let j index firms and let $\mathrm{J}(i,c)$ be the firm ID of worker i's employer in cross section c. The log wage regression we consider is

$$w_{ic} = \sum_{k=1}^{3} \xi_{1k} s_{ic}^k + \sum_{k=1}^{3} \xi_{2k} t_{ic}^k + \psi_i + \varphi_{\mathrm{J}(i,c)} + u_{ic}, \tag{12}$$

where w_{ic} is the log-wage, s_{ic} is job tenure, and t_{ic} is labor market experience. Tenure and experience enter via cubic polynomials. The parameters ψ_i and φ_j are unobserved time-invariant worker and firm effects and u_{ic} is the residual.

We estimate the parameters relating to tenure and experience by applying within-firm OLS to (12). Firm and worker effects are subsequently recovered from the resulting residuals in two steps (firm effects first, then worker effects). We normalize the empirical distribution of worker effects to have zero mean.

We include in the set of moments to match the estimated tenure and experience profiles, and the first four moments of the distributions of firm effects (employment-weighted), worker effects, and residuals. Finally, to further describe wage dynamics, we select sequences of consecutive within-job wage residuals (from spells containing at least five consecutive observations) and include residual autocovariances of order up to four over these observations. These moments convey information about the structural human capital accumulation function $g(\cdot)$ (via the wage-experience profile in (12)), the sampling distribution $F(\cdot)$ of firm types faced by job seekers (via the distribution of firm effects), the distribution of worker heterogeneity α, and the individual level productivity shock process ε.

5.3 Within-job wage growth

We further consider the autocorrelation structure of within-job wage growth, which is what the estimation of statistical models of earnings dynamics is typically based on (see e.g. Browning, Ejrnaes, and Alvarez 2010). For convenience, we condition the analysis on worker i staying in the same firm between experience levels t and $t+1$ and estimate the following auxiliary model for within-job wage growth:

$$\Delta w_{ic} = \zeta_1 + \zeta_2 \Delta t_{ic}^2 + \zeta_3 \Delta t_{ic}^3 + \Delta u_{ic}. \tag{13}$$

Neither equation (12) nor (13) has a structural interpretation: 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.

The auxiliary wage growth equation is estimated by OLS on the subsample of job spells with at least two consecutive annual wage observations. We include the estimated slope parameters, the standard deviation, skewness and kurtosis of the residuals, as well as the residual autocovariances up to the fourth order in the set of moments to match. These moments contributes to the identification of the structural human capital accumulation function $g(\cdot)$ and the individual level productivity shock process ε.

Comparison between the within-job wage dynamics described by the first-difference equation (13), and the equation in levels (12) further conveys information on the bargaining power parameter β. As will be established in Section 8, the bargaining power parameter β governs the magnitude of the response of wages to a job change — i.e., to a change in employer type p —, and to the receipt of an outside job offer not causing a job change — i.e., to a change in the random variable q in equation (7). Tenure effects and wage changes upon employer changes should thus convey information on β.

5.4 Firm-level value added

We finally include in the set of moments to match summary statistics of the distribution of value added, employment and mean wage across firms, weighing each firm observation by the inverse of the sampling probability. We also include the standard deviation of the growth rate of output per worker to identify the variance of the output shock z_{jc}.

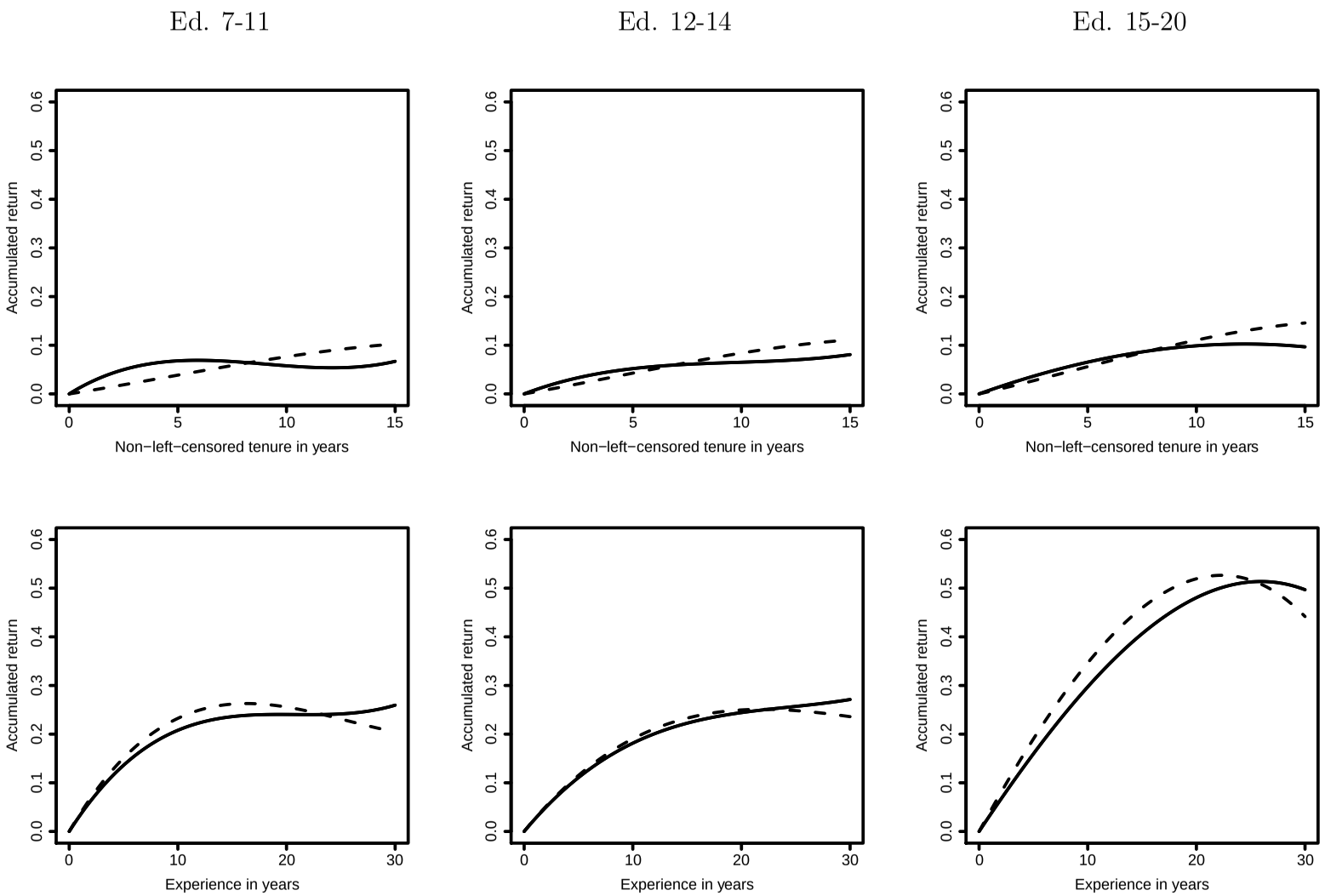
6 Model Fit

6.1 Labor market mobility

Table 2 reports annual transition probabilities based on (11). Our structural model replicates the observed job-to-job transition probabilities almost exactly, and also offers a reasonable fit to job-to-unemployment transition probabilities, especially for highly educated workers. It has some difficulty explaining the strong duration dependence in job-to-unemployment transition rates among low educated workers. Yet, duration dependence in unemployment-to-job transition rates is well captured by our model, although the fit is markedly better for groups of workers with low and high education levels.

6.2 Mincer wage equations

Figure 1 shows the experience and tenure profiles of individual wages as estimated from the Mincer-type auxiliary wage regression, equation (12). Solid lines depict profiles based on real data, while dashed lines relate to model-generated data. Finally, moments of the



Note: Solid line: real data; Dashed line: simulated data.

Figure 1: Cumulative wage-tenure and wage-experience profiles

Table 2: Fit of the survival functions: destination specific (simulated and real)

	Ed. 7-11		Ed. 12-14		Ed. 15-20	
	Sim.	Real	Sim.	Real	Sim.	Real
Unemployment-to-job transitions						
$1 - S_{UE}(12)$	0.3576	0.3735 (0.0097)	0.2810	0.5372 (0.0103)	0.3540	0.3759 (0.0249)
$S_{UE}(12) - S_{UE}(24)$	0.0665	0.0385 (0.0044)	0.1206	0.0506 (0.0050)	0.0518	0.0492 (0.0124)
$S_{UE}(24) - S_{UE}(36)$	0.0470	0.0176 (0.0032)	0.0763	0.0176 (0.0035)	0.0391	0.0241 (0.0097)
Job-to-job transitions						
$1 - S_{EE}(12)$	0.1346	0.1646 (0.0017)	0.1738	0.1787 (0.0010)	0.2121	0.2086 (0.0024)
$S_{EE}(12) - S_{EE}(24)$	0.1083	0.1137 (0.0016)	0.1284	0.1201 (0.0009)	0.1507	0.1319 (0.0020)
$S_{EE}(24) - S_{EE}(36)$	0.0866	0.0767 (0.0014)	0.0967	0.0905 (0.0008)	0.1098	0.1188 (0.0020)
Job-to-unemployment transitions						
$1 - S_{EU}(12)$	0.0518	0.0953 (0.0014)	0.0331	0.0544 (0.0006)	0.0251	0.0325 (0.0010)
$S_{EU}(12) - S_{EU}(24)$	0.0502	0.0470 (0.0011)	0.0330	0.0291 (0.0005)	0.0249	0.0211 (0.0010)
$S_{EU}(24) - S_{EU}(36)$	0.0464	0.0392 (0.0011)	0.0308	0.0250 (0.0005)	0.0234	0.0174 (0.0010)

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

firm and worker fixed effect distributions—again based on the auxiliary wage regression—are reported in Table 3.

