The Economic Journal, **116** (*June*), F245–F285. © Royal Economic Society 2006. Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

WAGES, MOBILITY AND FIRM PERFORMANCE: ADVANTAGES AND INSIGHTS FROM USING MATCHED WORKER–FIRM DATA*

John M. Abowd, Francis Kramarz and Sébastien Roux

To illustrate the wide applicability of longitudinal matched employer-employee data, we study the simultaneous determination of worker mobility and wage rates using an econometric model that allows for both individual and firm-level heterogeneity. The model is estimated using longitudinally linked employer–employee data from France. Structural results for mobility show remarkable heterogeneity with both positive and negative duration dependence present in a significant proportion of firms. The average structural returns to seniority are essentially zero, but this result masks enormous heterogeneity with positive seniority returns found in low starting-wage firms.

In this article, we reconsider the relation between earnings and mobility using newly developed longitudinal matched employer-employee data. Our research is positioned at the intersection of labour economics and human resource management; however, the techniques we use and the ideas we examine have broad applicability. We develop this idea immediately.

Longitudinally linked employer–employee data can be characterised as follows. Labour markets are used as a motivating example. The population frame from which such data are created is a record of all formal jobs in the economy over a specific time period. A job consists of an association between in individual (worker) and an employing entity (firm[1]). The longitudinally linked data are constructed following jobs over time and by adding information from two additional population frames: workers and firms. Longitudinal information from both of these sources is integrated into the job frame. Then, an analysis sample is constructed based on individuals, jobs or employers according to the question under study.

Successful integration depends upon the records in the job frame containing a person and a firm identifier, which must also be used by the records in the individual and employer frames, respectively. For each job, the match between the worker and the employing firm is fully specified by these identifiers. A direct consequence of this design is that once the integrity of the identifiers is established, the job frame describes the complete graph connecting workers and firms. The analyst can study the worker at a particular job in relation to other workers at the same firm and in relation to other employers that worker has had. Identification of most of the critical components of

* We thank Steve Machin and two anonymous referees for extremely helpful suggestions. We also thank participants at various seminars and conferences, in particular at Crest, Insee (JMS), Stockholm University, European Community in Brussels, CNRS at Caen and ESSLE in Ammersee. The data used in this article are confidential but the authors' access is not exclusive. For additional information, contact F. Kramarz (kramarz@ensae.fr). J. Abowd acknowledges financial support from the National Science Foundation (grant SBER 96-18111 to the NBER and SES 99-78093 to Cornell University) and from the National Institute on Aging (R01-AG18854-01).

[1] The firm may be either a physical establishment or a legal entity. In the data we use herein, both measures are available but most of the economic data we use are measured for the legal entity. So, we constructed the integrated data at this level.

employment outcome heterogeneity is a direct consequence of connectedness of the workers and firms in this graph.

In some economies, for example Sweden, the person identifier on the job record is used in multiple data sources (on wages, employment, education, health, among others). In other economies, for example France, the firm identifier on the job record is used in multiple data sources (on performance, inputs, innovation, skill-structure, among others). This commonality of identifiers is necessary, but not sufficient, for building a useful integrated employer–employee data structure. The identifiers must also be relatively error-free and consistent over time. Finally, in order to specify the characteristics of any analysis sample, at least one of the three frames (job, individual, or firm) must be a universe or census.

Our description, which was customised for labour economists, can be fully adapted to most other fields in economics. Let us give a few examples.

In education, there are students and establishments (schools). A 'job' is the enrolment relation between a student and a school. Ancillary information on students, for example demographic characteristics and family background, are added from the student frame. Characteristics of the establishment, for example teacher identity and other inputs, are added from the establishment frame. An analysis sample of students consists of following them over time and across different schooling establishments. Students potentially share the same teacher and study in the same establishment. Many existing data sources, often of administrative origin, have this exact structure. The economic and statistical techniques that we use here can be directly adapted to education economics.

In health economics, there are patients and hospitals. A 'job' is the inpatient spell of a particular patient at a given hospital. Ancillary information on patients is integrated from the patient frame. Ancillary information on hospitals is integrated from the hospital frame. An analysis sample of patients can be followed over time as they are treated by different hospitals. In addition, a prominent input of medical care is physicians. Their association with particular hospitals or clinics is a characteristic matched from the hospital (or clinic) frame. There are doctors in medicine, both in these hospitals and in private practices. And, within hospitals there are various services with different specialties. Once again, most of the techniques that we use herein have direct applications to health economics.

In any of these, fields similar questions can be posed and similar techniques can be used to answer them. The techniques we illustrate are based on an analysis of variance with two or more high-dimensional, non-orthogonal effects. The questions can be paraphrased as follows: 'What is the contribution of worker's (student's; patient's) observables and unobservables to the variance of wages (grades in math or French; medical tests, such as the level of cholesterol in the blood, or the costs of treatments)?' and 'What is the contribution of the firm (school; hospitals) to this variance?'. Put differently, in education, is it the school that makes the student good or is it the student that explains most of the variation in grades? In health care, is it the hospital (or the doctor) with the associated treatments and medicines that makes the patient healthy and cheap to treat or is it the patient who is inherently healthy or sick? The answer to these questions is fundamentally identified by the longitudinally linked data provided that the analysis sample is sufficiently connected, in the graph-theoretic sense, to estimate the decomposition. We provide a detailed example of how to conduct this analysis for wages with worker and firm effects below.

