Early Career Wage Profiles and Mobility Premiums

Jesper Bagger†
University of Aarhus

December 2006

Abstract

I study how early career wage trajectories are affected by job mobility using a sample of male labor market entrants. A long period of observation (up to 14 years) with detailed mobility data allows identification of the persistence of any initial mobility premium. The potential endogeneity of job mobility in relation to earnings is circumvented by explicitly modeling the processes governing transitions out of a job. The estimation procedure uses matched employer-employee data to control for unobserved worker specific and match specific effects. I find a sizeable, and rather persistent, mobility premium, although there are important differences across education groups. The analysis does not permit conclusions as to what causes the persistency, but it might result from employed workers having more bargaining power than unemployed workers, or from firms taking unemployment periods as signals of low productivity or low susceptibility of training.

Keywords: Mobility, Mobility Premium, Wage Profiles, Labor Market Entry, Job Search, Labor Market Transitions, Matched Employer-Employee Data.

JEL codes: C33, J33, J41, J63, J64

*I would like to thank Michael Rosholm, Peter Jensen, Jean-Marc Robin and Dimitris Mavridis for helpful comments and suggestions. Earlier versions of the paper were presented at the Economic Topics workshop at University of Aarhus, the GradInProWe workshop in Nyborg, 2004, the 7th IZA Summer School, the Xth Spring Meeting for Young Economists in Geneva and at Université de Paris 1. I thank the participants at these events for constructive comments and suggestions. Any remaining errors are mine.

†jbagger@econ.au.dk.
1 Introduction

The purpose of this paper is to study how early career earnings profiles relate to job mobility.\(^1\) Job search theory, as spelled out in e.g. Burdett (1978) and Burdett and Mortensen (1998), lays down that voluntary labor market mobility is triggered by workers' aspirations to higher wages. Indeed, previous studies have documented that job mobility in the early phases of workers' labor market careers are associated with a significant wage premium. Antel (1986) finds that a quit induces a wage gain of 20 percent using a sample of U.S. young men. Topel and Ward (1992) and Keith and McWilliams (1999), also using a sample of U.S. young men, report lower, but still substantial, returns to mobility of 10 percent and 8-11 percent, respectively. However, concurrent with the substantial returns to job mobility, workers also seem to have non-negligible within-job wage growth, a subject that is typically studied within the framework of the human capital theory, spelled out in e.g. Becker (1962). Topel (1991) finds that within-job wage growth for a U.S. worker with ten years of experience amounts to 7 percent in the first year in the job, and declines with job tenure. Topel and Ward (1992) report an average annual within-job wage growth rate among U.S. labor market entrants of about 14 percent (including aggregate wage growth).

In this paper I take these two empirical regularities together and study how within-job wage profiles in jobs \textit{initiated} via a job-to-job transition (henceforth \textit{jj}-jobs) differ from within-job wage profiles in jobs \textit{initiated} via a nonemployment-to-job transition (henceforth \textit{nj}-jobs).\(^2\) In other words, the focal point of this paper's analysis is the within-job persistence of any initial mobility premium (henceforth the \textit{jj}-premium). The closest related paper is Arulampalam (2001) who, using the British BHPS panel, find a wage penalty of about 6\% upon re-employment from a period of unemployment, which increases to 14\% over the first four years in the job, at which point it starts to vanish. The notions of \textit{premiums} associated with job-to-job transitions and \textit{penalties} associated with nonemployment-to-job transitions are of course flip-sides of the same coin. Compared to

---

\(^1\)The focus is kept on the earlier parts of workers' labor market careers for three reasons: First, the years following labor market entry are generally characterized by high job mobility and concomitant high wage growth, thus providing a good data source for the relationship between earnings and mobility. Second, the actions of labor market entrants are (arguably) less confounded with household decision issues; a subject that is not touched upon in this paper. Third, it conveniently rules out initial condition problems in the endogenous variables.

\(^2\)The terms “unemployment” and “nonemployment” will be used interchangeably.
Arulampalam’s study this paper focus exclusively on labor market entrants, use a longer (and different) panel with more precise information on individual labor market histories and apply a different estimation methodology which allows unobserved job specific effects to be correlated with e.g. tenure.

Bar the aforementioned study by Arulampalam, the issue of persistent mobility premiums has been largely overlooked in the empirical literature on individual earnings and mobility, but is potentially important for several reasons. First, it touches on the fundamental issue of wage determination. For example, it might be conjectured that workers quitting a job obtain better wage contracts than otherwise identical workers initiating their jobs from a period of unemployment due to the former groups better fallback option. This idea is pursued in Postel-Vinay and Robin (2002), who construct an equilibrium job search model, where firms engage in an offer-matching game for workers’ labor supply. According to this model, the starting wage in a job is higher for workers hired away from another firm vis-à-vis workers hired from unemployment, while the wage earned when leaving the job is unrelated to way the job was initiated. Hence, on average, workers coming from unemployment have steeper within-job wage profiles, and the initial $jj$-premium enjoyed by workers in $jj$-jobs thus vanishes as job tenure is accumulated.\(^3\) Second, if within-job wage growth is related to on-the-job training (Becker, 1962), the persistence of any initial $jj$-premium is informative on how unemployment periods influence the amount of training received in a subsequent job. On the one hand, firms might take unemployment periods as a signal of low susceptibility of training and consequently offer workers hired from the pool of unemployed workers less training, resulting in a highly persistent $jj$-premium.\(^4\) On the other hand, depreciation of human capital during unemployment periods implies that workers in $nj$-jobs \textit{ceteris paribus} have higher returns to training, causing firms to invest more heavily in these workers, and implying that wages in $nj$-jobs will catch up with wages in $jj$-jobs. Antel (1986) suggests that job search and (specific) training

\(^3\)If instead firms post wage contracts, Burdett and Coles (2003) have shown that the posted contracts differ only in starting points on an underlying baseline salary scale, and are not conditional on the workers’ previous labor market state. Hence, in the contract posting environment, once intrinsic heterogeneity has been adequately accounted for, there should be no initial $jj$-premium and no differences in wage profiles across $jj$-jobs and $nj$-jobs.

\(^4\)This is equivalent to arguing, as in Arulampalam (2001), that unemployment periods “scar” workers future earnings potential.
are mutually exclusive human capital investments. In this case, mobile workers are likely to have specialized in job search, and thus, should have flatter within-job wage trajectories.5

Ultimately, the persistence of any initial $jj$-premium, is an empirical question. To answer it requires an empirical model of earnings and mobility that is sufficiently flexible and rich in heterogeneity to isolate the structural dependence of wage growth on mobility from selection effects stemming from the endogeneity of mobility in relation to earnings. There is ample empirical evidence, see e.g. Antel (1986), Altonji and Shakotko (1987), Topel (1991) and Altonji and Williams (2004), suggesting that mobility behavior in the labor market is endogenous in relation to earnings. On this issue, Farber (1999, p. 2470) notes that

Since tenure is the result of a series of (non)quit decisions, an earnings function with tenure as an explanatory variable can be thought of as an inverse quit function in some respects. Thus, it is arbitrary to assign the wage as the dependent variable that is “explained” by tenure.

I account for endogenous mobility behavior through a simultaneous equations approach where reduced form equations of the wage process and the processes governing job-to-job transitions and job-to-nonemployment transitions are jointly estimated.6 The endogeneity takes the form of cross-equation correlations of unobserved heterogeneity components. The resulting empirical model bears similarities to the model developed in Lillard (1999).

The model is estimated using a sample of Danish male labor market entrants followed in a period of up to 14 years. The estimation procedure exploits a matched employer-employee feature of the data to allow for a two-way unobserved heterogeneity component, accounting for both individual specific effects and match specific effects in all the model’s relations7.

---

5This argument of course hinges on the strong assumption that job search cannot influence within-job wage profiles. Postel-Vinay and Robin (2002) show how job search might inflict on within-job wage profiles.

6A more structural (as opposed to reduced form) approach, with a tighter correspondence between economic theory and the empirical model, would have facilitated more precise inference and interpretation. However, structural micro-econometric models often put serious limitations on the amount of heterogeneity that can be incorporated, and moreover, often make strong assumptions on e.g. the arrival of information to workers. In developing the empirical model applied in this paper I aimed at a high descriptive accuracy, tilting the arguments in favor of flexible reduced form analysis.

7The terms “match”, meaning a match between a firm and a worker, and “job” will be used interchangeably.
The results of the paper confirm that quits are associated with a significant immediate $jj$-premium, and that workers experience a significant within-job wage growth, both in $jj$-jobs and in $nj$-jobs. Moreover, I find that the initial $jj$-premium is indeed rather persistent. Only among medium-educated workers do the wage of workers in $nj$-jobs tend to catch up with that of workers in $jj$-jobs. The $jj$-premium is increasing in tenure for low-educated workers. For high-educated workers the job-to-job premium seems to remain constant in the long run, although some catching up takes place during the first three years of tenure. Furthermore, the analysis confirms that young workers gradually work themselves into more durable employment relationships as their careers progress, and that workers can meaningfully be grouped into “movers” and “stayers”, with “stayers” earning higher wages than “movers”.