Both experience and tenure profiles are correctly picked up by our structural model, albeit with a slight tendency to overstate the degree of concavity of those profiles. Wages are positively correlated with experience in all three subsamples (second row in Figure 1). These correlations are quantitatively rather modest for the low-educated group, and become more substantial at higher education levels. The auxiliary wage regression also indicates a moderate, yet positive correlation between wages and tenure in all three subsamples (second row in Figure 1). The accumulated 'returns to tenure' thus measured is about 5 percent after five years in a job in all education groups. Past five years of tenure, the wage-tenure profiles flatten out, especially among workers with low and intermediate education.

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 effect at zero in all samples.) Moreover, dispersion of worker- (and, to

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	4.9106	4.9012 (0.0020)	5.0095	4.9998 (0.0018)	5.1971	5.2224 (0.0040)
Std. dev.	0.1690	0.1609 (0.0006)	0.1614	0.1480 (0.0004)	0.1767	0.1788 (0.0010)
Worker effects						
Std. dev.	0.1166	0.1190 (0.0005)	0.1291	0.1362 (0.0003)	0.1678	0.1683 (0.0009)
Residuals						
Std. dev.	0.1325	0.1380 (0.0003)	0.1282	0.1344 (0.0002)	0.1403	0.1455 (0.0005)
Within-job residual autocovariance						
Order 1	0.0018	0.0023 (0.0000)	0.0015	0.0024 (0.0000)	0.0018	0.0028 (0.0001)
Order 2	−0.0009	−0.0001 (0.0000)	−0.0006	0.0001 (0.0000)	−0.0007	0.0001 (0.0000)
Order 3	−0.0013	−0.0013 (0.0000)	−0.0011	−0.0010 (0.0000)	−0.0013	−0.0012 (0.0000)
Order 4	−0.0012	−0.0016 (0.0000)	−0.0011	−0.0014 (0.0000)	−0.0013	−0.0016 (0.0000)

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 tenure and experience, and fit to 3rd and 4th moments of distributions are available on request.

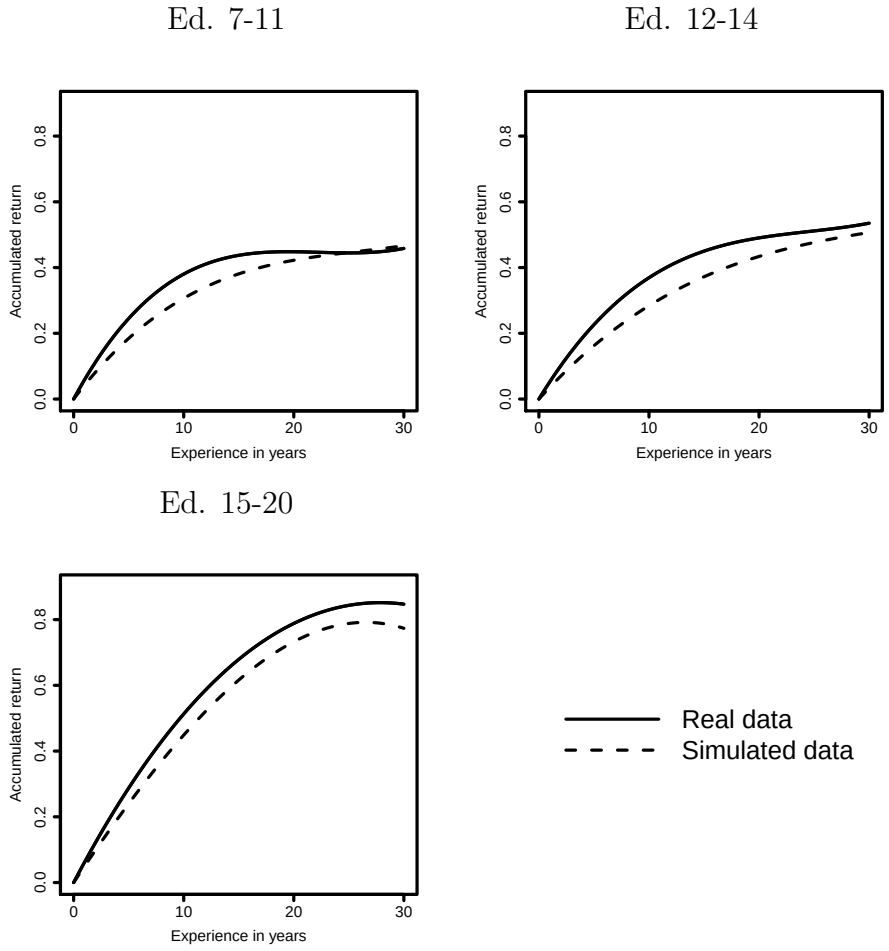


Figure 2: Cumulative experience profiles for wage growth

a smaller extent, firm-) effects tends to be slightly higher among more educated groups. Except for first-order residual wage autocovariances, which are underestimated, all numbers in Table 3 are accurately replicated by the structural model. We do not report or comment on the fit to 3rd and 4th-order moments of the distributions of firm and worker effects, and residuals (these moments are available on request).

6.3 Wage growth

Results from the auxiliary wage growth equation (13) 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. Table 4 further reports parameter estimates and moments of the residual distribution, including the autocovariance structure of wage growth residuals.

The profiles in Figure 2 combine the effects of tenure and experience within a job spell. As one would expect based on results from 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

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.	0.1229	0.1342 (0.0003)	0.1203	0.1271 (0.0002)	0.1232	0.1353 (0.0008)
Within-job residual autocovariance						
Order 1	−0.0041	−0.0045 (0.0000)	−0.0044	−0.0042 (0.0000)	−0.0043	−0.0037 (0.0001)
Order 2	−0.0016	−0.0008 (0.0000)	−0.0011	−0.0008 (0.0000)	−0.0013	−0.0008 (0.0000)
Order 3	−0.0004	−0.0006 (0.0000)	−0.0004	−0.0005 (0.0000)	−0.0005	−0.0006 (0.0000)
Order 4	−0.0003	−0.0006 (0.0000)	−0.0003	−0.0004 (0.0000)	−0.0004	−0.0005 (0.0000)

Note: Standard errors in parentheses; computed by bootstrapping the variance-covariance matrix of the real moments (10,000 replications). The estimated slope coefficients on experience, and fit to 3rd and 4th moments of distributions are available on request.

slightly underestimates those experience profiles, especially in the high education group.

In Table 4, the second-order moment of the distribution of residual wage growth is well captured by the structural model (the mean being normalized at zero). Residual autocovariances decline sharply between one and two lags, and are essentially zero at longer lags. As is typically found in studies of individual earnings dynamics based on individual or household data, this is suggestive of a low-order MA structure. Our structural model is once again able to replicate this feature of the data.

6.4 Firm-level value added

Results from the auxiliary equation (10) 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.

Overall, actual data exhibits a considerable amount of dispersion in average log labor productivity. 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 due to the fact that output per worker is calculated for the entire workforce, including both skilled and unskilled workers. It is thus an artifact of not observing the relevant productivity parameter p. Note that the

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	5.3484	5.3447 (0.0017)	5.3379	5.3461 (0.0010)	5.5527	5.5532 (0.0023)
Std. dev.	0.2963	0.3216 (0.0014)	0.3261	0.3478 (0.0008)	0.3636	0.4000 (0.0018)
Log wages, individual level (w)						
Mean	5.1770	5.1707 (0.0015)	5.2748	5.2809 (0.0008)	5.6518	5.6541 (0.0021)
Std. dev.	0.2734	0.2641 (0.0012)	0.2799	0.2574 (0.0006)	0.3496	0.3363 (0.0017)
Correlations and innovations						
Corr(y, w)	0.6103	0.2651 (0.0050)	0.6104	0.2631 (0.0028)	0.5521	0.2052 (0.0055)
Std. dev. within-job Δy	0.2246	0.2036 (0.0010)	0.2378	0.2124 (0.0006)	0.2614	0.2494 (0.0014)

Note: Standard errors in parentheses; computed by bootstrapping the variance-covariance matrix of the real moments (10,000 replications). Standard errors of Corr(y, w) were computed using the delta method. Fit to 3rd and 4th moments of distributions are available on request.

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. 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.

7 Structural Parameter Estimates

We now discuss structural parameter estimates, except for the monthly discount rate $\rho = 0.0050$ and attrition probability $\mu = 0.0018$, which were both fixed prior to estimation.