A second question can be posed for all these fields. 'Why do workers (students; patients) leave a firm (a school; a hospital or a doctor)?' Is it because the firm (the school; the hospital) is bad (low profits, low pay, bad working conditions; bad teachers or environment; bad equipment and doctors) or is it because the worker is not productive (does not test well; is less healthy) in this environment?'.

We also demonstrate how to answer this second question using the worker–firm pair in this article. More precisely, we study the connections between firm-level compensation, promotion, retention policies and firm-level performance. We begin by relating a worker's inter-firm mobility to firm-specific compensation policies. Then, in our empirical analysis we use newly developed econometric methods and fully-integrated French employer-employee data to estimate some of the critical parameters of the models. As others have noted, particularly for France and the US, the results tend to show enormous individual and firm-level heterogeneity in compensation, promotion, and retention policies. We characterise this heterogeneity by modelling its joint distribution in the populations of individuals and firms. Finally, we recover some of the structural parameters of the firm-level policies such as the central tendency of the 'return to seniority' parameter.

The labour economics literature has attempted to measure the average return to seniority in models with limited heterogeneity. Abraham and Farber (1987) were the first to demonstrate that heterogeneity in the model for employment duration induced an upward bias in the measured average return to seniority, specifically, jobs with a longer expected duration were likely to be better-paying jobs and, therefore, longer seniority would be associated with higher pay but the return to an additional year of seniority, holding constant the expected duration of the job, was much smaller than the measured average return to seniority, ignoring expected job duration. Brown (1989) showed that the return to seniority is not constant; rather, it is higher during the first years of a job and diminishes to zero at the end of the employee's self-declared training period. In a series of articles, Altonji and coauthors, (Altonji and Shakotko, 1987; Altonji and Williams, 1992, 1997) applied various econometric techniques that attempted to remove the bias in the average return to seniority due to unobserved heterogeneity in individual job durations. These estimates, very much in the spirit of Abraham and Farber (1987), also indicated that the measured average return was upward biased and that the true return was closer to zero. In contrast, Topel (1991) used a model that included the possibility of bias arising from individual job search. This bias goes in the opposite direction of the job-duration heterogeneity bias leading Topel (1991) to consider both upward and downward biases. He concluded that the bias was downward in the uncorrected average return to seniority. More complete models of the sources of heterogeneity in the return to seniority lead to distributions of estimates that display individual, firm and within-firm heterogeneity as in, for example, Abowd, Kramarz, and Margolis, (1999, AKM hereafter), who find substantial heterogeneity in the returns to seniority in France (all of the previously cited papers used American data) with an average return of zero for men and women. More recent work by Margolis (1996) and Dostie (2005), using French data, confirm that simultaneous modelling of individual and firm-level heterogeneity produces estimates of the average return to seniority that are lower than the uncorrected estimates.[2]

A topic equally important but much less studied is the role of heterogeneity in firm compensation, promotion and retention policies in determining individual, employer, and match outcomes. As Lazear (1995) noted in great detail, measures of firm performance – such as productivity, value-added and profit – and measures of employment outcomes – such as retention, promotion and salary raises – ought to be related to choices made by workers and firms regarding compensation policies, work rules, and alternative employment offers. AKM studied both productivity and profit outcomes in relation to firm-specific measures of compensation policy and individual compensation heterogeneity. They found that high-wage firms were associated with higher productivity and higher profits. High-wage workers were associated with higher productivity but not higher profits. Other studies are reviewed in Abowd and Kramarz (1999) and Lazear (1995).

We begin by presenting our estimating equations in Section 1. Section 2 describes our data. Section 3 presents estimation results for compensation and mobility parameters that account for potential mobility and heterogeneity biases that have plagued some of the previous analyses. The first set of results, which properly accounts for each of the biases, delivers estimates of the central tendency of compensation parameters, such as the return to seniority and the structural job duration, that may be interpreted as structural in the same sense as the studies cited above. The second set of results, which relate measures of the compensation, promotion and retention policies to measures of firm performance, must be seen as descriptive – not causal – but given the dearth of serious evidence they should be indicative of potential directions for further research.

1. Estimating Equations for Wages and Mobility

In this Section we present our econometric model. We adopt a very flexible representation of the career process within the firm. A career is modelled as a succession of wages from the start to the end of the job.[3] Statistically, the observation of a wage within a firm has two components: the value of the wage and the presence of the worker in that firm. Thus, we represent the career process within a firm as a sequence of wages and presence indicators in this particular firm. A worker i entering firm j at calendar date $t_0(i)$ and leaving at date $t_f(i)$ has a career represented by:

$$\begin{pmatrix} w_{it_0(i)-1} = ., P_{it_0(i)} = 0; w_{it_0(i)}, P_{it_0(i)} = 1; w_{it_0(i)+1}, P_{it_0(i)+1} = 1; \\ \ldots; w_{it_0(i)+s}, P_{it_0(i)+s} = 1; \ldots, w_{it_f(i)}, P_{it_f(i)} = 1; w_{it_f(i)+1} = ., P_{it_f(i)} = 0 \end{pmatrix}$$

where $P_{it_0}(i)+s$ is an indicator function equal to 1 if the worker is employed in this firm j at this date and 0 otherwise, w_{it} denotes the wage the worker receives, and =. indicates that the value is not observed (for the sake of simplicity, we do not include the j subscripts).