The remainder of the paper is organized as follows: section 2 presents the data, the construction of the analysis samples and stylized facts on job mobility and earnings. In section 3 I develop an empirical model of earnings and job mobility and present the resulting likelihood function. Section 4 presents the estimated model and contains the main analysis of the paper. Section 5 contains a brief conclusion.

2 The data

The data used for the empirical analysis consists of a ten percent random sample of male workers from the register based Danish matched employer-employee dataset IDA, merged with detailed data on individual labor market histories covering the period 1986 to 1999. Female workers are discarded at the onset in order not to confound the analysis with fertility and household production issues, which are usually believed to have important effects on females’ labor market behavior. IDA contains annual socio-economic information on workers and background information on employers, and covers the entire Danish population aged 16 to 69. This information is recorded annually and remains constant within a calendar year (except earnings information, see below). The labor market history data is also register based and is built from weekly information on unemployment.

*IDA: (in Danish) Integreret Database for Arbejdsmarkedsforskning, (in English) Integrated Database for Labor Market Research.
status and mandatory pension contributions made by employers.

The labor market history data identifies four labor market states: Employment (including a firm identifier\(^9\)), temporary unemployment (unemployment with recall), unemployment and non-participation. I treat unemployment and nonparticipation as a common state, nonemployment, and construct only job spells and nonemployment spells. Spells of temporary unemployment are disregarded and jobs interrupted by temporary unemployment are aggregated into a single job spell of duration equal to the sum of the durations of the original job spells. Very short nonemployment spells (viz. in practice nonemployment spells shorter than 5 weeks) are likely to be periods of transition before an already obtained job is initiated rather than “real” nonemployment and are disregarded. Consequently, a job-to-nonemployment transition followed by a nonemployment-to-job transition within four weeks is considered a job-to-job transition.

Earnings information consists of the annual average hourly wage of the job occupied in the last week of November. This implies that jobs that are never active in the last week of November in any year, i.e. a sizeable proportion of short jobs, will have no wage information. However, wage information is available at short tenures in jobs initiated just prior to the last week of November in any year. I augment the empirical analysis to take job spells with no wage information into account and thus retain these spell in the estimation data. The earnings information is truncated from below at a set of effective minimum wages and inflated to 1999 levels using Statistics Denmark’s consumer price index.

Background information on worker characteristics is annual and includes age, educational attainment, cohabiting status, area of residence and information on offspring. I also categorize employing firms according to industries (6 categories) and according to whether or not they are private sector firms. A notable missing piece of information is working hours.\(^{10}\)

\(^9\)Employers are identified at firm and plant level. I construct job spells on firm level.

\(^{10}\)In fact, the data does carry information on interval censored working hours, but this information is only recorded in November and thus cannot be used in jobs not active in November in any year. I therefore disregard the information.
2.1 Sample selection

**Labor market entry.** Using annual information on education status, I discard workers who are not observed to be in education or who are in education at the end of the observation period. This condition implies that workers transition from education to work, or nonemployment is observed. A worker’s entry-year is the last year in which he is observed to be in education. If the worker is working in the last week of his entry-year, and the job was initiated in the entry-year, the start of the job marks the worker’s labor market entry. If the job was initiated prior to the entry-year, the start of the first subsequent job is taken as the labor market entry. If the worker is nonemployed in the last week of the entry-year, the start of the first subsequent job marks labor market entry. Furthermore, to minimize the probability that workers return to education after the end of the observation period, workers who are not observed for at least 2 years (104 weeks) after the start of their designated entry-job are discarded. Finally, I condition on the entry-job being initiated at age 16-30.

**Additional sample selection.** Having thus obtained a sample of non-leftcensored labor market histories, I construct transition indicators and discard all nonemployment spells after assigning information on prior nonemployment periods to the retained employment spells. Next, I restrict the analysis to private sector jobs and truncate labor market histories upon entry into a public sector job or a job with missing or invalid public sector information. A private sector job that is terminated by a transition into a public sector job is treated as a censored job spell in the empirical analysis. Finally, I discard workers observed to take jobs in the agricultural sector and delete observations with missing information on relevant variables.

The sample is split according to educational attainment (9-11 years, 12-14 years and 15-18 years of education) and the remaining regressors are retained as parametric controls in the empirical analysis. The educational grouping is roughly in accordance to whether workers have less than a highschool education, a highschool or a vocational education or a short, medium or long higher education. The real wage distribution in each sample is truncated from above at the 97.5th per-
centile. The resulting datasets are thus flow samples of private sector male labor market entrants aged 16-30 years at labor market entry. Table 1 presents summary statistics on the three samples.

< Table 1 about here. >

2.2 Descriptive statistics

The descriptive analysis will focus on two features of the data, the first being job mobility, as it is essential for the separate identification of worker specific effects and match specific effects in the estimation procedure. In addition, a careful descriptive analysis of the mobility data will help the building of a reasonable empirical model of mobility. The other feature of the data attended to is the relationship between earnings and mobility.

Mobility. We start the descriptive analysis by noting from Table 1 that, on average, a worker is observed for between 3.5 and 5 years, and in that period, he holds three jobs. This is a first indication that the selected workers are highly mobile. Table 2, containing information on the average and standard deviations of the durations of completed and censored spells, provides further evidence on the mobility in the estimation samples. It shows that completed job durations are rather short and that there is substantial variability in job durations both between and within samples, with average job duration increasing in length of education across samples. In fact, the average completed job duration among high-educated workers (85 weeks) is twice as large as the average duration among low-educated workers (43 weeks). These short durations result from the focus on labor market entrants, a notoriously mobile group of workers, in Denmark, a country with high labor market mobility.11 Breaking the completed spells into spells ending in job-to-job transitions and spells ending in job-to-nonemployment transitions shows that highly educated workers are less (more) likely to make job-to-nonemployment transitions (job-to-job transitions), given that the job is terminated, compared to the two lower education groups. However, the highly educated workers seem to be less mobile, with a greater fraction of their job spells being rightcensored.12

11See Jolivet, Postel-Vinay and Robin (2006) for a recent cross-country study of labor market mobility covering Europe and the U.S.
12Recall that jobs that terminate with a transition to the public sector, are treated as rightcensored spells in the analysis.
An alternative representation of the mobility data is given in Figure 1, which plots Kaplan-Meier estimates of the quarterly job-to-job and job-to-nonemployment hazard functions. Figure 1 shows that both the unconditional job-to-job and the unconditional job-to-nonemployment hazard functions exhibit strong negative duration dependence among workers with 9-11 and 12-14 years of education. The negative duration dependence is less pronounced for workers with 15-18 years of education, with the job-to-job hazard function being virtually horizontal in this latter group.

Around one third of all new jobs among low- and medium-educated workers are terminated within one quarter of a year. This, and the strong decline in the hazard within the first year of tenure, is in accordance with the findings of Topel and Ward (1992) for U.S. male labor market entrants, without stratifying on educational attainment.

Earnings and mobility. Figure 2 illustrates plots of the average hourly wages against tenure in years for $jj$-jobs and $nj$-jobs, respectively. As expected, average starting wages in $jj$-jobs are higher than average starting wages in $nj$-jobs. In addition, Figure 2 reveals that the gap between average earnings in $jj$-jobs and $nj$-jobs, the $jj$-premium, remains significant even after five years tenure. This holds for all three samples. Taken at face value, Figure 2 suggests that the $jj$-premium is rather persistent. This paper seeks to test whether the stylized fact shown in Figure 2 results from a structural relationship in the labor market (say, wage determination or differences in human capital investments) or from selection effects due to heterogeneity among workers and jobs. For this purpose an empirical model of the relationship between earnings and mobility is developed in section 3.

---

Recall that $jj$-jobs are jobs initiated by a job-to-job transition, while $nj$-jobs are initiated by a nonemployment-to-job transition.
3 An empirical model of earnings and mobility

The main difficulty in the econometric modeling of earnings and job mobility is to allow for the potential endogeneity of mobility with respect to earnings. I get around this problem by joint modeling of the wage and transition processes: First, an empirical model of earnings conditional on mobility behavior and a marginal model of mobility behavior are developed. Second, the likelihood function, which ensues from coupling the conditional earnings model with the marginal mobility model, is presented. The model is non-behavioral, or reduced-form, and the potential endogeneity of earnings with respect to mobility is introduced via cross-relation correlations of heterogeneity components rather than through explicit modeling of the behavior of firms and workers. The third step in the development of the empirical model is the precise specification of the unobserved heterogeneity components capturing the endogeneity issues, and the derivation of the empirically implementable unconditional likelihood function. The resulting empirical model is similar in structure to the model developed in Lillard (1999).

Let \( i = 1, \ldots, I \) index workers, \( j = 1, \ldots, J_i \) index jobs of worker \( i \) and \( n = 1, \ldots, N_{ijn} \) index observations on worker \( i \)'s \( j \)'th job. Hence, the worker index \( i \) nests the job index \( j \) which in turn nests the observation index \( n \). In the empirical analysis a “job” is defined by employer identity and time, i.e. a job is any position with a given employer, and two job spells with the same employer is treated as two different matches. The conditioning on exogenous variables, such as household characteristics, industry dummies etc., is suppressed in the notation of this section.