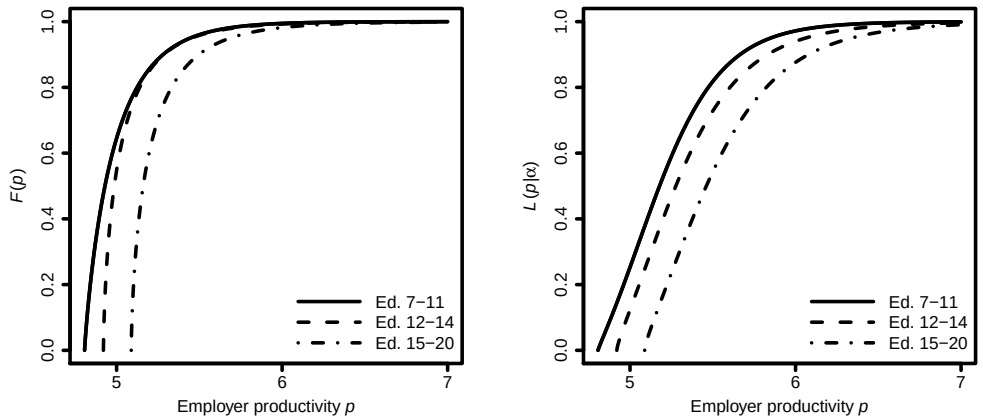
7.1 Worker Heterogeneity

Table 6 contains the estimated standard deviation of the distribution of the time-invariant component of worker heterogeneity, α, which we interpret as fixed innate worker ability. Within-group dispersion in ability is increasing from low to high educated workers. Interestingly, the structural model has a 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 equa-

Table 6: Structural parameter estimates: Wages and productivity

	Ed. 7-11	Ed. 12-14	Ed. 15-20
Worker type distribution $H(\alpha) = \mathcal{N}(0, \sigma_\alpha^2)$			
σ_α	0.0659 (0.0262)	0.0946 (0.0066)	0.1415 (0.0113)
Sampling distribution $F(p) = 1 - \exp\left(-\left[\nu_1(p - \nu_0)\right]^{\nu_2}\right)$			
$\nu_0 = p_{\min}$ (location)	4.8066 (0.0022)	4.9202 (0.0016)	5.0884 (0.0044)
ν_1 (scale)	5.3796 (0.0015)	8.9990 (0.0000)	8.1575 (0.0002)
ν_2 (shape)	0.8897 (0.0192)	0.7000 (0.0087)	0.6924 (0.0225)
Productivity shocks $\varepsilon_{it} = \eta\varepsilon_{it-1} + u_{it},\ u_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$			
η	0.7890 (0.0189)	0.7022 (0.0112)	0.7325 (0.0372)
σ_u	0.0829 (0.0099)	0.0992 (0.0086)	0.0934 (0.0250)
Workers' bargaining power			
β	0.3178 (0.0025)	0.2985 (0.0012)	0.2971 (0.0039)
Value added equation $y_{jt} = \chi_0 + \chi_1 p_j + z_{jt}, z_{jt} \sim \mathcal{N}(0, \sigma_p^2)$			
χ_0	0.8778 (0.0348)	0.4185 (0.0218)	0.4308 (0.0457)
χ_1	0.8551 (0.0067)	0.9223 (0.0041)	0.9256 (0.0084)
σ_p	0.1612 (0.0059)	0.1709 (0.0032)	0.1877 (0.0066)
Experience accumulation function $g(t) = \gamma_1 t + \gamma_2 t^2 + \gamma_3 t^3$			
γ_1	0.0241 (0.0003)	0.0136 (0.0003)	0.0268 (0.0006)
$\gamma_2 \times 10$	−0.0113 (0.0002)	−0.0039 (0.0002)	0.0011 (0.0004)
$\gamma_3 \times 1000$	0.0144 (0.0004)	0.0017 (0.0004)	−0.0216 (0.0009)

Note: Standard errors in parentheses. The monthly discount rate ρ and the monthly attrition rate μ are fixed at values of 0.0050 and 0.0018.



Note: The plot in the right panel is done for a worker of median ability α among employed workers.

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

tion captures the persistence generated by the $AR(1)$ idiosyncratic shock ε_{it}. A worker's ability also conditions his labor market transition probabilities, a set of parameters we consider later in Subsection 7.3.

Table 6 further reports estimates of the parameters of the monthly $AR(1)$ idiosyncratic productivity process, ε_{it}. The main feature of those estimates is that they are quite similar across all three education groups. We further note that the reported AR coefficients are based on a period length of one month, and translate into much smaller annual coefficients of 0.047, 0.040 and 0.042 for the low, medium and high-education groups, respectively.

7.2 Firm Heterogeneity

The estimates of the postulated reduced-form relationship (10) between observed firm-level value added and the underlying firm type p are reported in the fifth panel of Table 6. We note that, as expected, the estimated relationship between observed value added and firm type is increasing.

Estimates of the parameters of $F(\cdot)$ are reported in the second panel of Table 6. Perhaps more directly informative are the implied mean and variances of the relating sampling distributions. The mean sampled log-productivity is 5.00 for workers with 7-11 years of schooling, 5.06 for workers with 12-14 years of schooling and 5.25 for workers with 15-20 years of schooling (all in log terms). The corresponding standard deviations are 0.22, 0.20 and 0.23. 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. The same ranking appears to hold 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 $L(p|\alpha)$, evaluated at the median value of α among employed workers. $L(p|\alpha)$ is deduced from the estimated sampling distributions $F(p)$ and transition parameters μ, δ, and $\lambda_1(\alpha)$ using equation (9), that is $L(p|\alpha) = (\mu+\delta)F(p)/[\mu+\delta+\lambda_1(\alpha)\overline{F}(p)]$. The right panel of Figure 3 shows that the same FOSD-ordering holds for these cross-sectional distributions, thus confirming the presence of positive sorting by education.

7.3 Labor Market Mobility

The probability that a type-α worker receives a job offer while unemployed is $\lambda_0(\alpha)$. If employed, the job offer probability is $\lambda_1(\alpha)$. Jobs are destroyed at rate δ, independent of ability (a normalization). When a job is destroyed, a type-α worker immediately finds a substitute job with probability $\kappa(\alpha)$. Hence, the probability of a type-α employed worker entering unemployment is $\delta[1-\kappa(\alpha)]$. Figure 4 plots $\widehat{\lambda}_0(\alpha)$, $\widehat{\lambda}_1(\alpha)$, and $\widehat{\delta}[1-\widehat{\kappa}(\alpha)]$ as functions of $\alpha \in [-1.96 \times \widehat{\sigma}_\alpha, 1.96 \times \widehat{\sigma}_\alpha]$. We see from Figure 4 that the dependence of individual level transition probabilities on ability has the expected sign: more able workers have higher job offer arrival rates both as unemployed and as employed, and are less likely to experience unemployment. This is true for all three education groups. We discuss the implications for labor market sorting in the next section.

Consider the top-left panel in Figure 4 which plots the job offer arrival probabilities for unemployed workers. By assumption, this is also the job finding probability. All three education groups have a small fraction of workers with very low job finding probability. These workers are long-term unemployed. At the same time, a rather large fraction of the population of workers have a monthly job finding probability close to one. These workers find jobs almost instantly when they become unemployed.

The top-right panel in Figure 4 shows the offer arrival probabilities for employed workers. These are ordered across education groups as one would expect when we consider low-ability workers in each group: higher educated workers receive offers on-the-job more frequently than less educated workers. However, the slopes are steeper for less educated workers, so that high-ability workers tend to have similar on-the-job offer arrival rates, irrespective of education. Finally, the bottom-left panel in Figure 4 plots the job-to-unemployment transition probability. They mirror the on-the-job offer arrival rates, education and ability providing the expected insurance against unemployment.

Putting together results on worker heterogeneity, firm heterogeneity, and mobility, we see that the model generates labor market sorting among employed workers separately within each education group. More able workers receive offers more frequently and therefore, in steady state, will hold jobs at better employers than less able workers. Because $\lambda_1'(\alpha) > 0$ a strong form of sorting occurs within education groups: the distribution of