[2] Other countries could exhibit downward biases rather than upward biases.

[3] A job is defined as a continuous history of presence within a firm. Precise assignment to tasks cannot be identified in the data that we use.

We are trying to be flexible in modelling the worker's career. Consequently, we do not derive a path for (w, P) from a preferred theoretical model.[4] Instead, we decompose the career process into two parts: a starting wage equation, that captures the external market effects, and a firm-specific model for wages and mobility, which depends upon the market effects from the starting wage equation. This firm-specific model and its estimation firm by firm is the core and originality of this article. The firm-specific parameters of the model will most often be interpreted as an expression of the firm-specific wage and retention policies.

As part of this approach, we do not account for firm-specific hiring policies, because the data do not tell us which workers contacted the firm but were not hired or which workers rejected a job offer. This is clearly a fascinating question but beyond our reach given the current data. Identification would be based entirely on prior assumptions.

1.1. *The Entry Wage Equation*

Let us first consider a worker's initial period at the firm. The entry wage can be decomposed into two components. The first component, which is individual-specific, depends on worker characteristics and their prices in the labour market. Whatever the firm the worker contacts, this component is interpreted as the mean of what that worker can expect from the labour market, given characteristics at that time and no firm-specific experience (tenure =0). It includes all elements of the worker that are transferable from one firm to another: experience, unobserved abilities and so forth. The individual-specific part of the wage rate will be denoted as the 'market' wage.

The second component of the wage rate corresponds to the job-specific part of the wage that results from the interaction of the worker and the firm. It should thus depend on the job-specific productivity and on shocks that may affect the firm. The firm-specific wage policy is also included in this component.

For each employment spell, the starting-wage equation at date $t_0(i)$, when worker i enters firm $j = J[i, t_0(i)]$, is given by

$$\log w_{it_0(i)} = \mathbf{X}_{it_0(i)}\boldsymbol{\beta} + \theta_i + \psi_{J[i,t_0(i)]} + \varepsilon_{it_0(i)} \tag{1}$$

$\mathbf{X}_{it_0}(i)$ denotes the variables describing the individual and the labour market that are time-varying but seniority invariant, θ_i denotes a person effect. Equation (1) also includes the initial component of the firm-specific compensation policy, $\psi_{J[i,t_0(i)]}$ at hire, as measured in the entry wage. The component $\varepsilon_{it_0(i)}$ is a zero mean error term reflecting, among other things, the initial productivity of the match. All person-specific elements are transferable among firms. Hence, worker i when moving from firm j to firm j' loses all elements of pay that are specific to firm j and receives at entry in firm j' the opportunity wage given by (1) using the new firm effect, ψ_j and at the calendar date $t'_0(i)$, the $\mathbf{X}$ variables at this calendar date.

As in Topel (1991), the returns to experience are directly estimated with this equation. For this reason, $\mathbf{X}_{it_0(i)}$ includes experience variables as well as other variables related to observed characteristics of the individual and market characteristics. In this

[4] The interested reader will find a theoretical rationale of our set of equations based on a very limited set of assumptions in a previous version of this work. This theoretical strategy is not the only way to get our reduced form equations presented below in this Section. Hence, we do not present this 'theory' here.

case, a bias may occur by not considering worker heterogeneity with respect to previous employment (past jobs).[5] To resolve this problem we introduce the number of previous jobs as an explanatory variable.

The starting-wage equation is estimated by full least squares based on the technique described in Abowd *et al.* (2003). These methods jointly estimate the fixed time-varying, individual and firm effects. A graph-theoretic algorithm is applied to produce the identifiable estimated person and firm effects. We include all observations that are at the beginning of a job (first year) for each worker. The coefficients $\hat{\boldsymbol{\beta}}, \hat{\theta}_i, \hat{\psi}_j$ are treated as known parameters for the firm-specific compensation policy estimates.

1.2. *The Firm-specific Model for Wages and Mobility*

Throughout the life of the job spell, worker i and firm j jointly decide whether to separate or to continue the match. In our approach, and given the available data, quits and layoffs are empirically identical. The worker's wage is observed after entry if and only if the worker and firm pair jointly decide to continue the match. This process is very much in the spirit of work by Jovanovic (1979), Flinn (1986), Topel and Ward (1992), Buchinsky *et al.* (2002) for micro-matching models or earlier work by Lillard and Willis (1978), Mincer and Jovanovic (1981) for the whole economy.