3.1 Earnings

I start out by considering wage determination. Log-wages \( w \) are made conditional on a set of potentially endogenous observable worker characteristics related to job mobility, \( z \), and a set of unobservable conditioning variables \( \eta \). The log-wage observations are assumed to arise from the linear model:

\[
w_{ijn} = \beta_0 + z_{ijn}\beta + \eta_{ijn},
\]  

(1)
where $\beta_0$ is the constant term. The vector $z_{ijn}$ contains the central variables of the analysis: tenure $t$, experience at job initiation $a$ (such that actual labor market experience equals $t + a$); an indicator for whether the job was initiated via a job-to-job transition $jj$-job; an interaction of tenure and $jj$-job; tenure in any preceding job (implicitly interacted with $jj$-job and denoted $pt$); the duration of any preceding unemployment period (implicitly interacted with $(1 - jj$-job) and denoted $pud$); and finally, the number of job-to-job transitions made since the last period of unemployment (i.e. the job’s rank in the ongoing employment cycle). Workers coming from unemployment have rank equal to zero. Hence,

$$z_{ijn} \beta = \beta_1 t_{ijn} + \beta_2 t_{ijn} \times jj$-job_{ij} + \beta_3 jj$-job_{ij} + \beta_4 pt_{ij} + \beta_5 pud_{ij} + \beta_6 rank_{ij} + \beta_7 a_{ij}.$$  (2)

In the empirical implementation of (1), tenure and experience at job initiation are included as fourth order and second order polynomials, respectively.\textsuperscript{14} Notice that (1) does not allow separate identification of tenure and experience effects in wages. In this paper the central distinction is between within-job and between-job wage growth, and not the disentanglement of tenure and experience effects in earnings growth.\textsuperscript{15} For future reference, notice that $jj$-job, duration of any preceding job, the current job’s rank, duration of any preceding unemployment spell and experience at job initiation are predetermined at job initiation.

The unobserved component $\eta$ in (1) is orthogonally decomposed in a worker specific component $\mu$, a job specific component $\nu$ and a purely transitory normal iid component $\varepsilon$:

$$\eta_{ijn} = \mu_i + \nu_{ij} + \varepsilon_{ijn}, \quad \varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2).$$  (3)

Hence, (1), (2) and (3) form a log-normal wage equation with a two-way error structure. The worker specific effect $\mu$ reflects unobserved (to the econometrician) time-invariant person specific traits such as “ability”, “personal drive” or “perseverance” that are likely to correlate with individual productivity, and the job specific component $\nu$ reflects unobserved (to the econometrician) time-invariant match specific traits that correlate with earnings such as “match quality”. The error term

\textsuperscript{14}More flexible specifications of the tenure-effect on earnings, through the use of splines, were abandoned because the added flexibility hampered separate identification of the effect of job mobility on the level and the slope of the subsequent wage-tenure profile at short tenures.

\textsuperscript{15}Manning, 2003, ch. 6 contains a discussion of the within-job/between-job wage growth distinction versus the returns to tenure/returns to experience distinction.
\( \varepsilon \) can be thought of as capturing fluctuations in earnings coming from random shocks to productivity and preferences etc. In the empirical analysis the unobserved components will be treated as random effects (although possibly correlated with the elements in \( z \)). Abowd, Kramarz and Margolis (1999) estimate a wage equation similar to (1), with firm effects in place of match effects in (3), using fixed effects techniques.\(^{16}\) The random effect approach taken here, although more restrictive in terms of the admissible correlations between observed and unobserved components, provides direct evidence on the role played by unobserved factors in shaping the relationship between wages and mobility.

### 3.2 Mobility

Since tenure is the time spent in a certain state (viz. in a given job), it is conveniently modeled in a hazard model framework. Observed tenure \( t \) is already introduced. Now, let \( T \) be the random variable from which \( t \) emanates with distribution function \( F(t) \), density function \( f(t) \), and hazard function \( h(t) = f(t)/\bar{F}(t) \), where \( \bar{F}(t) = 1 - F(t) \).\(^{17}\)

A job can either terminate in a job-to-job transition or in a job-to-nonemployment transition, and job spells thus follow a competing risk model. Let \( T^0 \) be the latent duration until a job-to-nonemployment transition occurs and let \( T^1 \) be the latent duration until a job-to-job transition occurs (with hazards \( h^0 \) and \( h^1 \), and distributions \( F^0 \) and \( F^0 \), respectively). The competing risk model stipulates that \( T = \min\{T^0, T^1\} \) implying that \( h(t) = h^0(t) + h^1(t) \) and \( \bar{F}(t) = \bar{F}^0(t) \bar{F}^1(t) \).\(^{18}\)

Each of the transition specific hazard functions are assumed to belong to the class of mixed proportional hazard (MPH) models.\(^{19}\) Hence, the hazard functions are multiplicative in a baseline hazard component \( b_k(\cdot) \), capturing the duration dependence of the hazard rate; a systematic component \( \exp\{\tilde{z}_{ijn}\gamma_k\} \), scaling the baseline hazard function according to the observed regressors \( \tilde{z}_{ijn} \);

---

\(^{16}\)If firms operate a constant return to scale technology with labor as the only variable input, the notion of a firm is not well-determined and firm effects are indistinguishable from match effects.

\(^{17}\)Since \( \bar{F}(t) = \exp\{-\int_0^t h(u)du\} \) the hazard function fully characterizes the distribution of tenure. From a practical point of view this relation is useful since an observed spell of duration \( t \) is either censored, with likelihood contribution \( \bar{F}(t) \), or non-censored, with likelihood contribution \( f(t) = h(t)\bar{F}(t) \).

\(^{18}\)Censoring is routinely handled in a competing risk framework, in fact, random censoring can be interpreted as a third and non-modeled transition type with associated latent duration, say \( T^* \), so that \( T = \min\{T^0, T^1, T^*\} \) (see Lancaster, 1990).

\(^{19}\)Van den Berg (2001) provides a thorough description of the MPH model.
and a unobserved, time-invariant mixing component $\varphi_{ij}^k$:

$$h^k(t_{ij|n}|\tilde{z}_{ijn}, \varphi_{ij}^k) = b^k(t_{ij|n}) \exp \left\{ \tilde{z}_{ijn} \gamma^k \right\} \varphi_{ij}^k, \quad k = 0, 1, \quad (4)$$

with $\tilde{z}_{ijn}$ being the subset of predetermined variables from the set of potentially endogenous conditioning variables $z_{ijn}$.

The baseline hazard functions $b^k(\cdot)$ are piecewise constant functions on $M$ baseline intervals, which are taken as given and identical across transition types; $\mathcal{M} = \{1, ..., M\}$ is the index set associated with the baseline intervals. If $q : \mathbb{R}_+ \mapsto \mathcal{M}$ is the mapping of durations into baseline intervals and $\exp\{\delta^k\}$ is the $M \times 1$ vector of destination specific baseline values (i.e. the level of the baseline hazard in any given interval), the baseline hazard function can be written as:

$$b^k(t_{ijn}) = \exp \left\{ \iota(q(t_{ijn})) \delta^k \right\}, \quad k = 0, 1, \quad (5)$$

where $\iota^m \equiv (0, ..., 0, 1, 0, ..., 0)$ is a conformable vector with the “1” appearing in the $m$’th column.

The mixing component $\varphi_{ij}^k$ is assumed to consist of a worker specific effect $\mu_{i}^k$ and a job specific effect $\nu_{ij}^k$ such that

$$\varphi_{ij}^k = \exp\{\mu_{i}^k + \nu_{ij}^k\}, \quad k = 0, 1. \quad (6)$$

Imposing (5) and (6) on (4) produces the following expression for the transition specific job hazard functions:

$$h^k(t_{ij|n}|\tilde{z}_{ijn}, \mu_{i}^k, \nu_{ij}^k) = \exp \left\{ \iota(q(t_{ij|n})) \delta^k + \tilde{z}_{ijn} \gamma^k + \mu_{i}^k + \nu_{ij}^k \right\}, \quad k = 0, 1. \quad (7)$$

To obtain a closed form solution to the survivor functions let $t_{m}^*$ be the length of the $m$’th baseline interval and define the (row-)vector valued function $p : \mathbb{R}_+ \mapsto \prod_{m \in \mathcal{M}} [0, t_{m}^*]$, as the mapping of durations into a $1 \times M$ vector whose $m$’th entry is the time spent in the $m$’th baseline interval. Then, upon assuming that regressors $\tilde{z}_{ijn}$ are constant between observation points, integration of (7) produces the following expression for the transition specific survivor function:

$$\overline{F}^k(t_{ij|n}|\tilde{z}_{ijn}, \mu_{i}^k, \nu_{ij}^k) = \exp \left\{ -\sum_{s=1}^{n} \exp \left\{ \tilde{z}_{ijn} \gamma^k + \mu_{i}^k + \nu_{ij}^k \right\} [p(t_{ij|s}) - p(t_{ij|s-1})] \exp\{\delta^k\} \right\}, \quad (8)$$
for $k = 0, 1$, where $\tilde{z}_{ijn} \equiv \{\tilde{z}_{ij}s\}_{s=1}^n$ denotes the regressor trajectories up to and including observation $n$ on individual $i$’s $j$’th job. The overall job survivor function $\tilde{F}(t_{ijn}|\cdot)$ conditional on observable and unobservable factors is given as $\tilde{F}(t_{ijn}|\cdot) = \tilde{F}^0(t_{ijn}|\cdot)\tilde{F}^1(t_{ijn}|\cdot)$.