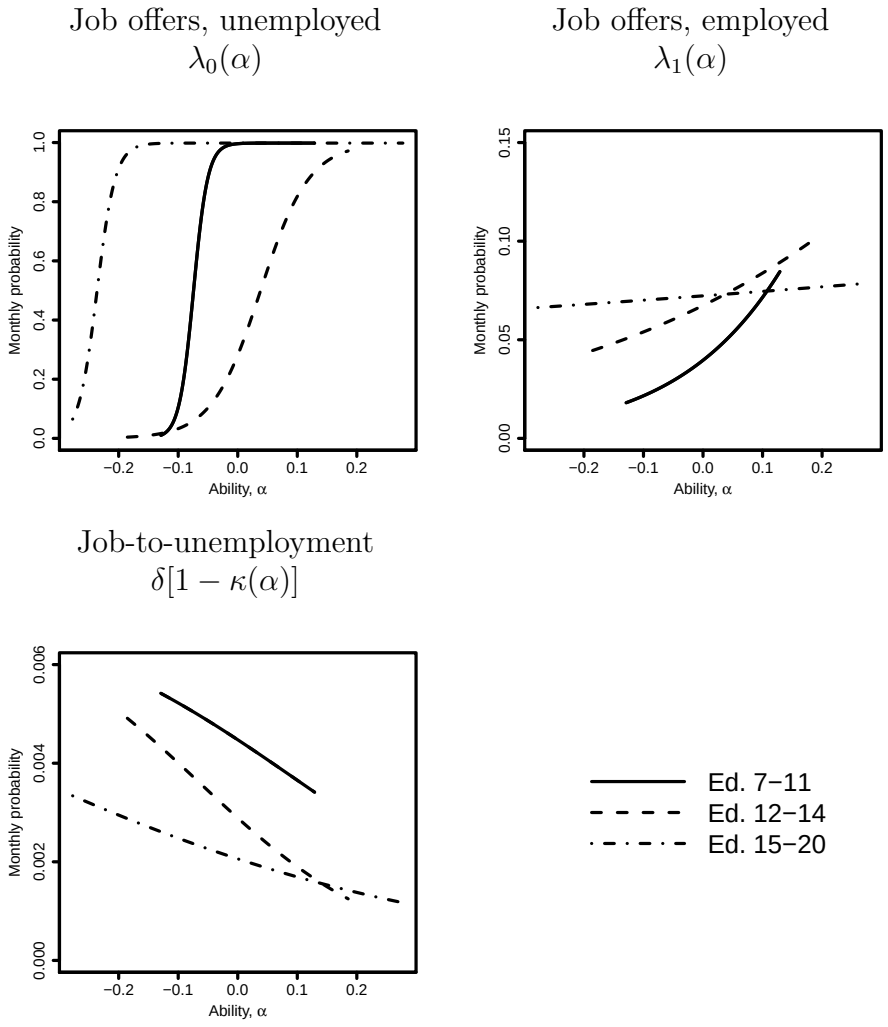


Figure 4: Labor market mobility parameters as a function of worker ability

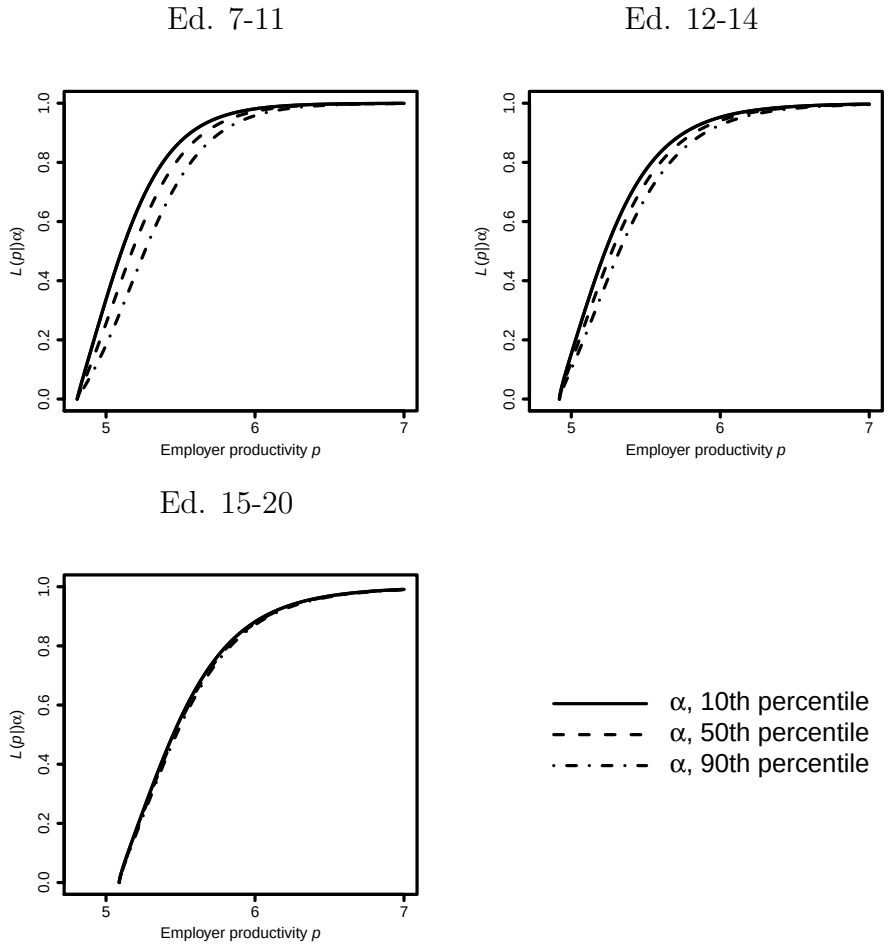


Figure 5: Sorting of workers across employers

employer types among worker of ability α, $L(p|\alpha)$, is increasing in α in the FOSD sense. We thus capture the spirit of the sorting mechanism introduced by Bagger and Lentz (2012), who estimate an equilibrium search model where sorting arises endogenously through workers' heterogeneous choices of search intensity depending on the properties of the match production function, worker types and employer types. In our simplified specification, search intensity is exogenous and is only allowed to depend on worker type.

Figure 5 plots $L(p|\alpha)$ for α equal to the 10th, the 50th and the 90th percentile in the population distribution of α, for each of the three education groups, and suggests that sorting, as defined above, is a feature of the labor market for workers with low and intermediate education, but does not appear in the labor market for highly educated workers. This of course is just another manifestation of the fact, already apparent in Figure 4, that job offer arrival rates are not strongly type-dependent among highly educated workers. From Figure 5, it appears that within-education group sorting is in fact strongest among workers with the lowest education. Interestingly, our estimated model suggests that within-education group heterogeneity (as measured by σ_α) among workers and employers is increasing in education. Nonetheless, this increased heterogeneity does

not translate into increased labor market sorting.

7.4 Worker Bargaining Power

Our estimates of worker bargaining power β are reported in the fourth panel of Table 6 above, and suggest that workers from all three education group have virtually equal bargaining power, around $\beta \simeq 0.3$.

Using Danish data, Bagger, Christensen, and Mortensen (2012) present industry- and occupation specific estimates of workers' bargaining power parameter in a bargaining model in the spirit of Stole and Zwiebel (1996). They find that the bargaining power parameter is difficult to pin down precisely, but the range of estimates they present are broadly consistent with the estimates reported in Table 6. Closer to us, Cahuc, Postel-Vinay, and Robin (2006) estimate a model similar to ours with no experience accumulation on French data. They find that workers in low skilled occupations have virtually no bargaining power with bargaining power increasing from less to more skilled occupations. While there are a number of differences between the Cahuc, Postel-Vinay, and Robin (2006) paper and this one (our use of Danish rather than French data, our stratification on education rather than occupations, our different way of including firm-level output data in the estimation), to which the discrepancy in bargaining power estimates may be partly attributed, the most important substantive innovation of this paper is the inclusion of human capital accumulation into the model. Our results suggest that, once we allow for human capital accumulation with different profiles between education groups, the model does not need to resort to exogenous differences in bargaining power to accommodate differences in wage profiles across education groups.

Even though we find that all education groups have roughly equal bargaining power, we should note that a worker's steady state share of match output, i.e. the piece rate, depends on a number of structural parameters in addition to β, most notably the probability that the worker obtains an outside offer $\lambda_1(\alpha)$ (see (6)). Since $\lambda_1(\alpha)$ is higher on average for the intermediate and high education groups, this may offset the lower β for these groups. Indeed, our estimates imply that, for the median α in the population, the range (depending on experience t) of steady-state shares of output are 0.81-0.83, 0.82-0.85, and 0.81-0.83 for low, medium and high educated workers, respectively.

7.5 Human Capital Accumulation

Table 6 finally reports estimates of the deterministic trend in individual human capital accumulation, $g(t)$. For added legibility, those trends are also plotted in Figure 6. There are qualitative similarities between education categories in human capital accumulation patterns. For all education categories, the pace of human capital accumulation is fastest

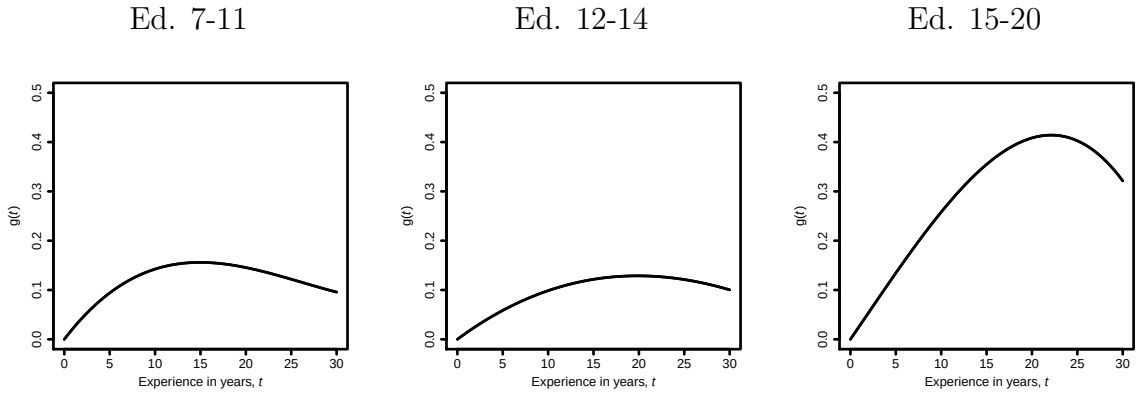


Figure 6: Structural human capital-experience profile ($g(t)$)

in the first ten years of a career, after which it 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 15 years, raising their productivity by a total of 15 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. At 30 years of experience, cumulated productivity growth for low-educated workers stands at 10 percent. At the other extreme, workers with more than 15 years of schooling grow about 40 percent more productive over the first 25 years of their careers. The human capital profile then declines for these highly educated workers toward the end of their working lives. At 30 years of experience, the accumulated productivity growth is still up to around 33 percent. The experience accumulation of workers in the intermediate education group is similar to that of low educated workers. In the next section we look at the implications of these productivity profiles for post-schooling wage growth.

8 Wage Profile Decomposition

8.1 Wage Growth

Workers accumulate human capital with experience. At the same time, job-to-job transitions tend to gradually reallocate workers to better jobs as they grow more experienced (in spite of reallocation shocks occasionally causing involuntary job mobility from high to low productivity jobs). Finally, wages also increase within job spells due to contract renegotiations prompted by outside offers.