At date t for a worker with seniority s, i.e., $t = t_{0(i)} + s$, after subtracting the market wage (the sum of the effect of the market variables as measured by $\mathbf{X}_{it}\hat{\boldsymbol{\beta}}$ and of the individual fixed effect $\hat{\theta}_i$) the wage and mobility process can be expressed using the following equations:

$$
\begin{aligned}
R_{itj} &= \mathbf{1}_{R^*_{itj}>0} \\
R^*_{itj} &= \mathbf{Q}^{s(i,t)}_{itj}\boldsymbol{\gamma}_j + \boldsymbol{v}^{s(i,t)}_{ij} \\
\log w_{it} - \mathbf{X}_{it}\hat{\boldsymbol{\beta}} - \hat{\theta}_i &= \mathbf{Z}^{s(i,t)}_{itj}\boldsymbol{\beta}_j + \varepsilon_{it}
\end{aligned}
\tag{2}
$$

where we substitute $j = J(i, t)$ for notational clarity, R^*_{itj} is a latent variable expressing mobility out of firm j at date t when positive ($R = 1 - P$, using the above definition), $\mathbf{Q}^{s(i,t)}_{itj}$ is a vector of seniority-dependent variables that affect the separation decision, quit or layoff, γ_j is the firm-specific parameter vector describing the dependence of the separation decision on $\mathbf{Q}$, and $v^{s(i,t)}_{ij}$ is a mean zero error term reflecting productivity shocks. For the wage equation, $\mathbf{Z}^{s(i,t)}_{itJ(i,t)}$ denotes the variables indexed on seniority in the firm and and ε_{it} denotes a statistical residual.

Let us consider the system (2) at the entry into the firm. Since the worker is observed in the firm at the first period, there is no choice for this variable to be negative, and the first observed wage is the entry wage.

$$
\begin{aligned}
R_{it_{0(i)}j} &= 0 \\
\log w_{it_{0(i)}} - \mathbf{X}_{it_{0(i)}}\hat{\boldsymbol{\beta}} - \hat{\theta}_i &= \psi_j + \varepsilon_{it_{0(i)}}.
\end{aligned}
$$

Hence, the wage equation in this system is fully consistent with the entry wage equation (1) since all variables $\mathbf{Z}^{s(i,t)}_{itJ(i,t)}$ are equal to zero when s is zero, at entry, except for the

[5] For instance, Dustmann and Meghir (2003) restrict their estimation of such an equation to displaced workers.

intercept, ψ_j. For any firm j, because all workers used in the estimation have zero seniority at entry, there is no mobility equation at the date of entry.

Our system makes job seniority endogenous, as noted by many authors (Abraham and Farber, 1987; Altonji and Shakotko, 1987; Altonji and Williams, 1992, 1997, among others). Better workers may leave the firm because the firm is low-wage and they received a better outside offer. Others might leave when the firm considers that they are overpaid (given their productivity) and imposes a wage reduction. Should the first argument be true, selection within the firm leads to under-sampling the best workers. The second argument leads to over-sampling them. These types of selection biases have potentially severe consequences on the wage equation estimates.

The econometric identification of the selection process relies on the specification and estimation of the correlation matrix between the various residuals together with exclusion restrictions among the different equations (variables present in the wage equation and absent from the mobility equations, and conversely).[6]

For example, a worker hired at date $t_0(i)$, who stays exactly two periods in the firm is modelled with the following mobility and wage equations:

$$\begin{aligned} R^*_{it_0(i)+1j} &= \mathbf{Q}^1_{it_0(i)+1j}\gamma_j + v^1_{ij} < 0 \\ \log w_{it_0(i)+1} - \mathbf{X}_{it_0(i)+1}\hat{\boldsymbol{\beta}} - \hat{\theta}_i - \hat{\psi}_j &= \mathbf{Z}^1_{it_0(i)+1j}\boldsymbol{\beta}_j + \varepsilon_{it_0(i)+1} \\ R^*_{it_0(i)+2j} &= \mathbf{Q}^2_{it_0(i)+2j}\gamma_j + v^2_{ij} < 0 \\ \log w_{it_0(i)+2} - \mathbf{X}_{it_0(i)+1}\hat{\boldsymbol{\beta}} - \hat{\theta}_i - \hat{\psi}_j &= \mathbf{Z}^2_{it_0(i)+2j}\boldsymbol{\beta}_j + \varepsilon_{it_0(i)+2} \\ R^*_{it_0(i)+3j} &= \mathbf{Q}^3_{it_0(i)+3j}\gamma_j + v^3_{ij} > 0. \end{aligned} \tag{3}$$

In order to model the statistical relations between past wages and mobility or, similarly, between future wages and mobility, we assume that the following correlation structure holds:

$$\begin{matrix} v^1_{ij} \\ \varepsilon_{it_0(i)+1} \\ v^2_{ij} \\ \varepsilon_{it_0(i)+2} \\ v^3_{ij} \end{matrix} \rightsquigarrow \mathrm{N}\left(\mathbf{0}, \begin{bmatrix} 1 & \rho_{1j} & 0 & 0 & 0 \\ \rho_{1j} & \sigma^2_j & \rho_{2j} & 0 & 0 \\ 0 & \rho_{2j} & 1 & \rho_{1j} & 0 \\ 0 & 0 & \rho_{1j} & \sigma^2_j & \rho_{2j} \\ 0 & 0 & 0 & \rho_{2j} & 1 \end{bmatrix}\right).$$

A simple rewriting of the correlation matrix based on the normality assumption, presented in Appendix A, is useful for estimation since the likelihood does not involve multiple integration of the normal distribution.