3.3 The likelihood function

A compact notation for the different sets of variables is needed. Recall that $w$ is the log-wage and $z$ is the set of potentially endogenous mobility variables, with $\tilde{z}$ being predetermined at job initiation. Bold font indicates the inclusion of the whole past trajectory of the relevant set of variables; for example $z_{ijn} = \{z_{ij}s\}_{s=1}^n$. Let $u = (\mu, \mu^0, \mu^1)'$ be the set of worker specific heterogeneity components and let $v = (\nu, \nu^0, \nu^1)'$ be the set of job specific heterogeneity components.

The conditional likelihood function. Let $c^0$ and $c^1$ indicate job termination by a job-to-nonemployment transition and a job-to-job transition, respectively. The data contains observations with no valid wage information that still carry useful information on mobility behavior. I assume that wage information is missing at random and define the indicator variable $d_{ijn}$ taking the value 1 if the $n$’th observation on worker $i$’s $j$’th job has valid wage information, and the value 0 otherwise. Then, denoting the density of observed wages by $g(\cdot)$ and the set of parameters introduced so far by $\tilde{\theta}$, the likelihood contribution from an observation $\{w_{ijn}, z_{ijn}\}$ conditional on the heterogeneity components $u$ and $v$ is given as:

$$L_{ijn}(\tilde{\theta}; u', v') = g(w_{ijn}|z_{ijn}, \mu, \nu)^{d_{ijn}}$$
$$h^0(t_{ijn}|\tilde{z}_{ijn}, \mu^0, \nu^0)^{c_{ijn}^0} \tilde{F}^0(t_{ijn}|\tilde{z}_{ijn}, \mu^0, \nu^0)/\tilde{F}^0(t_{ijn-1}|\tilde{z}_{ijn-1}, \mu^0, \nu^0)$$
$$h^1(t_{ijn}|\tilde{z}_{ijn}, \mu^1, \nu^1)^{c_{ijn}^1} \tilde{F}^1(t_{ijn}|\tilde{z}_{ijn}, \mu^1, \nu^1)/\tilde{F}^1(t_{ijn-1}|\tilde{z}_{ijn-1}, \mu^1, \nu^1),$$

where $t_{ij0} \equiv 0$. The likelihood function (9) is the joint density of earnings, durations and transition types and is made up of three components: line one is the contribution from wages conditional on mobility behavior, line two is the contribution from the job-to-job transition process, and line three is the contribution from the job-to-nonemployment transition process. Line two and three together form the distribution of durations and transitions.
To further break down (9), consider observation $n$ on worker $i$’s $j$’th job, and suppose that the ensuing job survives the time interval covered by the observation in question, i.e. $c_{ijn}^0 = c_{ijn}^1 = 0$. The likelihood contribution is the density at the observed wage $w_{ijn}$, $g(w_{ijn}|·)$, times the probability that the job survives tenure $t_{ijn}$, conditional on having survived to tenure $t_{ijn-1}$. In the competing risk model this probability is given as $\frac{\bar{F}_0(t_{ijn}|·)}{\bar{F}_0(t_{ijn-1}|·)}$ times $\frac{\bar{F}_1(t_{ijn}|·)}{\bar{F}_1(t_{ijn-1}|·)}$. If, instead, the job ends in a job-to-nonemployment transition in observation $n$, such that $c_{ijn}^0 = 1$ and $c_{ijn}^1 = 0$ (and by definition $n = N_{ij}$), the likelihood contribution would be the density at the observed wage $w_{ijn}$, $g(w_{ijn}|·)$, times the density that the job ends in a job-to-nonemployment transition at $t_{ijn}$, conditional on the job having survived up to $t_{ijn-1}$. This density is simply $h_0(t_{ijn}|·)$ times $\frac{\bar{F}_0(t_{ijn}|·)}{\bar{F}_0(t_{ijn-1}|·)}$ times $\frac{\bar{F}_1(t_{ijn}|·)}{\bar{F}_1(t_{ijn-1}|·)}$. The likelihood contribution in the case of job-to-job transitions follows by similar reasoning.

The empirical model for wages and job mobility developed above gives explicit expressions for each of the components in (9). Imposing (1), (2), (3), (7) and (8) on (9) produces the following conditional likelihood contribution from an observation $\{w_{ijn}, z_{ijn}\}$:

$$L_{ijn}(\tilde{\theta}; u', v') = \frac{1}{\sqrt{2\pi\sigma^2_\varepsilon}} \exp\left\{ -\frac{1}{2} \frac{(w_{ijn} - \beta_0 - z_{ijn}\beta - \mu - \nu)^2}{\sigma^2_\varepsilon} \right\} \frac{d_{ijn}}{d_{ijn}} \exp\left\{ \gamma_0(t_{ijn})\delta^0 + z_{ijn}\gamma^0_0 + \mu^0 + \nu^0 \right\} \frac{d_{ijn}}{d_{ijn}} \exp\left\{ \gamma_1(t_{ijn})\delta^1 + z_{ijn}\gamma^1_1 + \mu^1 + \nu^1 \right\} \frac{d_{ijn}}{d_{ijn}} \exp\left\{ -\exp\left\{ z_{ijn}\gamma^0_0 + \mu^0 + \nu^0 \right\} [p(t_{ijn}) - p(t_{ijn-1})] \exp\{\delta^0\} \right\} \exp\left\{ -\exp\left\{ z_{ijn}\gamma^1_1 + \mu^1 + \nu^1 \right\} [p(t_{ijn}) - p(t_{ijn-1})] \exp\{\delta^1\} \right\}.$$

(10)

To make (10) empirically implementable the unobserved effects $u = (\mu, \mu^0, \mu^1)'$ and $v = (\nu, \nu^0, \nu^1)'$ must be integrated out of (10).

**Unobserved heterogeneity.** The empirical model is formulated with a two-way error structure in all its three building blocks and the unrestricted model thus has six heterogeneity components; the three individual specific effects $u = (\mu, \mu^0, \mu^1)'$ and the three match specific effects $v = (\nu, \nu^0, \nu^1)'$. The covariance matrix of the heterogeneity components could in principle be left unrestricted. However, in practice, estimation of the unrestricted model is likely to be a very arduous task; one
that I will not take on in this paper. Instead, a number of restrictions on the covariances between
the heterogeneity components are imposed. First, it is assumed that the individual specific effects \( u' \)
are independent of the match specific effects \( v' \). Second, it is assumed that the job specific effects are
serially uncorrelated, also within an employment cycle.\(^{20}\) Third, it is assumed that the heterogeneity
components in the job-to-job transition process and the job-to-nonemployment process are linearly
related through a pair of factor loading models, such that \( \mu^0 = \xi \mu^1 \) and \( \nu^0 = \zeta \nu^1 \) for some pair
of scalar loading coefficients \((\xi, \zeta)\). Fourth, it is assumed that the heterogeneity components, now
effectively reduced to the pairs \( u = (\mu, \mu^1)' \) and \( v = (\mu, \mu^1)' \), are normally distributed:
\[
\begin{equation}
\begin{aligned}
    u &\sim \mathcal{N}(0, \Sigma), \\
    \Sigma &\equiv 
    \begin{pmatrix}
        \sigma_1^2 & \sigma_{12} \\
        \sigma_{12} & \sigma_2^2
    \end{pmatrix}
    \tag{11}
    \\
    v &\sim \mathcal{N}(0, \Omega), \\
    \Omega &\equiv 
    \begin{pmatrix}
        \omega_1^2 & \omega_{12} \\
        \omega_{12} & \omega_2^2
    \end{pmatrix}
\end{aligned}
\end{equation}
\]
which completes the specification of the empirical model.