Making use of the wage equation (7) and the characterization of wage dynamics in (8) we can calculate period-to-period wage growth for each one of the five regimes of

equation (8) as follows. 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:

$$\mathbb{E}(\Delta h_{i,t+1}|t, p_{it}, q_{it}) = \frac{g(t+1) - g(t)}{1 - \mu - \delta(1-\kappa)}, \tag{14}$$

  which is a deterministic function of experience.

- Within-job spell wage mobility (always upward):

$$\frac{\lambda_1}{1 - \mu - \delta(1-\kappa)} \int_{q_{it}}^{p_{it}} [F(p_{it}) - F(x)]\phi(x)dx. \tag{15}$$

  Employers give their workers wage increases to keep them from accepting outside job offers, thus creating positive apparent returns to tenure.

- Voluntary between-job wage mobility:

$$\frac{\lambda_1}{1 - \mu - \delta(1-\kappa)} \left[\int_{p_{it}}^{p_{\max}} \overline{F}(x)[1 - \phi(x)]dx + \overline{F}(p_{it}) \int_{q_{it}}^{p_{it}} \phi(x)dx \right], \tag{16}$$

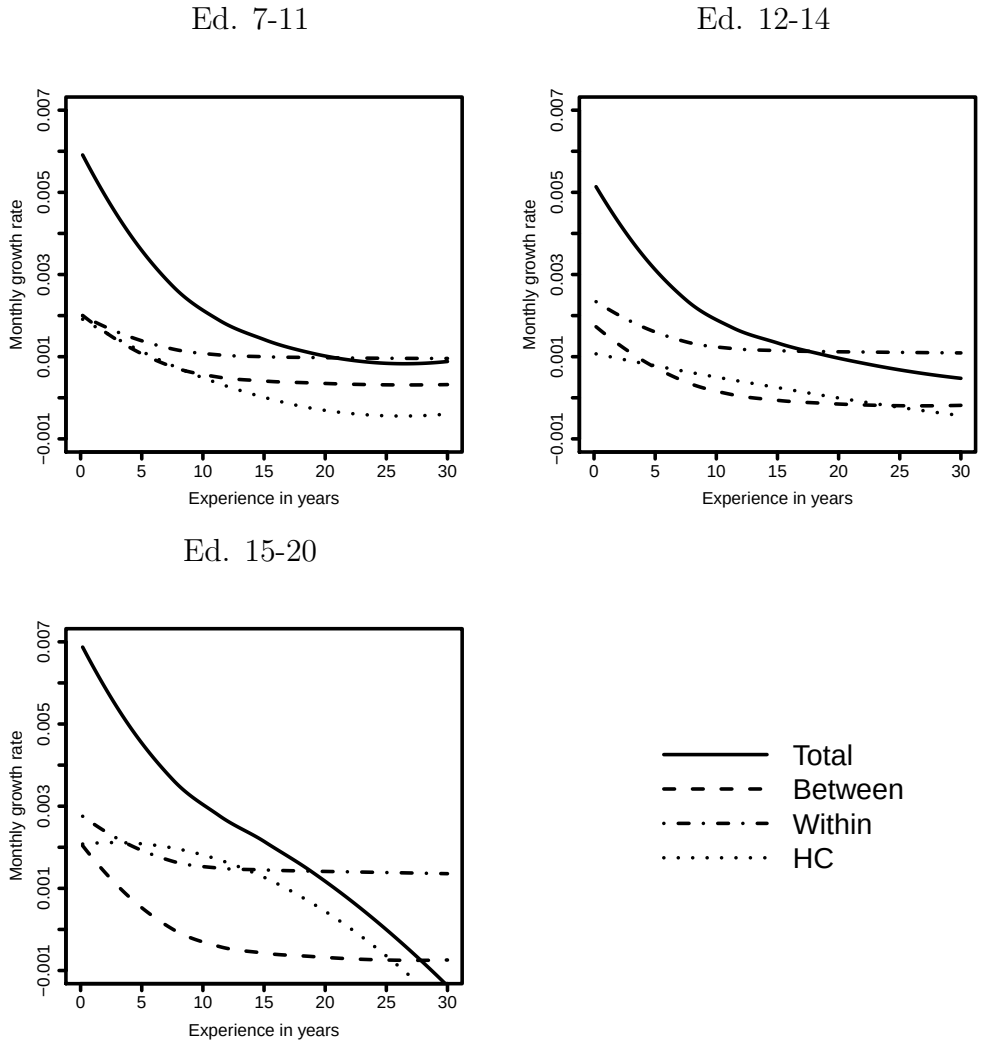
  where $\phi(x)$ was defined in equation (6). This is the expected instantaneous wage change following a *voluntary* job mobility. The negative component in the first integral, apparent in $1 - \phi(x)$, reflects the fact that workers are willing to give up a share of the surplus now in exchange for higher future earnings when moving from a less to a more productive employer. Yet, on average, this negative effect is cancelled out and job-to-job mobility is on average associated with a wage increase.

- Involuntary between-job wage mobility:

$$\frac{\delta\kappa}{1 - \mu - \delta(1-\kappa)} \left[\mathbb{E}p - p_{it} - \int_{p_{\min}}^{p_{\max}} \overline{F}(x)\phi(x)dx + \int_{q_{it}}^{p_{it}} \phi(x)dx \right]. \tag{17}$$

  This is the expected instantaneous wage change following an *involuntary* job mobility. In contrast to voluntary job changes, the worker's outside option is unemployment in this case, making wage cuts more frequent.

Finally, the conditioning variables q_{it} and p_{it} can be integrated out using the equilibrium distributions derived in Appendix A and given in equations (9) above. 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, two terms reflecting the impact of interfirm competition for workers, both within and between job spells, and a



Note: The plots are done for a worker of median ability $\alpha = 0$.

Figure 7: Sampling (left panel) and employer (right panel) distributions

term reflecting the impact of reallocation shocks. Search frictions and wage renegotiation generate wage/tenure profiles independently of human capital accumulation as employers raise wages in response to outside job offers as workers receive them.

Figure 7 gives a graphical rendering of that structural decomposition as a function of work experience. Figure 7 also plots total wage growth $\mathbb{E}\left(\Delta w_{i,t+1}|t\right)$. All plots are done for workers with median ability $\alpha_i = 0$. The solid/dashed/dash-dotted lines represent the contributions of between- and within-job wage mobility due to between-employer competition for workers, and human capital accumulation, respectively. The between-job component includes contributions from voluntary and involuntary job mobility.

Experience profiles of all three structural components of wage growth have similar qualitative shapes across education groups, and show that the observed concavity of wage-experience profiles results from the combination of a rate of human capital accumulation that declines with experience, and a concave impact of movements up and down the job

ladder. The latter reflects the fact that inexperienced workers tend to start at the bottom of the job ladder, thus facing a relatively high probability of sampling better jobs, whereas more experienced workers are likely to be nearer the top and have very little chance of drawing offers from even more productive firms.

The contribution of human capital accumulation, reflecting our estimates of $g(t)$ (Figure 6), declines more or less steadily throughout the working life in all education groups. It is largest for highly educated workers and decreasing with experience for all three education groups (and becomes negative past 15-20 years of experience).

As for the contribution of job search, both the within- and between-job spell components also decline with experience, although most of the decline occurs within the first ten years of a career. This initial ten-year period can be interpreted as the 'job-shopping' phase of a working life, during which workers wade their way up the job ladder, after which they settle in a high-quality job and use outside offers to generate gradual wage increases, reaping the benefits from competition between employers. The within-job component (reflecting contract renegotiations within a job spell) always dominates the between-job component, especially so toward the end of a career. The between-job component eventually becomes negative for highly educated workers due to involuntary job mobility.

8.2 Cumulative Wage-Experience Profiles

To offer a slightly different perspective on career wage dynamics, we can further use the structural model for an additive statistical decomposition of wage-experience profiles, as follows. The simple additive structure of our structural wage equation (7) implies that:

$$\mathbb{E}\left(w_{it}|t\right) = \mathbb{E}\left(\alpha_i|t\right) + g(t) + \mathbb{E}\left(p_{it}|t\right) + \mathbb{E}\left(r_{it}|t\right). \tag{18}$$

Note that mean worker ability among employed workers depends on experience through non-random selection into employment (since the job finding and layoff rates depend on worker ability α), so that $\mathbb{E}\left(\alpha_i|t\right)$ is, a priori, not constant across experience levels. For notational convenience, we further define $w(t) = \mathbb{E}\left(w_{it}|t\right)$, $\alpha(t) = \mathbb{E}\left(\alpha_i|t\right)$, $p(t) = \mathbb{E}\left(p_{it}|t\right)$, and $r(t) = \mathbb{E}\left(r_{it}|t\right)$. We can then use our estimated structural model to simulate all four components of (18): $w(t) = \alpha(t) + g(t) + p(t) + r(t)$.

This decomposition is interesting as a way of describing the cross-sectional wage distribution among workers of a given experience level. It is also in line with common practice in the literature (e.g. Topel, 1991, Altonji and Williams, 2005, Dustmann and Meghir, 2005, Altonji, Smith, and Vidangos, 2009). Altonji et al. propose what is perhaps the richest dynamic model of individual wages, mobility and hours worked estimated to date. It differs from our model in that it does not model labor market frictions and their consequences on equilibrium wage formation. Figure 1 in that paper decomposes expected

career wage growth into the sum of the effects of age (potential experience), job tenure, and the gains from job shopping as measured by a match-specific random component. They find that general human capital explains about 60% of wage growth while tenure and job shopping account for approximately equal shares of the remainder. Although the decomposition of Altonji et al. is similar to ours, it is based on a different model and different data, thus rendering further comparison with our results difficult.