The correlation matrix above accounts for the possible correlation between mobility residuals, v^t_{ij} and both past and contemporaneous wage residuals, ε^{t-1}_{ij} and ε^t_{ij}, respectively. We do not allow for a direct correlation between past and present wage residuals since such a correlation is already captured by the person effect.[7]

The crucial point in our approach should now be clear: all parameters of the wage and mobility equations, apart from the starting-wage equation, are

[6] Even though the normality assumption suffices for identification.

[7] Recall that the person effect is removed from the log-wage variable as shown above.

firm-specific. For instance, in the wage equation there are returns-to-seniority parameters that are similar for all workers employed in the firm and may differ from the returns-to-seniority parameters estimated in other firms. In addition, the firm-specific returns to seniority are allowed to vary with the sex and education of the workers in that firm. More generally, since the estimation is done separately for each firm, the estimated parameters can be used to characterise the promotion and retention policies of the firm.

There are two kinds of exclusion variables included in the mobility equation (and not in the wage equation). The first one is classic, it captures measurable individual heterogeneity with respect to previous employment spells. These effects are modelled using the number of previous jobs and the duration of the most recent previous job. The other group of variables in the mobility equation, but not the wage equation, reflects the position of the worker in the age distribution at entry in the firm. This variable is inspired by the literature on internal labour markets (Doeringer and Piore, 1976). According to these theories, some firms have restricted entry ports. In such firms, hiring (entry to the firm) occurs at a young age and is associated with a particular career path at the firm. Workers in the lower part of the age distribution at entry should be less mobile on this theory. By contrast, when firms hire workers at various ages and when the worker-firm pair is concerned about the quality of the match, one expects to see more separation among the workers who are younger at hire. In each case, the position of the worker in the age distribution at entry is a good candidate for a variable to be to identify mobility – it affects the mobility process without directly interfering with the wage formation process.

The firm-specific mobility and wage process requires enough within-firm variability for the effect of each of the explanatory variables to be identified. For this reason, estimation is conducted in firms with sufficient observations. We set a minimum within-firm sample size of 200 observations. We use maximum likelihood estimation firm by firm. Parameters can only be estimated when there is enough within-firm variation in workers' observed characteristics. For instance, there is no point in estimating a male-specific mobility intensity when the firm only comprises males. Hence, we use the following strategy. Before estimation, we automatically locate all firms for which one of our explanatory variables has insufficient within-firm variation. This variable is withdrawn from the explanatory variables of the relevant firm-level equation. Of the above 4,000 firms, 45% have at least one such variable. For some variables, such as the male indicator in the mobility equation, the coefficient is identified for more than 99% of the firms. However, the proportion falls to 82% for a variable such as the returns to seniority for workers with 10 years or more of seniority. The coefficient is then set to zero in the firm-level analyses and to 'missing value' in Tables 1 to 6 describing the parameters.

2. Data Description

We use data from the Déclarations Annuelles des Données Sociales (DADS), a 1/25th sample of the French work force with information from 1976–96 on the matched worker-firm side and data from the BRN on the performance of the firm side. We describe these data in turn.

2.1. *The DADS*

The DADS are a large collection of matched employer–employee information collected by INSEE (Institut National de la Statistique et des Etudes Economiques). The data are based on a mandatory employer report of the gross earnings of each employee subject to French payroll taxes. The universe includes all employed persons. Our analysis sample covers all individuals employed in French enterprises who were born in October of even-numbered years, excluding civil servants. Our extract runs from 1976 to 1996, with 1981, 1983 and 1990 excluded because the extracts were not created for those years. The initial data set contained 16 million observations, each corresponding to a unique enterprise-individual-year combination. The observation includes an identifier that corresponds to the employee (called NNI below) and an identifier that corresponds to the enterprise (SIREN). For each observation, we have the exact starting date and end date (day of the year and year) of the job spell in the establishment and an indicator for full-time/part-time/intermittent/at home work-status of the employee. Each observation also includes, in addition to the variables listed above, the individual's sex, month year and place of birth, current occupation, total net nominal earnings during the year, annualised gross nominal earnings during the year for the individual. Employer characteristics are the location and industry of the employing establishment.

2.1.1. *Observation selection, variable creation, data editing and imputation*

An observation is identified by a combination of two identifiers, the firm id and the person id. The SIREN number has an internal structure that allows us to check for coding errors. The NNI number has no such internal control. Even though 90% of recently-submitted DADS files are sent by the responding firm using electronic media, the situation in the 1980s was quite different. At that time INSEE had to perform the data entry from paper records. Data entry errors in the NNI occurred (for example, exchanging two digits of the NNI, error in one of the digits and so forth). This phenomenon is well-known at INSEE but, despite many attempts, no general way of solving this problem was found.[8] As a consequence, some observations have a NNI-year-SIREN combination such that no other observation has the same NNI. As a joint product, some NNI-SIREN combinations have a unique missing year. Consider now the case of a worker with observations in, say, 1978 and 1980 in the same enterprise (SIREN) but no observation for 1979. To be true, this would mean that the worker would be employed until some date in 1978 (depending on the number of days worked, December 31 most likely) and also employed after some date in 1980 (depending on the number of days worked, January 1 most likely) in this firm but not employed at all during year 1979. This is very improbable because of the regulations governing layoffs in France, in which workers may be recalled by their previous employers even after some period of unemployment. The suggestions of D. Verger (chief of the Division Revenus, in charge of the DADS at the beginning of the 1990s) led us to adopt the following solution. Whenever an observation was missing in a given year when the same NNI-SIREN combination existed for the preceding and the following year, we created an observation for the missing year with the same NNI-SIREN combination. (This added 193,148 observations.) Earnings are

[8] The computer intensive method of Abowd and Vilhuber (2005) was not an option when the data for this article were created.

computed as the geometric mean of the preceding and following wages (in real terms). All other variables are taken at their preceding year value.