Integrating out the heterogeneity components yields the following expression for the unconditional
likelihood contribution from a set of observations on worker \( i \):
\[
\begin{equation}
\begin{aligned}
    L_i(\theta) = \frac{1}{\pi} \int_{\mathbb{R}^2} \prod_{j=1}^{J_i} \left\{ \frac{1}{\pi} \int_{\mathbb{R}^2} \prod_{n=1}^{N_{ij}} L_{ijn}(\tilde{\theta}; [\sqrt{2}\Sigma^{1/2} y]'', [\sqrt{2}\Omega^{1/2} x]') e^{-x'x} dx \right\} e^{-y'y} dy,
    \\
\end{aligned}
\end{equation}
\]
where \( \theta \) is the full set of parameters to be estimated, including the parameters defining the distribution of unobservables, \( L_{ijn}(\cdot) \) is given by (10), \( \Sigma^{1/2} \) and \( \Omega^{1/2} \) are the Cholesky decompositions of \( \Sigma \) and \( \Omega \), and \( y \sim \mathcal{N}(0, \mathbf{I}_2) \) and \( x \sim \mathcal{N}(0, \mathbf{I}_2) \) are bivariate standard normal vectors. The integrals in (13) are dealt with by use of product-rule Gauss-Hermite quadratures (Judd, 1998, ch. 7). Thus, (13) is approximated with the weighted sum
\[
\begin{equation}
\begin{aligned}
    L_i(\theta) \approx \frac{1}{\pi} \sum_{k_1=1}^{K} \sum_{k_1=1}^{K} \tau_{k_1}^0 \tau_{k_2}^1 \left[ \prod_{j=1}^{J_i} \frac{1}{\pi} \sum_{k_3=1}^{K} \sum_{k_4=1}^{K} \tau_{k_3}^0 \tau_{k_4}^1 \prod_{n=1}^{N_{ij}} L_{ijn}(\tilde{\theta}; [\sqrt{2}\Sigma^{1/2} y_{k_1,k_2}]', [\sqrt{2}\Omega^{1/2} x_{k_3,k_4}]') \right] 
    \\
\end{aligned}
\end{equation}
\]
where \( y_{k_1,k_2} = (y_{k_1}, y_{k_2})' \) and \( x_{k_3,k_4} = (x_{k_3}, x_{k_4})' \), with the individual entries being one-dimensional
Gauss-Hermite quadrature nodes, and the \( \tau \)'s the associated weights. In the empirical implemen-

\(^{20}\)Seeing that workers are likely to move systematically from “bad” jobs to “good” jobs this assumption is admittedly
not consistent with a job search model.
Finally, recall that the set of potentially endogenous variables in the wage equation (2) includes the duration of any preceding unemployment period, even though unemployment spells are not modeled in the empirical analysis. The (non-)modeling of the unemployment periods is valid insofar as the unobservables, in the nonemployment-to-job transition process, are linearly related to the unobservables in the job-to-job transition process.

4 Results

This section presents the estimated model. The full set of parameters is large and for the sake of clarity, I have omitted the parameter estimates associated with the exogenous controls in the wage equation and the hazard models from the presentation. Generally, these estimates are plausible in terms of signs and magnitudes, and they are available upon request. Instead, I will focus on the effect of job mobility on wages, the effect of predetermined mobility variables on future mobility, and on the distributions of unobserved heterogeneity.

4.1 Wages conditional on mobility behavior

Table 3 presents the estimated parameters of the potentially endogenous mobility variables in the wage equation. The set contains a number of predetermined mobility variables, namely a \( jj \)-job indicator, the job’s rank in the ensuing employment cycle, tenure in any preceding job, duration of any preceding unemployment period and experience at job initiation. The rank of the job in the employment-cycle is the number of job-to-job transitions made since the last period of unemployment. Hence, workers in \( nj \)-jobs have rank in employment-cycle equal to zero. In addition, the potentially endogenous mobility controls in the wage equation include tenure and tenure interacted with the \( jj \)-job indicator. These variables are not predetermined at job initiation. The predetermined mobility variables affect only the starting wage (or the job’s wage level), while the remaining mobility variables affect the slope of the job’s wage profile.

\[21\text{The main drawback of multi-dimensional numerical integration by product rules is that the number of function evaluations increase rapidly with the number of quadrature points in each direction. In this case, with two two-dimensional integrals and five points in each dimension, the total number of function evaluations at each step in the iterative optimization scheme applied to (14) is } 5^4 = 625.\]
Starting wages and job initiation. The parameter estimates reported in Table 3 reveal that $jj$-jobs have significantly higher starting wages than $nj$-jobs. To quantify the wage hike, compare the starting wage of a worker who makes his first job-to-job transition in an employment cycle after one year of tenure in his previous job, to that of a worker initiating his job after 12 weeks of unemployment. Among workers with 9-11 years of education, the starting wage in $jj$-jobs is 2.5 percent higher than that in $nj$-jobs. The corresponding figures for workers with 12-14 and 15-18 years of education are 5.5 percent and 8 percent, respectively.

Additional job-to-job transitions (improving the current job’s rank in the employment cycle) does not affect starting wages in the new jobs significantly for low-educated workers. Among medium- and high-educated workers, every additional job-to-job transition increases starting wages in the new jobs by around 0.6 percent and 2.8 percent, respectively.

Tenure in the preceding job also has a positive effect on the wage level in the current job. The estimated model implies that an additional year of tenure in a preceding job increases earnings in the current job by around 0.75 percent, with no substantial differences across education groups. The duration of preceding unemployment spells also has similar effects across education groups, the effect being positive, albeit small, and in fact insignificant among high-educated workers.\footnote{That unemployment periods have a positive effect on subsequent earnings might result from the fact that workers’ entry jobs (necessarily) are initiated by a nonemployment-to-job transition where the worker is recorded as having zero weeks of prior unemployment. If working involves some kind of human capital accumulation that is not completely lost upon entry into unemployment, one would expect workers making nonemployment-to-job transitions after a period of unemployment to have higher starting wages than workers initiating their very first job. In other words, the positive effect of duration of prior unemployment periods might be due to inadequate controls for experience.}

Finally, consider experience at job initiation. In all samples there is a concave relationship between earnings in a job and experience at job initiation with the experience profile peaking at 14 years, 10 years and 11 years for workers with 9-11, 12-14 and 15-18 years of education, respectively. The parameter estimates imply that 10 years of experience increase starting wages in new jobs by 18 percent, 12 percent and 27 percent for workers with 9-11, 12-14 and 15-18 years of experience, respectively.
Summing up, I find that: (i) the first job-to-job transition in an employment cycle is associated with a sizable gain in starting wages; (ii) additional job-to-job transitions also increase the starting wage in the new job, although this effect is quantitatively small compared to the gain obtained via the first job-to-job transition; and (iii) the longer the preceding job, the higher the starting wage in the current job. Although these observations hold true across education groups, the quantitative importance of job mobility in generating starting wage hikes is increasing in educational attainment.

**Persistence of the jj-premium.** The predetermined mobility variables considered so far only shift the level of wages, and are not informative on the within-job persistence of the jj-premium. The parameters of tenure and tenure interacted with the jj-job indicator in Table 3 shed light on this issue. Wald-tests of the joint significance of the interaction terms reject the null at conventional significance levels in all three education groups. Hence, the data reveal significant differences between within-job wage trajectories in jj-jobs and nj-jobs, and thus indicates that the jj-premium does not remain constant as tenure is accumulated.

Table 4 reports the estimated jj-premium as a function of tenure, by comparing the wage of a worker in a jj-job, who is in the second job in the employment cycle and whose previous job lasted one year, with the wage of a worker who initiated a job after 12 weeks of unemployment. In this case, the initial jj-premium of 2.5 percent among workers with 9-11 years of education evolves into a 5 percent advantage after 2 years of tenure. After 10 years of tenure the premium has risen to more than 6 percent. Among workers with 12-14 years of education, the initial jj-premium of around 5.5 percent is slowly eroded as the job progresses because nj-jobs offer more favorable within-job wage trajectories at short tenures; in fact, after five years of tenure, the jj-premium is insignificant. Finally, among workers with 15-18 years of education the initial jj-premium is more than 8 percent, and after ten years of tenure it is around 7 percent. In between the jj-premium is insignificant, albeit positive, between three and six years of tenure.

*Table 4 about here.*

The presentation of the estimated relationship between job initiation and subsequent within-job wage profiles is rounded off by Figure 3, which plots the implied wage-tenure profiles for the
situation considered above. In addition to presenting the evidence in Table 4 graphically, Figure 3 shows that workers experience a non-negligible within-job wage growth. Restricting attention to the “baseline” profiles of $nj$-jobs, it is seen that for low-educated workers, five (ten) years of tenure raise earnings by around 10 percent (13 percent). Among the medium-educated workers, the returns to five (ten) years of tenure is 14 percent (18 percent), and finally, for the high-educated workers, five (ten) years of tenure raise earnings by 30 percent (40 percent). These figures are topped with the $jj$-premium when considering $jj$-jobs.

In summary, the persistence of the initial $jj$-premium differs across education groups. Among low-educated and high-educated workers, the $jj$-premium exhibits a high degree of persistence (even though, for high-educated workers the $jj$-premium in fact becomes insignificant at 3 to 6 years of tenure), while the persistence is less pronounced among medium educated workers, where wages in $nj$-jobs catch up with wages in $jj$-jobs, but only after 5 years of tenure. The analysis thus revealed that young workers tend to move from “low-wage” jobs to “high-wage” jobs via a process of job-to-job transitions, and that only among medium-educated workers do subsequent within-job wage growth manage to neutralize the initial wage differential in $jj$-jobs and $nj$-jobs.