We present results based on simulations of 100,000 career trajectories. From those simulations, it emerged that selection effects are quantitatively negligible, i.e. $\alpha(t)$ stays very close to zero at all experience levels. We thus omit that component of (18) in the analysis that follows. Figure 8 plots the remaining three components of (18), together with the overall wage-experience profile $w(t)$, all normalized at zero at one month of experience for readability. Specifically, the figure plots the average over 100,000 simulated trajectories of $w(t) - w(1)$, $g(t) - g(1)$, $p(t) - p(1)$, and $r(t) - r(1)$, as functions of t.

It is immediately striking from Figure 8 that the average piece rate in a cross-section of employed workers of given experience, $r(t)$, is almost constant across experience levels. In other words, the wage-experience profile $w(t)$ is almost entirely explained by the combination of human capital accumulation, and an increasing mean employer type, $p(t)$. On average, more experienced workers earn the same piece rate as younger workers, but are employed at higher-productivity firms than younger workers: they receive an equal share of a higher output, and thus earn higher wages. As for the role of human capital, Figure 8 confirms that it is most important in explaining the wage-experience profiles of the high-education group, where $g(t)$ accounts for about a half of $w(t)$, roughly on a par with $p(t)$. The weight of human capital is smaller in the lower two education groups, where $g(t)$ accounts for a fifth to a quarter of $w(t)$.

The structural decomposition in Figure 7 revealed that within-job search quantitatively dominates between-job search in generating wage growth. However, the decomposition in Figure 8 shows that job shopping, i.e. $p(t)$, is quantitatively more important than the dynamics in piece rate contracts, $r(t)$, in generating wage growth over the life cycle. This reflects the fact that changes to p are cumulative whereas changes to r are not: within a job, the piece rate increases stochastically with experience, but between-jobs, the piece rate may fall as a result of a transition from a less to a more productive employer. Those two effects wash out, so that piece rates end up being uncorrelated with experience.

The structural wage equation (7) further allows for a decomposition of the cross-sectional variance of wages, conditional on experience:

$$\begin{aligned}\mathbb{V}(w_{it}|t) = \mathbb{V}(\varepsilon_{it}|t) + \mathbb{V}(p_{it}|t) + \mathbb{V}(r_{it}|t) + 2\mathrm{Cov}(p_{it}, r_{it}|t) \\ + \mathbb{V}(\alpha_i|t) + 2\mathrm{Cov}(p_{it}, \alpha_i|t) + 2\mathrm{Cov}(r_{it}, \alpha_i|t). \quad (19)\end{aligned}$$