Because of the 1982 and 1990 Census, the 1981, 1983 and 1990 DADS data were not available. We used the same principle as the one described above to impute missing observations. Hence, imputation was performed only for those individuals that were present in the same firm in 1980 and 1982 or 1982 and 1984 or 1989 and 1991. (This added 759,017 observations to the sample.) All variables were imputed as above. A more precise description of the construction of some variables in the data set are presented in the working paper version of this article.

Finally, as in AKM, we eliminated observations for which the logarithm of the real annualised total compensation cost was more than five standard deviations away from its predicted value based on a linear regression model of this variable on sex, region, experience, and education (see the data appendix in AKM).

After these selections and imputations, the final data set contains 13,770,082 observations, corresponding to 1,682,080 individuals and 515,557 firms.

2.1.2. *Seniority imputations for 1976*

To estimate the starting wage equation (1), the wage equation for observations with zero seniority, we concentrate on all job observations for the year of the individual's entry into the firm. Because of the left-censoring in the creation of the DADS, there is a problem with observations in 1976 – no information is available about the actual start date for individuals who were already employed at a particular firm at the start of this year. We use AKM's imputed seniority as an indicator of recent hire or not. We assumed that all workers with strictly less than two years of imputed seniority in 1976 'just entered' their employing firm. Since short-term contracts were virtually non-existent in those years, most workers had relatively long employment spells.[9] After selecting only those observations corresponding to the year of entry in new job, we have 4,616,093 observations. These include 1,535,758 individuals (some persons are only employed by their 1976 employer and never leave it) and 480,360 firms (some firms only employ workers with 2 years of seniority in 1976; in general these are very small firms with only one DADS individual).

2.2. *The BRN and Other Firm-level Variables*

2.2.1. *Within-firm employment, seniority and skill structure*

To measure the employment structure of our French firms, we construct firm-level variables based on some of our entry wage equation estimates. First, we take the estimated $\hat{\boldsymbol{\beta}}$ from the entry wage equation to generate the variable $\mathbf{X}\hat{\boldsymbol{\beta}}$ for each person at entry to the firm. Then, we generate the various quartiles of the within-firm distribution of this variable. These quartile boundaries are used to control for a mixture of the experience and time-varying skill structure of the firm since they are based on a univariate measure derived from the market wage. Similarly, we compute the quartiles of the within-firm distribution of the estimated person-effect, an effect which is the sum of

[9] Other definitions of 'just entered' in 1976 did not materially change in the estimated coefficients of the starting wage equation.

an unobserved component and of the time-invariant observed characteristics (education and sex).

Finally, and reminiscent of the dual labour market literature (Doeringer and Piore, 1976), we compute the within-firm distribution of age of the workers at entry. The idea is that some firms may hire most of their workers after they completed education (internal labour markets) whereas other firms hire workers at all ages.

2.2.2. *Indicators of performance*

Indicators of performance are drawn from an administrative firm-level data source, the annual report on profitability and employment by enterprises (Bénéfices Industriels et Commerciaux, BIC) for all firms reporting under the BRN regime (Bénéfices Rééls Normaux), the usual reporting regime for most firms (only the very small firms, in particular the zero employees firms, are excluded). The data cover 1985 to 1996. The variables listed below, and described extensively in AKM, are averaged over the period for all available years from the BRN.

Variables that we will consider in the following are:

- Firm-level employment (expressed in thousands of workers),
- Real capital stock per employee (defined as total assets divided by the industry-specific price index of physical capital per employee) also called the capital-labour ratio,
- The fraction of stable employees (measured as the within-firm average probability of being employed next year conditional on being employed this year in the firm) also called the stability index,
- Real value added inclusive of labour costs per employee, and
- Operating profit per unit of capital.

3. Estimation Results

3.1. *The Starting-wage Equation*

Estimation of the starting wage equation (1) requires some identification assumptions. The individual and firm effects are identified, relative to the overall constant, in fixed-effect estimation when the observations are part of a connected group. Ancillary assumptions are required to identify the person effects, firm effects, and in some cases the residual. First, in the case where the individual has at least two for the many small disconnected groups in the French data.[10] First, the individual effect is estimated separately from the residual only for workers whose entry wages are observed for at least two jobs. Second, for those workers who are observed in a single job, the residual is set to zero and the individual effect is set at the non explained part of the wage as a consequence. Because we focus hereafter on relatively large firms for which at least 200

[10] The largest connected group covers more than 95% of the total. The many small disconnected groups are a consequence of the data being a 1/25th sample. We identified the person effects in the disconnected groups by setting their within-group average to zero. We identified the firm effects in the disconnected groups by including the group mean in the firm effect and setting the overall average firm effect to zero.

observations are available, the estimated firm fixed effect is relatively well estimated; see Abowd *et al.* (2003) for the standard errors formulae.