4.2 Mobility behavior

The empirical analysis also provides insight on mobility behavior, unconditional on the wage, through the estimated competing risk MPH model. The focal point of the analysis will be the effects of predetermined mobility variables on future mobility. However, first consider the plots of the estimated hazard functions for the median worker in Figure 4. After controlling for observed regressors and unobserved worker and match specific effects, the three job-to-job transition hazard functions exhibit almost no duration dependence, a picture in stark contrast to the pronounced negative duration dependence of the unconditional hazard plots in Figure 1. Controlling for heterogeneity in the job-to-nonemployment hazard function does not alter the conclusion drawn from the unconditional plots in Figure 1: The likelihood of lay-offs decline with tenure.

In summary, the persistence of the initial $jj$-premium differs across education groups. Among low-educated and high-educated workers, the $jj$-premium exhibits a high degree of persistence (even though, for high-educated workers the $jj$-premium in fact becomes insignificant at 3 to 6 years of tenure), while the persistence is less pronounced among medium educated workers, where wages in $nj$-jobs catch up with wages in $jj$-jobs, but only after 5 years of tenure. The analysis thus revealed that young workers tend to move from “low-wage” jobs to “high-wage” jobs via a process of job-to-job transitions, and that only among medium-educated workers do subsequent within-job wage growth manage to neutralize the initial wage differential in $jj$-jobs and $nj$-jobs.

4.2 Mobility behavior

The empirical analysis also provides insight on mobility behavior, unconditional on the wage, through the estimated competing risk MPH model. The focal point of the analysis will be the effects of predetermined mobility variables on future mobility. However, first consider the plots of the estimated hazard functions for the median worker in Figure 4. After controlling for observed regressors and unobserved worker and match specific effects, the three job-to-job transition hazard functions exhibit almost no duration dependence, a picture in stark contrast to the pronounced negative duration dependence of the unconditional hazard plots in Figure 1. Controlling for heterogeneity in the job-to-nonemployment hazard function does not alter the conclusion drawn from the unconditional plots in Figure 1: The likelihood of lay-offs decline with tenure.

In summary, the persistence of the initial $jj$-premium differs across education groups. Among low-educated and high-educated workers, the $jj$-premium exhibits a high degree of persistence (even though, for high-educated workers the $jj$-premium in fact becomes insignificant at 3 to 6 years of tenure), while the persistence is less pronounced among medium educated workers, where wages in $nj$-jobs catch up with wages in $jj$-jobs, but only after 5 years of tenure. The analysis thus revealed that young workers tend to move from “low-wage” jobs to “high-wage” jobs via a process of job-to-job transitions, and that only among medium-educated workers do subsequent within-job wage growth manage to neutralize the initial wage differential in $jj$-jobs and $nj$-jobs.

4.2 Mobility behavior

The empirical analysis also provides insight on mobility behavior, unconditional on the wage, through the estimated competing risk MPH model. The focal point of the analysis will be the effects of predetermined mobility variables on future mobility. However, first consider the plots of the estimated hazard functions for the median worker in Figure 4. After controlling for observed regressors and unobserved worker and match specific effects, the three job-to-job transition hazard functions exhibit almost no duration dependence, a picture in stark contrast to the pronounced negative duration dependence of the unconditional hazard plots in Figure 1. Controlling for heterogeneity in the job-to-nonemployment hazard function does not alter the conclusion drawn from the unconditional plots in Figure 1: The likelihood of lay-offs decline with tenure.

In summary, the persistence of the initial $jj$-premium differs across education groups. Among low-educated and high-educated workers, the $jj$-premium exhibits a high degree of persistence (even though, for high-educated workers the $jj$-premium in fact becomes insignificant at 3 to 6 years of tenure), while the persistence is less pronounced among medium educated workers, where wages in $nj$-jobs catch up with wages in $jj$-jobs, but only after 5 years of tenure. The analysis thus revealed that young workers tend to move from “low-wage” jobs to “high-wage” jobs via a process of job-to-job transitions, and that only among medium-educated workers do subsequent within-job wage growth manage to neutralize the initial wage differential in $jj$-jobs and $nj$-jobs.

4.2 Mobility behavior

The empirical analysis also provides insight on mobility behavior, unconditional on the wage, through the estimated competing risk MPH model. The focal point of the analysis will be the effects of predetermined mobility variables on future mobility. However, first consider the plots of the estimated hazard functions for the median worker in Figure 4. After controlling for observed regressors and unobserved worker and match specific effects, the three job-to-job transition hazard functions exhibit almost no duration dependence, a picture in stark contrast to the pronounced negative duration dependence of the unconditional hazard plots in Figure 1. Controlling for heterogeneity in the job-to-nonemployment hazard function does not alter the conclusion drawn from the unconditional plots in Figure 1: The likelihood of lay-offs decline with tenure.
Table 5 contains the estimated parameters associated with the predetermined mobility variables in the systematic part of the hazard functions. However, since the quantitative effects of the controls cannot be read directly from Table 5, consider instead Table 6 tabulating expected job durations in five different setups. Setup A shows expected durations in a \(nj\)-job initiated after 12 weeks of unemployment. Setup B shows expected durations for a \(jj\)-job of rank one (i.e. the worker occupying the job has made one job-to-job transition since his last unemployment period), with tenure in the previous job being 9 months. Setup C relate to a \(jj\)-job of rank 2, where tenure in the previous job is one year. Finally, setup D considers a \(jj\)-job, this time of rank 3, and with the preceding job lasting 15 months. Hence, going from setup B to setup D, the workers have made more job-to-job transitions and the previous job lasted longer. Experience at job initiation is held constant at 3 years for all setups and the exogenous regressors are held constant at the median values. Likewise, the unobserved heterogeneity components are fixed at their means.

< Tables 5 and 6 about here. >

The top panel of Table 6 shows the expected duration to exit by any transition. The first thing to note is that the estimated model predicts quantitatively important differences in job stability across educational groups, most notably between workers with 15-18 years of education and the two remaining groups. High educated workers have more stable jobs, the expected duration of a job in this strata being up to twice as long as that of workers with the shortest educations. Comparing across the setups in the top panel of Table 6, reveals that \(nj\)-jobs are more unstable than \(jj\)-jobs.

The gains in expected job duration from the first job-to-job transition in an employment cycle, where the previous job lasted nine months (i.e. comparing row 1 and 2 in Table 6), vary across educational groups from 11 weeks among workers with 15-18 years of education to 33 weeks among workers with 12-14 years of education. The effect of additional job-to-job transitions and increases in the completed tenure in the previous job, i.e. going to setup C and D, increases expected job durations further in all education groups.

The two lower panels of Table 6 show expected durations to exit by a job-to-job transition and to exit by a job-to-nonemployment transition, respectively. The transition specific expected job
durations exhibit the same pattern as the overall job durations, i.e. the increase in expected job
duration from making job-to-job transitions comes from both longer expected duration to a quit
and longer expected duration to a layoff.23

In conclusion, the evidence presented in Table 6 suggests that young workers gradually work
themselves into more durable employment relationships as their labor market careers progress by
way of job-to-job transitions. This finding is in line with that of Topel and Ward (1992) for U.S.
male labor market entrants.

4.3 Unobserved heterogeneity

Table 7 presents the estimated unobserved heterogeneity distributions. Recall that the worker spe-
cific effects and the match specific effects are assumed independent. The worker specific components,
respectively the match specific components, in the wage equation and the job-to-job transition haz-
ard function are joint normal with correlation coefficient \( \rho_\sigma = \sigma_{12}/(\sigma_1\sigma_2) \) and \( \rho_\omega = \omega_{12}/(\omega_1\omega_2) \).
Both the worker specific component and the match specific component in the job-to-nonemployment
hazard function are assumed to be perfectly correlated with the corresponding components in the
job-to-job hazard function via a pair of factor loading models, with loading coefficients \( \xi \) (worker
effects) and \( \zeta \) (match effects).

< Table 7 about here. >

**Residual log-wage variance.** The bottom panel in Table 7 decomposes residual log-wage vari-
ance in worker heterogeneity, job heterogeneity and idiosyncratic shock heterogeneity. Worker
heterogeneity accounts for around 50 percent of residual log-wage variance among high-educated
workers, while it only accounts for around 20 percent among the low- and medium-educated work-
ers. Match specific effects account for between 45 and 50 percent of log wage variance among
workers with less than 15 years of education, but only around 25 percent for workers with 15-18
years of education. Finally, the idiosyncratic component is quantitatively more important among
the low skilled, accounting for around 35 percent of log-wage variance among workers with 9-11

---

23The very high expected durations to exit by layoff reflects the fact that the job-to-nonemployment baseline hazard
rate is virtually zero at long tenures.
years of education, with the corresponding figure dropping to around 30 percent and 25 percent among workers with 12-14 and 15-18 years of education, respectively. In short, worker heterogeneity is the most important determinant of residual wage dispersion for high educated workers, while match specific effects account for the lion’s share of residual wage dispersion among low- and medium-educated workers.