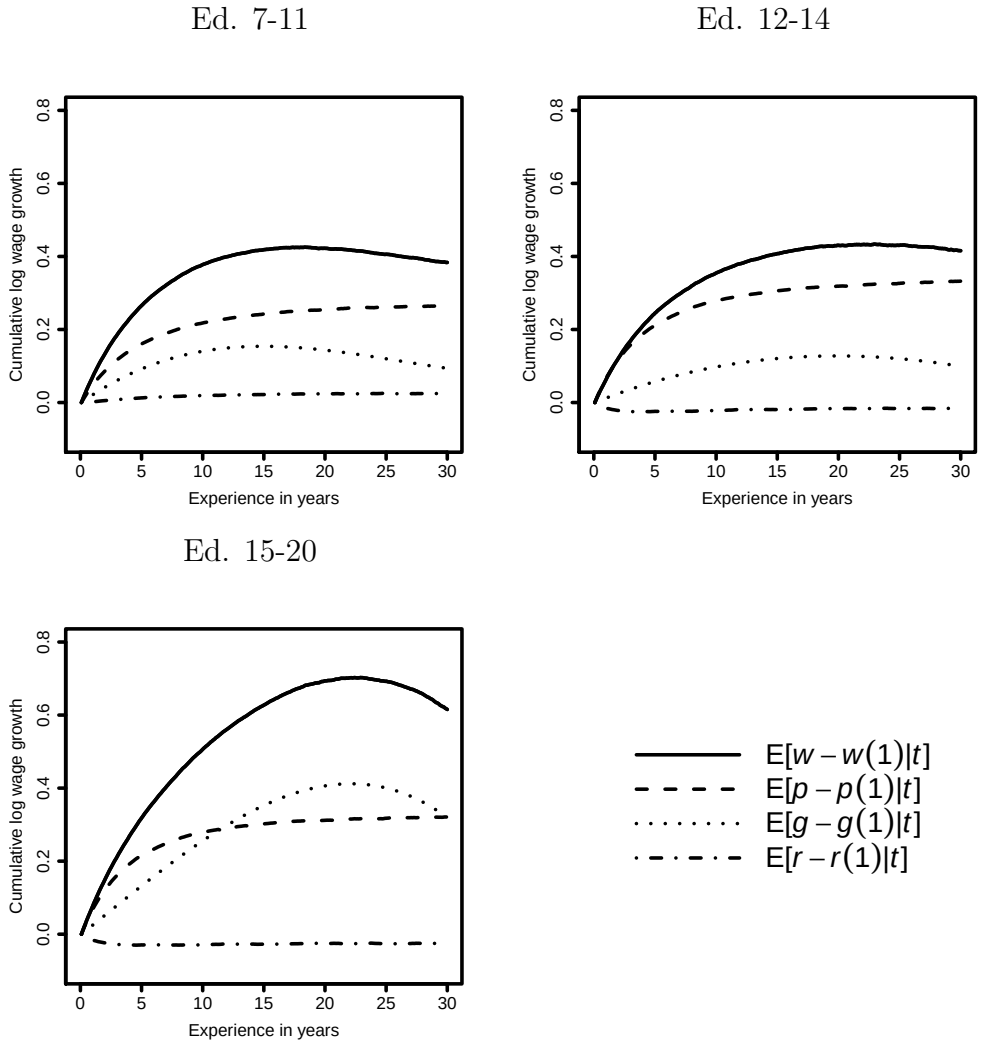


Figure 8: Log-wage component profiles: conditional mean

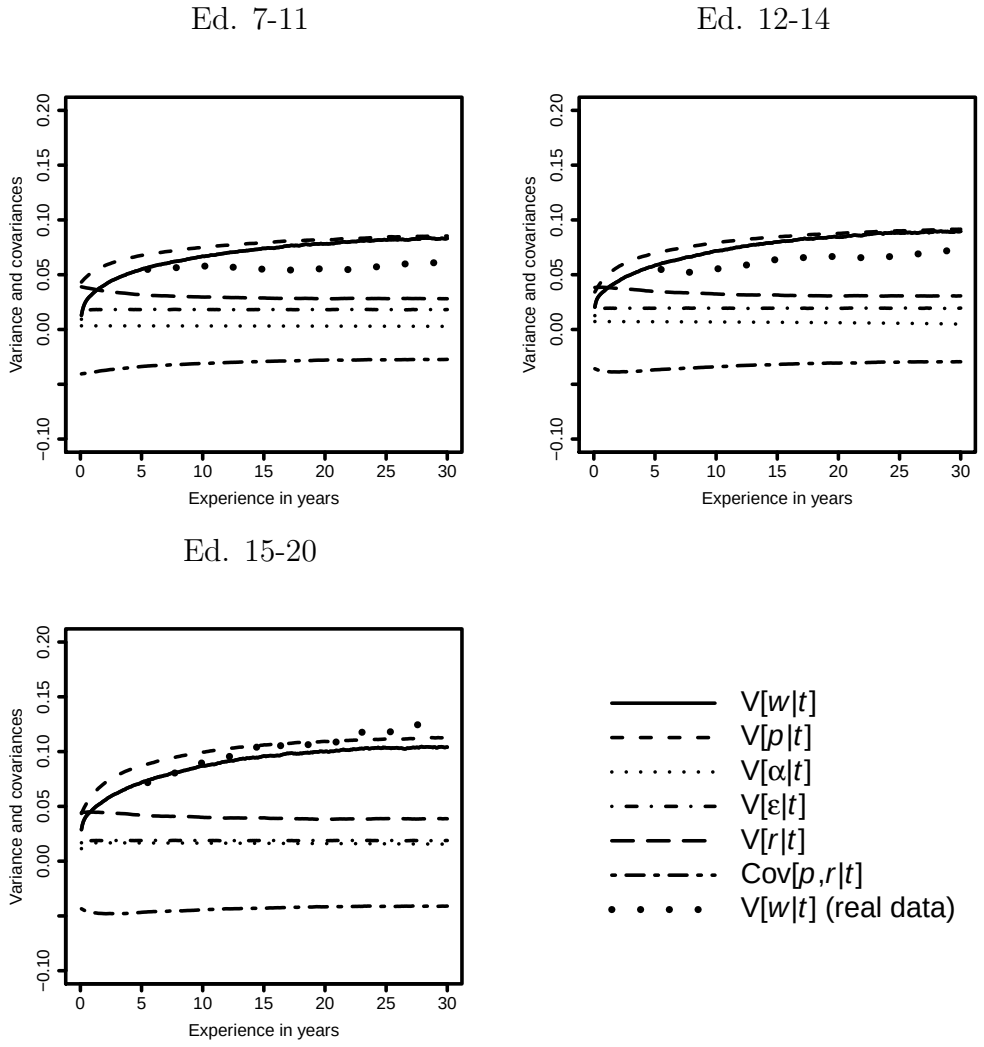


Figure 9: Log-wage component profiles: conditional variance

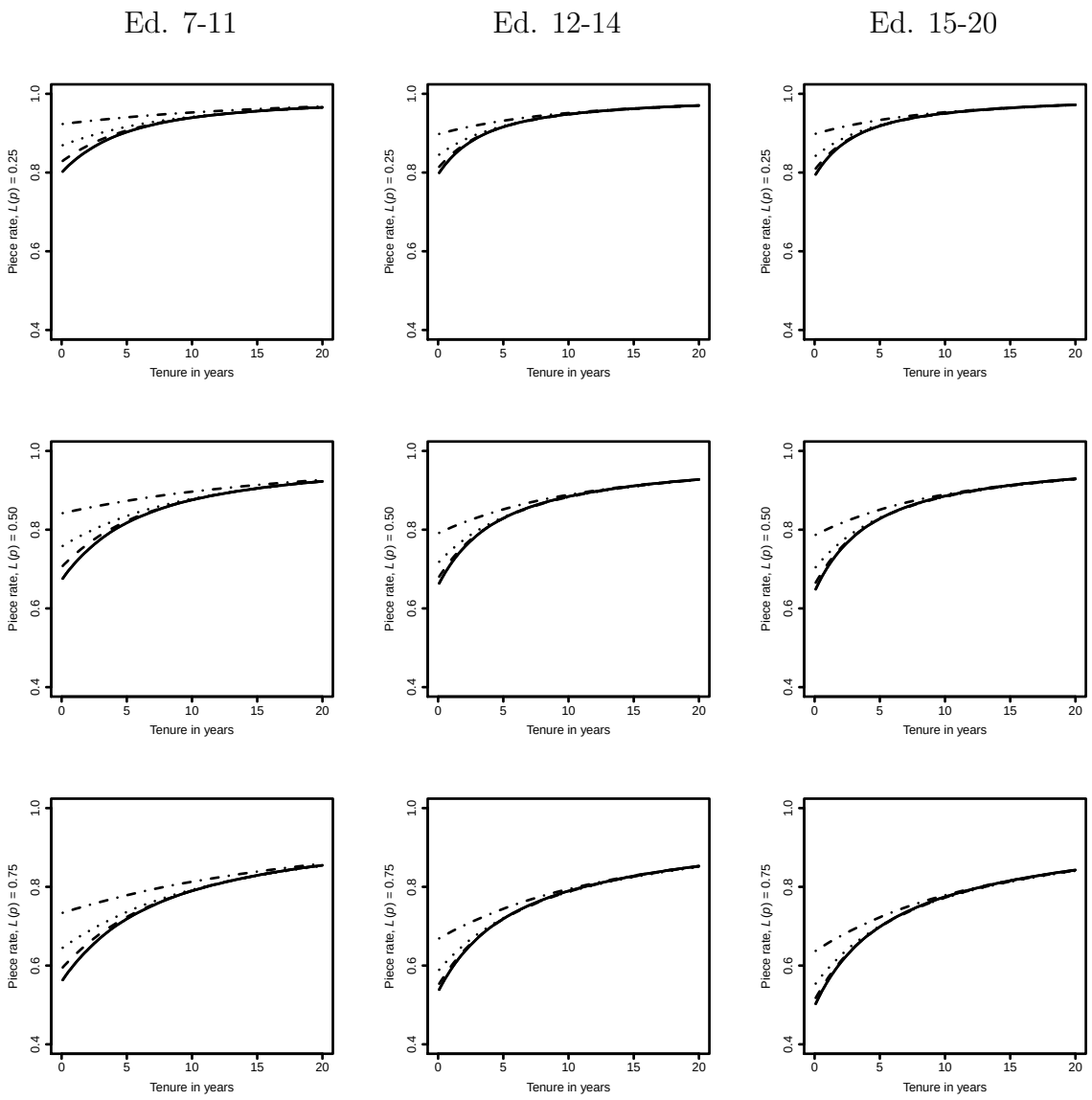
The variance and covariance terms involving the fixed worker heterogeneity term α_i reflect non-random selection of worker types into employment. Simulations again reveal that those selection effects are very small: $\mathbb{V}(\alpha_i|t)$ stays roughly constant, and $\text{Cov}(p_{it}, \alpha_i|t)$ and $\text{Cov}(r_{it}, \alpha_i|t)$ are both very close to zero across experience levels. To avoid cluttering the graphs, we omit the latter two covariance terms from the following analysis.

Figure 9 then plots the remaining five variance components in (19) plus the overall conditional variance. The figure also shows a plot of a non-parametric regression of log wage variance on experience, constructed directly from the raw data. Comparison of the model-predicted conditional log wage variance with the non-parametric regression line (dotted) shows that the model captures cross-sectional wage dispersion very well at all levels of experience for the group of highly educated workers, and has a slight tendency to overstate wage dispersion at high experience levels in the lower two education groups. This is an encouraging validation of the model, as those conditional variances were not among the moments fitted for estimation. Figure 9 then reveals that cross sectional log-wage variance is increasing with labor market experience at a declining rate. Interestingly, this phenomenon is driven almost exclusively by increased dispersion in employer productivity, i.e. by $\mathbb{V}(p|t)$.

Figure 9 further indicates that the variance decomposition is similar across education groups. There is considerable dispersion in employer types and substantial dispersion in piece rates among a cross section of employed workers of a given experience. Both of those components have sizeable positive contributions to cross-sectional wage dispersion. Those positive contributions are partially offset by the negative covariance between employer type and piece rate, reflecting the fact that workers are prepared to accept lower piece rates to work at more productive employers. The contribution of individual productivity shocks ε_{it} to wage dispersion is comparatively modest, as is the contribution of ability α. Finally, it is interesting to observe from Figures 8 and 9 that the conditional variance of piece rates, like their conditional mean, does not vary much with experience (it declines very slightly).

9 Returns to Tenure

Positive "returns to tenure" arise in our structural model because piece rates are gradually revised upward within a job spell as workers receive 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 7. This average profile, however, conceals 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



Note: The plots are done for a worker of median ability $\alpha = 0$. Each column plots a different education group. The solid line is $q = p_{\min}$ at initiation, the dashed line is q equal to the 1st quartile in $G(q|p)$, the dotted line is q equal to the 2nd quartile in $G(q|p)$, and the dash-dotted line is q equal to the 3nd quartile in $G(q|p)$.

Figure 10: Piece rate profiles

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 benefiting 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 the 1st, 2nd and 3rd quartiles of the $L(p)$ distribution as our set of firm types p. Then, for each of those p, we consider four different piece rates corresponding to $q = p_{\min}$, and q equal to the 1st, 2nd and 3rd quartiles of $G(q|p)$, the distribution of renegotiation thresholds within a type-p firm ($q = p_{\min}$ yields the piece rate obtained by workers just hired from unemployment). For each of those (p, q) pairs we then simulate career trajectories for $100,000$ workers over 20 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 10.

As expected, more productive firms tend to offer lower starting piece rates and steeper subsequent tenure profiles, with little differences across 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 about 10 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).

10 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 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 decomposition of individual wage growth are qualitatively similar across education groups, but reveals that more educated workers have higher total wage growth. This reflects both more rapid human capital accumulation and, at the early stage of worker's

career, higher returns to job search. We also find that both human capital accumulation and job search contribute to the concavity of wage-experience profiles. The contribution from job search to wage growth, both within- and between-job, declines within the first ten years of a career, a period that we identify as the 'job-shopping' phase of a working life. Workers subsequently settle into high-quality jobs and use outside offers to generate gradual wage increases, thus reaping the benefits from competition between employers. Indeed, the within-job component always dominates the between-job component, but especially so after ten years of labor market experience.

We supplement the structural decomposition of monthly wage growth by decomposing the mean and variance of log wages, conditional on experience level, into components of our structural wage equation: worker ability, employer productivity, human capital, invidiual level producticity shocks, and the piece rate contract. The wage-experience profile is almost entirely explained by human capital accumulation and an increasing mean employer productivity (due to 'job shopping'). Indeed, among highly educated workers, human capital accounts for about half of the accumulated growth at all experience levels. The weight of human capital is smaller among workers with medium or low education (a fifth to a quarter). Cross sectional log-wage variance is increasing with labor market experience at a declining rate (in the data and in the model). This is almost exclusively driven by increased dispersion in employer productivity. The level of log-wage variance is explained by dispersion in employer productivity and in piece rate contracts. Dispersion in worker ability and individual productivity shocks contribute only little to cross sectional log-wage dispersion.

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.

Altonji, J. G., A. Smith, and I. Vidangos (2009): "Modeling Earnings Dynamics," NBER Working Papers 14743, National Bureau of Economic Research, Inc.

Altonji, J. G., and N. Williams (2005): "Do wages rise with job seniority? A reassessment," *Industrial and Labor Relations Review*, 58(3), 370–397.

Bagger, J., B. J. Christensen, and D. T. Mortensen (2012): "Wage and Labor Productivity Dispersion: The Roles of Total Factor Productivity, Labor Quality, Capital Intensity, and Rent Sharing," Discussion paper, Royal Holloway, University of London.

Bagger, J., and R. Lentz (2012): "An Equilibrium Model of Wage Dispersion and Sorting," Discussion paper, Royal Holloway, University of London.

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.

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.

Stole, L. A., and J. Zwiebel (1996): "Intra-Firm Bargaining under Non-Binding Contracts," *The Review of Economic Studies*, 63, 375–410.

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.

Veramendi, G. (2011): "Labor Market Dynamics: A Model of Search and Human Capital Accumulation," Discussion paper, Arizona State University.