The explanatory variables ($\mathbf{X}_{it}$) are actual labour force experience (up to a quartic), an Ile de France region indicator, a full-time vs part-time indicator, and year indicators. These variables are fully interacted with sex. To control for the endogenous number of starting-wage observations (some workers churn more and, therefore, start more jobs), we control for the past number of jobs in the equation (using indicator variables). In addition, we include both a person and a firm effect. Since both person and firm effects are high-dimensional, standard estimation techniques, which rely on matrix inversions or sweep operators, cannot be applied. Instead, we use a minimisation technique based on the standard conjugate gradient algorithm that solves the OLS normal equations directly; see Abowd *et al.* (2003).

Starting wage results are presented in Appendix B. Since this regression controls for fixed person and firm effects, the reported coefficients identify the effects of changes in the time-varying attributes. Results show that the returns to experience are larger for men than for women. In addition, the first job pays less than all subsequent jobs. Not surprisingly, jobs in the Ile de France region are better paid. Finally, the time indicators show a decreasing trend in real starting wages, reflecting the worsening of entry conditions for young workers in the French labour market. The starting wage equation can also be viewed as an attempt to capture initial heterogeneity, at the date of the worker's hire into the firm, in the spirit of Heckman (1981).

3.2. *The Firm-specific Mobility and Wage Equations*

For each of the 5,000 firms in the sample for which there are enough individual observations to proceed, we estimate the set of equations described in (3) by maximum likelihood. Convergence occurs for 4,015 firms using the automated maximisation programs. We did not try to re-estimate models for those firms where the algorithm did not converge.[11] Firm-specific parameters can only be estimated when there is enough within-firm variation in workers' observed characteristics (see the discussion in Section 1).

The firms used in these estimates are not representative of all French firms in two ways. The estimation sample over-represents large firms, which tend to survive longer, because of our requirement that at least 200 worker-year observations per firm be available. In addition, firm-specific observations may correspond to many different workers in a few years or to fewer workers observed for more years. Thus the within-firm composition of the sample reflects the average completed duration of job spells at that firm. We can assess these potential sources of non-representativeness in our firm sample by comparing our firms with the universe of firms contained in the BRN, our

[11] In previous versions of this research, we succeeded in obtaining convergence for 95% of firms after multiple attempts but none of the subsequent results were changed by inclusion of these firms. In addition, we were less confident in some of the coefficients obtained for those firms where convergence was obtained in the supplementary searches. For these reasons, we decided to restrict attention, in this article, to firms for which the maximum likelihood procedure converged immediately using a grid search for the starting values of the correlation coefficients ρ_{1j} and ρ_{2j}.

sample covers 36% (30% in the BRN) of the workers working in the private sector in 1985 (1996) and 83% (75%) of firms employing more than 1,000 workers.

We turn now to interpretation of the estimated firm-specific determinants of wage growth and mobility. We have estimated values for the effects of different worker characteristics for each firm. These estimates require some caution in interpretation. Our main interest lies in the between-firm heterogeneity of these effects. Indeed, if the effects could be ascertained by direct measurement (e.g., a firm survey), then, the data would directly reveal the between firm variability. Of course, direct survey measurement of a concept like the rate of increase of mobility as a function of seniority is fraught with its own difficulties – justifying our reliance on indirect statistical measures.

The firm-specific effects of interest in our work must be estimated from samples of workers at the firm. Hence, all observed variation in the estimated effects confounds both between-firm and within-firm variation. The between firm variation reflects differences in firm-level policies. The within-firm variation is estimation error arising from the incompleteness of our samples or the inadequacy of our models. Although it will not be easy to make inferences about the between-firm differences, we will try to do this using three methods. First, we summarise the between-firm heterogeneity in the estimated effects (Tables 1 and 2) while also drawing attention to the heterogeneity in the statistical precision of the estimated effects (Tables 3 and 4). Second, we simulate the nonlinear within-firm effects, in effect treating the within-firm parameter estimates as modes of the posterior distribution of the parameters. The simulations allow us reduce the influence of the estimation error on inferences about between-firm heterogeneity by averaging the heterogeneity measures over many simulated firm-level outcomes. Finally, we analyse the between-firm correlations of the estimated effects using a method that explicitly adjusts the covariance matrices for estimation error in the firm-level effects. While focusing on the dispersion of measures of statistical significance is not a conventional technique for summarising multi-level models like ours, it is instructive because it shows the reader the hazard of concluding that large differences in estimated effects necessarily arise from between-firm heterogeneity – estimation error within firms is also important.

We begin with the first assessment, direct interpretation of the distribution of firm-specific effects and their estimation error. The distribution of estimated firm-specific effects for the mobility equation is provided in Table 1. The distribution of the firm-specific effects for the wage equation is in Table 2. Tables 3 and 4 show the distribution of the conventional measure of statistical significance (the Student-*t*) for the effects in the mobility and wage equations, respectively.