**Individual specific effects.** Now turn to the top panel in Table 7, and consider worker specific heterogeneity. The loading coefficients in the job-to-nonemployment hazard functions are significantly positive in all three strata. Thus, workers with high propensity to quit, also have a high risk of being laid off. This suggests a partition of workers into “movers” with high likelihood of job termination, and “stayers” with low likelihood of job termination. Although the size of the estimated correlation coefficient of worker effects in wages and in mobility behavior $\rho_\sigma$ varies considerably across education groups, the sign is consistently negative. Hence, the data indicates that “movers”, *ceteris paribus*, tend to earn lower wages than “stayers”. However, there are important quantitative differences across education groups. Among low-educated workers, the correlation between the wage component and the mobility component is rather weak, with $\rho_\sigma$ being less than 0.10. The correlation is stronger for medium-educated workers with $\rho_\sigma$ being 0.22, and is indeed very strong for the highest educated workers, where $\rho_\sigma$ equals 0.80. Thus, even though a mover-stayer interpretation of the relationship between wages and mobility cannot be refuted by the data, it seems to be stronger among the high-educated workers. The statement that “stayers” earn higher wages does not contradict the conclusion that job-to-job transitions are associated with a sizeable (and rather persistent) *jj*-premium. Indeed, this latter conclusion is valid *conditional* on worker (and match) specific heterogeneity.

**Match specific effects.** Match specific effects enter significantly in both wages and mobility behavior. One would expect “high-wage” jobs to terminate at a lower rate than “low-wage” jobs, and indeed, the estimated model confirms this, with negative point estimates of $\rho_\omega$ in all three strata, even though the correlation is insignificant among high-educated workers. It should be
pointed out that the reduced form model applied here does not allow for serial correlation of match
effects within employment cycles. Indeed, allowing for such correlations is an interesting extension
of the analysis, and I conjecture that this will reduce the estimated $jj$-premium.

5 Conclusion

This paper adds to the existing empirical literature on the relation between wages and job mobility
by paying particular attention to the persistence of any initial mobility premium ($jj$-premium). I
find a sizeable initial $jj$-premium, (ranging from 2.5 percent to 8 percent in a specific, but “typical”,
situation considered in the paper) and the premium turned out to be rather persistent. Only among
medium-educated workers does the $jj$-premium vanish, and even in this case the catch-up period
is 5 years. The $jj$-premium is increasing in tenure for low-educated workers. For high-educated
workers the job-to-job premium seems to remain constant in the long run, although some catching
up takes place during the first three years of tenure. Hence, young workers move from “low-wage”
jobs to “high-wage” jobs, and these job-to-job transitions have lasting effects on workers’ wages.
This is in line with the findings of Arulampalam (2001). The reduced form analysis does not permit
conclusions as to what causes the persistent mobility premiums, but in the motivation of the paper
I noted that it might come from already employed workers having more bargaining power than
unemployed workers, or that firms take unemployment periods as a signal of low productivity or
low susceptibility of training.

The estimated model also provides evidence on mobility behavior, unconditional on the wage.
This part of the analysis confirmed that young workers gradually work themselves into more durable
employment relationships as their careers progress. All conclusions were reached conditional on
worker and match specific effects. An analysis of the correlated unobserved worker and match
specific effects in the wage equation and the competing risk job hazards justified a grouping of
workers into “movers” and “stayers”, with “stayers” earning higher wages than “movers”. Not
surprisingly, with respect to match specific effects, it turned out that “bad” matches tend to break
up sooner than “good” matches.
References


Table 1: Summary statistics on estimation samples

<table>
<thead>
<tr>
<th></th>
<th>Ed. 9-11</th>
<th>Ed. 12-14</th>
<th>Ed. 15-18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of observations</td>
<td>35,223</td>
<td>68,368</td>
<td>14,474</td>
</tr>
<tr>
<td>Number of observations with nonmissing wage</td>
<td>16,905</td>
<td>40,961</td>
<td>10,257</td>
</tr>
<tr>
<td>Number of workers</td>
<td>5,408</td>
<td>8,918</td>
<td>2,104</td>
</tr>
<tr>
<td>Number of jobs</td>
<td>17,146</td>
<td>30,217</td>
<td>5,219</td>
</tr>
<tr>
<td>Average number of jobs per worker</td>
<td>≈ 3</td>
<td>≈ 3</td>
<td>≈ 3</td>
</tr>
<tr>
<td>Average number of years in the sample</td>
<td>≈ 3.5</td>
<td>≈ 4.5</td>
<td>≈ 5</td>
</tr>
</tbody>
</table>

Table 2: Job durations—counts, means and S.E.’s

<table>
<thead>
<tr>
<th></th>
<th>Ed. 9-11</th>
<th>Ed. 12-14</th>
<th>Ed. 15-18</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>Mean</td>
<td>S.E.</td>
<td>Mean</td>
</tr>
<tr>
<td>Jobs, completed/wks.</td>
<td>12,737</td>
<td>43</td>
<td>61</td>
</tr>
<tr>
<td>Jobs, exit by JJ-trans./wks.</td>
<td>9,335</td>
<td>44</td>
<td>61</td>
</tr>
<tr>
<td>Jobs, exit by JN-trans./wks.</td>
<td>3,402</td>
<td>40</td>
<td>53</td>
</tr>
<tr>
<td>Jobs, right censored*/wks.</td>
<td>4,409</td>
<td>92</td>
<td>111</td>
</tr>
</tbody>
</table>

*Right censored includes jobs with exit to a public sector job.
Figure 1: Kaplan-Meier estimates of job hazard functions—by exit states
Figure 2: Raw data—within-job earnings growth—jj-jobs and nj-jobs
Table 3: Estimated wage equation—mobility behavior

<table>
<thead>
<tr>
<th></th>
<th>Ed. 9-11</th>
<th>Ed. 12-14</th>
<th>Ed. 15-18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tenure (in years)</td>
<td>−0.062</td>
<td>0.0279*</td>
<td>0.0534*</td>
</tr>
<tr>
<td></td>
<td>(0.0080)</td>
<td>(0.0046)</td>
<td>(0.0094)</td>
</tr>
<tr>
<td>Tenure² (in years²)</td>
<td>0.0113*</td>
<td>0.0024</td>
<td>0.0033</td>
</tr>
<tr>
<td></td>
<td>(0.0053)</td>
<td>(0.0018)</td>
<td>(0.0039)</td>
</tr>
<tr>
<td>Tenure³ (in years³)</td>
<td>−0.0014*</td>
<td>−0.0006*</td>
<td>−0.0010*</td>
</tr>
<tr>
<td></td>
<td>(0.0005)</td>
<td>(0.0003)</td>
<td>(0.0006)</td>
</tr>
<tr>
<td>Tenure⁴ (in years⁴)</td>
<td>0.0000*</td>
<td>0.0000*</td>
<td>0.0000*</td>
</tr>
<tr>
<td></td>
<td>(0.0000)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
</tr>
<tr>
<td>Tenure × jj-job (in years)</td>
<td>0.0401*</td>
<td>0.0045</td>
<td>−0.0350*</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
<td>(0.0062)</td>
<td>(0.0148)</td>
</tr>
<tr>
<td>Tenure² × jj-job (in years²)</td>
<td>−0.0178*</td>
<td>−0.0070*</td>
<td>0.0033</td>
</tr>
<tr>
<td></td>
<td>(0.0047)</td>
<td>(0.0028)</td>
<td>(0.0064)</td>
</tr>
<tr>
<td>Tenure³ × jj-job (in years³)</td>
<td>0.0025*</td>
<td>0.0010*</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0067)</td>
<td>(0.0004)</td>
<td>(0.0010)</td>
</tr>
<tr>
<td>Tenure⁴ × jj-job (in years⁴)</td>
<td>−0.0001*</td>
<td>0.0000*</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>(0.0060)</td>
<td>(0.0000)</td>
<td>(0.0000)</td>
</tr>
</tbody>
</table>

Wald(interactions) \( \sim \chi^2(4) \) 18.86 267.27 66.88

*An "*" indicates significance at the 5 percent level.