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 (5) and integrate by parts on the r.h.s. to obtain:

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

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:

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

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

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

A.2 Derivation of the mobility piece rate

Substitution of (A4) into (3) 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 (3) 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).

Substitution of (5) into (3) produces the following implicit definition of $q(\cdot)$:

$$\begin{aligned} r = -(1-\beta)\left[p - q(r,h_t,p)\right] &- \int_{q(r,h_t,p)}^{p} \frac{\lambda_1(1-\beta)^2\overline{F}(x)\,dx}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)} \\ &- \int\int_{q(r,h_t,p)}^{q(r,h_{t+1},p)} \frac{(1-\mu-\delta)(1-\beta)}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)}\,dx\,dM(h_{t+1}|h_t) \end{aligned} \quad \text{(A5)}$$

where $M(\cdot|h_t)$ 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 (A5) has a simple, deterministic (indeed constant), consistent solution $q(r,p)$ implicitly defined by:

$$r = -(1-\beta)\left[p - q(r,p)\right] - \int_{q(r,p)}^{p} \frac{\lambda_1(1-\beta)^2\overline{F}(x)\,dx}{\rho+\delta+\mu+\lambda_1\beta\overline{F}(x)}. \quad \text{(A6)}$$

Now even though (A5) implies no direct dependence of $q(\cdot)$ on t and h_t, other, nondeterministic solutions to (A5) may still exist. We ignore the possibility of more sophisticated expectational mechanisms in this paper, and concentrate on the deterministic solution (A6).

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 (7), 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:

$$u = \frac{\mu + \delta(1-\kappa)}{\mu + \delta(1-\kappa) + \lambda_0}. \tag{A7}$$

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

$$(\lambda_0 + \mu) u a_0(t) = \delta(1-\kappa)(1-u)a_1(t) \tag{A8}$$

$$(1-u)a_1(t) = [1 - \mu - \delta(1-\kappa)](1-u)a_1(t-1) + \lambda_0 u a_0(t-1) \tag{A9}$$

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

$$(\mu + \lambda_0) u a_0(0) = \mu. \tag{A10}$$

Jointly solving those three equations, one obtains:

$$a_1(t) = \left[\mu + \frac{\mu\delta(1-\kappa)}{\mu + \lambda_0}\right]\left[1 - \mu - \frac{\mu\delta(1-\kappa)}{\mu + \lambda_0}\right]^{t-1}. \tag{A11}$$

The corresponding cdf is obtained by summation:

$$A_1(t) = \sum_{\tau=1}^{t} a_1(\tau) = 1 - \left[1 - \mu - \frac{\mu\delta(1-\kappa)}{\mu + \lambda_0}\right]^{t}. \tag{A12}$$

(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 (A8): $A_0(t) = \frac{\mu[\mu+\delta(1-\kappa)+\lambda_0]}{[\mu+\delta(1-\kappa)][\mu+\lambda_0]} + \frac{\delta(1-\kappa)\lambda_0}{[\mu+\delta(1-\kappa)][\mu+\lambda_0]} A_1(t)$.

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

$$L(p|t)a_1(t) = [1 - \mu - \delta - \lambda_1 \overline{F}(p)]L(p|t-1)a_1(t-1) + \left[\frac{\lambda_0\delta(1-\kappa)}{\mu + \lambda_0} + \delta\kappa\right] F(p)a_1(t-1). \tag{A13}$$

Since (A11) implies:

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

one can rewrite (A13) as $L(p|t) = \Lambda_1(p)L(p|t-1) + \Lambda_2 F(p)$, with:

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

This last equation solves as:

$$L(p|t) = \left[\Lambda_1(p)^{t-1} + \Lambda_2\frac{1-\Lambda_1(p)^{t-1}}{1-\Lambda_1(p)}\right]F(p). \tag{A14}$$

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

$$L(p) = \frac{(\mu+\delta)F(p)}{\mu+\delta+\lambda_1\overline{F}(p)}. \tag{A15}$$

Conditional distribution of piece rates. Equation (6) states that piece rates are of the form $r = r(q,p)$. Thus the conditional distribution of piece rates within a type-p firm is fully characterized by the distribution of threshold values q in a type-p firm, $G(q|p,t)$, which we now derive. For $t > 1$, the flow-balance equation determining $G(q|p,t)$ is given by:

$$\begin{aligned} G(q|p,t)\ell(p|t)a_1(t) = {} & [1-\mu-\delta-\lambda_1\overline{F}(q)]G(q|p,t-1)\ell(p|t-1)a_1(t-1) \\ & + \lambda_1 L(q|t-1)a_1(t-1)f(p) + \delta\kappa a_1(t-1)f(p) \\ & + [\mu+\delta(1-\kappa)]a_0(t-1)f(p), \end{aligned} \tag{A16}$$

where $\ell(p|t) = L'(p|t)$ is the conditional density of firm types in the population of employed workers corresponding to the cdf in (A14). Rewriting this last equation in the case $q = p$, so that $G(q|p,t) = 1$, yields the differential version of (A13):

$$\begin{aligned} \ell(p|t)a_1(t) = {} & (1-\mu-\delta-\lambda_1\overline{F}(p))\ell(p|t-1)a_1(t-1) \\ & + \lambda_1 L(p|t-1)a_1(t-1)f(p) + \delta\kappa a_1(t-1)f(p) \\ & + [\mu+\delta(1-\kappa)]a_0(t-1)f(p). \end{aligned} \tag{A17}$$

Dividing (A16) and (A17) 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:

$$G(q|p,t) = \frac{\ell(q|t)/f(q)}{\ell(p|t)/f(p)} \qquad \text{for} q \in [p_{\min}, p]\,,\ t > 1. \tag{A18}$$

The unconditional version, (A19), obtains by similar reasoning:

$$G(q|p) = \frac{\ell(q)/f(q)}{\ell(p)/f(p)} = \left[\frac{\mu+\delta+\lambda_1\overline{F}(p)}{\mu+\delta+\lambda_1\overline{F}(q)}\right]^2, \qquad \text{for } q \in [p_{\min}, p]. \tag{A19}$$

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.[14]

[14]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

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). Still we do observe firms with 0-4 employees and lump them together with firms with 5-9 employees.

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

- Annual value added/FTE observations are transformed into hourly measures by scaling annual value added/FTE by 12×166.33 hours[15] and the strata-specific distributions are trimmed by recoding the top and bottom 1% to missing. We re-adjust nominal variables (wages and value added) to the 2003 level using Statistic Denmark's CPI.

- The Master Panels contains data on firms' annual value added which we convert into an hourly measure. The value added data comes from a survey put together by Statistics Denmark using a known sampling scheme, which we must take into account by appropriately re-weighting moments of the value added distribution. We select the 1999-2003 cross sections and keep only observations on jobs with wage information. We split 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. Statistics Denmark's sampling scheme for the accounting data is such that, within each bin, a random sample of employers is selected to submit their accounting data. Rather than using the sampling probabilities used by Statistics Denmark, we compute from our sample the actual fractions of employees 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. The empirical sampling probabilities in our Master Panel are relatively close to the sampling probabilities applied by Statistics Denmark.

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.

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.

[15] 166.33 hours being the monthly norm for a full time job.

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 $\mathcal{N}(0, \sigma_\alpha^2)$. This fixes the α-dependent transition parameters λ_0, λ_1, and κ for each of the I simulated workers. Next, we assign labor market states (employed or unemployed) to workers according to (A7), 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 (A12). Given workers' labor market states and experience t we assign employer productivity. Unemployed workers are assigned productivity b independent of t while employed workers with experience t are assigned employer productivity p according to $L(p|t)$ defined by (A14). The productivities of the last firms from which the workers were able to extract the whole surplus in the offer matching game—the q's—are drawn (conditional on p and $t > 1$) from $G(q|p, t)$ defined by (A18). Unemployed workers and employed workers with experience $t = 1$ are assigned $q = b$. Finally, 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 $p_j = F^{-1}(j/(J+1))$.[16] 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 $p_{j(i,t)} = \arg\min_{x \in \{p_1, ..., p_J\}} |L(x|t) - u_i|$. Similarly, worker i's q is assigned (conditional on $p = p_j$ and $t > 1$) as $q_{it} = q_{it}(p_j) = \arg\min_{x \in \{p_1, ..., p_{j-1}\}} |G(x|p_j, t) - v_i|$, 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 (10)) 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), receive a job destruction shock, which may lead to an immediate job-to-job transition (probability $\delta\kappa$) or a transition into unemployment (probability $\delta(1 - \kappa)$), or leave the sample (probability μ).[17] Each time an unemployed worker receives an offer, we record a change of status, the productivity of the new employer[18] (p'), an increase in experience and

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

[17] Recall that λ_0, λ_1, and κ are functions of worker ability α.

[18] 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

we set the worker's duration of stay in his current spell to one. Finally, we set $q(\cdot) = p_{\min}$. When an employed worker (with employer type p) receives an offer, this results in a job-to-job transition if $p' > p$, in which case we record the productivity p' of the new employer, set $q(\cdot) = p$, the worker's tenure at the new firm to one and increment the worker's experience. In case $q(\cdot) < p' \leq p$, the worker does not change firms. However we need to update the worker's productivity threshold $q(\cdot)$ to p', and also increment the worker's tenure and experience.[19] 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 (7) 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).

state distributions.

[19] If the worker makes a job-to-job transition immediately following a job destruction shock, we record the productivity p' of the new employer and set $q(\cdot) = p_{\min}$. Otherwise, the transition is treated as a voluntary job-to-job transition.