Table 4: Job Mobility Premium

<table>
<thead>
<tr>
<th></th>
<th>Ed. 9-11</th>
<th>Ed. 12-14</th>
<th>Ed. 15-18</th>
</tr>
</thead>
<tbody>
<tr>
<td>At job start</td>
<td>0.0243*</td>
<td>0.0567*</td>
<td>0.0831*</td>
</tr>
<tr>
<td></td>
<td>(0.0082)</td>
<td>(0.0049)</td>
<td>(0.0106)</td>
</tr>
<tr>
<td>1 year of tenure</td>
<td>0.0490*</td>
<td>0.0553*</td>
<td>0.0618*</td>
</tr>
<tr>
<td></td>
<td>(0.0056)</td>
<td>(0.0036)</td>
<td>(0.0078)</td>
</tr>
<tr>
<td>2 years of tenure</td>
<td>0.0516*</td>
<td>0.0454*</td>
<td>0.0289*</td>
</tr>
<tr>
<td></td>
<td>(0.0063)</td>
<td>(0.0040)</td>
<td>(0.0086)</td>
</tr>
<tr>
<td>3 years of tenure</td>
<td>0.0431*</td>
<td>0.0319*</td>
<td>0.0156</td>
</tr>
<tr>
<td></td>
<td>(0.0065)</td>
<td>(0.0041)</td>
<td>(0.0084)</td>
</tr>
<tr>
<td>4 years of tenure</td>
<td>0.0321*</td>
<td>0.0184*</td>
<td>0.0117</td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0043)</td>
<td>(0.0085)</td>
</tr>
<tr>
<td>5 years of tenure</td>
<td>0.0245*</td>
<td>0.0075</td>
<td>0.0162</td>
</tr>
<tr>
<td></td>
<td>(0.0076)</td>
<td>(0.0045)</td>
<td>(0.0088)</td>
</tr>
<tr>
<td>6 years of tenure</td>
<td>0.0233*</td>
<td>0.0007</td>
<td>0.0271*</td>
</tr>
<tr>
<td></td>
<td>(0.0082)</td>
<td>(0.0045)</td>
<td>(0.0096)</td>
</tr>
<tr>
<td>7 years of tenure</td>
<td>0.0300*</td>
<td>−0.0012</td>
<td>0.0423*</td>
</tr>
<tr>
<td></td>
<td>(0.0091)</td>
<td>(0.0047)</td>
<td>(0.0100)</td>
</tr>
<tr>
<td>8 years of tenure</td>
<td>0.0419*</td>
<td>0.0013</td>
<td>0.0572*</td>
</tr>
<tr>
<td></td>
<td>(0.0106)</td>
<td>(0.0056)</td>
<td>(0.0124)</td>
</tr>
<tr>
<td>9 years of tenure</td>
<td>0.0548*</td>
<td>0.0068</td>
<td>0.0672*</td>
</tr>
<tr>
<td></td>
<td>(0.0126)</td>
<td>(0.0068)</td>
<td>(0.0145)</td>
</tr>
<tr>
<td>10 years of tenure</td>
<td>0.0616*</td>
<td>0.0127</td>
<td>0.0668*</td>
</tr>
<tr>
<td></td>
<td>(0.0155)</td>
<td>(0.0079)</td>
<td>(0.0161)</td>
</tr>
</tbody>
</table>

*An "*" indicates significance at the 5 percent level.
Figure 3: Estimated within-job wage profiles—jj-jobs and nj-jobs.
Figure 4: Estimated hazard functions for median workers—by exit states
<table>
<thead>
<tr>
<th>Transition type</th>
<th>Ed. 9-11</th>
<th>Ed. 12-14</th>
<th>Ed. 15-18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiation by jj transition (jj-job)</td>
<td>-0.1136*</td>
<td>0.1779*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0407)</td>
<td>(0.0618)</td>
<td></td>
</tr>
<tr>
<td>Exp. at job initiation (in years)</td>
<td>0.2394*</td>
<td>0.1966*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0202)</td>
<td>(0.0344)</td>
<td></td>
</tr>
<tr>
<td>Exp.² at job initiation (in years²)</td>
<td>-0.0235*</td>
<td>-0.0224*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0622)</td>
<td>(0.0045)</td>
<td></td>
</tr>
<tr>
<td>Job’s rank in employment-cycle</td>
<td>-0.0007</td>
<td>-0.1226*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0125)</td>
<td>(0.0200)</td>
<td></td>
</tr>
<tr>
<td>Tenure in prec. job (in years)</td>
<td>-0.1600*</td>
<td>-0.2478*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0200)</td>
<td>(0.0368)</td>
<td></td>
</tr>
<tr>
<td>Dur. of prec. u-spell (in weeks)</td>
<td>-0.0032*</td>
<td>0.0052*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0012)</td>
<td>(0.0015)</td>
<td></td>
</tr>
</tbody>
</table>

*A “*” indicates significance at 5 percent level.

Table 5: Predetermined mobility variables in job hazards

<table>
<thead>
<tr>
<th>Transition type</th>
<th>Ed. 9-11</th>
<th>Ed. 12-14</th>
<th>Ed. 15-18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected duration to exit by any transition (in weeks)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (nj-job, 12 weeks preceding unempl.)</td>
<td>62</td>
<td>63</td>
<td>127</td>
</tr>
<tr>
<td>B (jj-job, 2nd job in cycle, 9 mths. ten. in preceding job)</td>
<td>79</td>
<td>96</td>
<td>138</td>
</tr>
<tr>
<td>C (jj-job, 3rd job in cycle, 12 mths. ten. in preceding job)</td>
<td>86</td>
<td>102</td>
<td>151</td>
</tr>
<tr>
<td>D (jj-job, 4th job in cycle, 15 mths. ten. in preceding job)</td>
<td>94</td>
<td>108</td>
<td>163</td>
</tr>
</tbody>
</table>

| Expected duration to exit by job-to-job transition (in weeks) |           |           |           |
| A (nj-job, 12 weeks preceding unempl.)              | 88       | 73        | 146       |
| B (jj-job, 2nd job in cycle, 9 mths. ten. in preceding job) | 109      | 114       | 153       |
| C (jj-job, 3rd job in cycle, 12 mths. ten. in preceding job) | 115      | 118       | 163       |
| D (jj-job, 4th job in cycle, 15 mths. ten. in preceding job) | 120      | 124       | 172       |

| Expected duration to exit by job-to-nonempl. transition (in weeks) |           |           |           |
| A (nj-job, 12 weeks preceding unempl.)              | 413      | 1,875     | 1,496     |
| B (jj-job, 2nd job in cycle, 9 mths. ten. in preceding job) | 532      | 2,480     | 2,254     |
| C (jj-job, 3rd job in cycle, 12 mths. ten. in preceding job) | 668      | 3,012     | 3,162     |
| D (jj-job, 4th job in cycle, 15 mths. ten. in preceding job) | 833      | 3,636     | 4,424     |

Table 6: Expected job durations—by exit states
### Worker specific effects:
- Worker effects in wage equation ($\sigma_1$)  
  - Ed. 9-11: $0.1103^*$  
  - Ed. 12-14: $0.1079^*$  
  - Ed. 15-18: $0.1891^*$  
  - SE: $(0.0025)$  
  - Test Statistic: $(0.0031)$  
- Worker effects in job-to-job hazard ($\sigma_2$)  
  - Ed. 9-11: $0.5598^*$  
  - Ed. 12-14: $0.6045^*$  
  - Ed. 15-18: $0.2435^*$  
  - SE: $(0.0334)$  
  - Test Statistic: $(0.0267)$  
- Correlation coefficient ($\rho_{\sigma}$)  
  - Ed. 9-11: $-0.0893^*$  
  - Ed. 12-14: $-0.2198^*$  
  - Ed. 15-18: $-0.7647^*$  
  - SE: $(0.0460)$  
  - Test Statistic: $(0.0288)$  
- Job-to-nonemployment loading coefficient ($\xi$)  
  - Ed. 9-11: $1.3915^*$  
  - Ed. 12-14: $1.2749^*$  
  - Ed. 15-18: $3.3613^*$  
  - SE: $(0.1049)$  
  - Test Statistic: $(0.0704)$

### Match specific effects:
- Match effects in wage equation ($\omega_1$)  
  - Ed. 9-11: $0.1590^*$  
  - Ed. 12-14: $0.1630^*$  
  - Ed. 15-18: $0.1293^*$  
  - SE: $(0.0021)$  
  - Test Statistic: $(0.0019)$  
- Match effects in job-to-job hazard ($\omega_2$)  
  - Ed. 9-11: $0.7476^*$  
  - Ed. 12-14: $1.1928^*$  
  - Ed. 15-18: $0.9560^*$  
  - SE: $(0.0758)$  
  - Test Statistic: $(0.0512)$  
- Correlation coefficient ($\rho_{\omega}$)  
  - Ed. 9-11: $-0.2423^*$  
  - Ed. 12-14: $-0.1689^*$  
  - Ed. 15-18: $-0.0260$  
  - SE: $(0.0410)$  
  - Test Statistic: $(0.0191)$  
- Job-to-nonemployment loading coefficient ($\zeta$)  
  - Ed. 9-11: $0.5257^*$  
  - Ed. 12-14: $0.1326$  
  - Ed. 15-18: $0.6492$  
  - SE: $(0.1907)$  
  - Test Statistic: $(0.0859)$  

### Idiosyncratic effects in wage equation, $\sigma_\varepsilon$
- Ed. 9-11: $0.1369^*$  
- Ed. 12-14: $0.1281^*$  
- Ed. 15-18: $0.1309^*$  
- SE: $(0.0006)$  
- Test Statistic: $(0.0006)$

### Log-wage variance decomposition ($\sigma_1^2 + \omega_1^2 + \sigma_\varepsilon^2$)
- Worker specific effects ($\sigma_1^2$)  
  - 21.65%  
- Match specific effects ($\omega_1^2$)  
  - 44.99%  
- Idiosyncratic effects ($\sigma_\varepsilon^2$)  
  - 33.35%

"A "*" indicates significance at the 5 percent level.

Table 7: Unobserved heterogeneity\(^a\)

---

\(^a\)