Most parameters are easy to interpret. In both the mobility and wage growth equations, seniority effects are estimated as splines. For the mobility equation, the spline is based on the following functional form for the probability of moving (leaving the firm):

$$\begin{aligned}\Pr(\text{Moving}|sen) = F\{&a_1 sen \times \mathbf{1}(0 \leq anc < 2)\\ &+ [2a_1 + a_2(sen - 2)] \times \mathbf{1}(2 \leq anc < 5)\\ &+ [2a_1 + 3a_2 + a_3(sen - 5)] \times \mathbf{1}(5 \leq anc < 10)\\ &+ [2a_1 + 3a_2 + 5a_3 + a_4(sen - 10)] \times \mathbf{1}(10 \leq anc)\}\end{aligned}$$

where $\mathbf{1}(\cdot)$ denotes an indicator function and *sen* denotes seniority. The spline function is linear, continuous everywhere, with changing slopes at 2, 5, and 10 years of seniority.

Table 1

Distribution of the Estimated Firm-specific Parameters for the Mobility Equation

	Mean	Std. Dev.	q1	q5	q25	q50	q75	q95	q99
Male	0.0161	0.3941	−0.7923	−0.3597	−0.0912	0.0056	0.1364	0.4055	0.8390
Part-time	−0.3628	0.6742	−1.7191	−1.0100	−0.5661	−0.3320	−0.1141	0.2063	0.8001
First period constant (1976–80)	−0.2607	1.2618	−3.1858	−1.9496	−0.8989	−0.3854	0.3271	1.5661	2.9530
Second period constant (1982–9)	0.0382	6.3362	−2.6178	−1.3875	−0.4819	−0.0323	0.8422	1.9353	3.5082
Third period constant (1991–6)	0.0363	6.4468	−2.9012	−1.4282	−0.4572	0.0133	0.8102	1.7774	3.2716
Person effect (from starting wage equation)	−0.1035	0.2629	−0.8511	−0.4912	−0.2221	−0.0890	0.0168	0.2669	0.6697
Experience	−0.0083	1.1331	−0.4069	−0.2149	−0.0743	0.0005	0.0478	0.1644	0.3339
$(\text{Experience}/10)^2$	0.0712	2.7688	−3.4281	−1.5255	−0.6348	−0.0146	0.6320	2.3532	4.6965
$(\text{Experience}/100)^3$	−0.0553	0.7575	−2.0674	−0.9559	−0.2472	−0.0039	0.2116	0.5314	1.3101
$(\text{Experience}/1000)^4$	0.0010	1.6188	−0.1587	−0.0545	−0.0199	0.0041	0.0358	0.1307	0.2925
Tenure (less than two years)	0.1796	24.8821	−0.8842	−0.4720	−0.2315	−0.0307	0.2385	1.2529	2.5486
Tenure (2 to 5 years)	0.3839	24.6520	−1.1096	−0.4240	−0.0601	0.0619	0.2156	0.7018	1.6256
Tenure (5 to 10 years)	0.1683	19.0219	−0.9527	−0.2717	−0.0691	0.0035	0.1687	0.7267	1.5974
Tenure (more than 10 years)	0.1788	4.4410	−0.4514	−0.0377	0.0374	0.1266	0.2716	0.7433	1.3999
Tenure × Male	−0.0074	5.0466	−0.4469	−0.1348	−0.0295	−0.0007	0.0147	0.1119	0.4647
Tenure × Low general education	0.1774	31.1999	−1.1699	−0.4384	−0.1318	−0.0498	−0.0190	0.2843	1.4458
Tenure × High general education	0.1037	17.5362	−1.4725	−0.5356	−0.0853	−0.0233	0.0409	0.4159	1.6970
Low general education	0.2575	1.8298	−1.8120	−0.7692	0.0170	0.2116	0.5095	1.2942	2.4651
High general education	0.3246	1.4654	−2.2733	−0.8746	−0.0095	0.2616	0.6040	1.5403	3.3289
Entry in first quartile of the age at entry distribution	0.4194	0.6432	−1.2314	−0.4548	0.1124	0.4609	0.6852	1.3098	2.1209
Entry in second quartile of the age at entry distribution	0.3016	0.5181	−1.1092	−0.3991	0.0824	0.3016	0.5058	1.0220	1.7105
Entry in third quartile of the age at entry distribution	0.1996	0.3867	−0.8599	−0.3132	0.0384	0.2111	0.3659	0.7134	1.1937
Number of previous jobs	0.0522	1.0663	−0.1097	−0.0145	0.0102	0.0277	0.0565	0.1646	0.3287
Seniority in the previous job	−0.1936	0.4149	−1.3476	−0.6890	−0.1972	−0.0930	−0.0670	−0.0392	0.0030

There is no overall intercept in the equation. Between-firm distribution of the estimated parameters for the mobility equation. The model is estimated by maximum likelihood, jointly with the equation in Table 2, separately for each firm. For each firm in the sample, there is a set of estimated parameters used to compute the distribution. Parameters are only estimated for those firms in which there is enough within-firm variability. Number of observations (firms): 3,951. Sources: DADS.

Table 2

Distribution of the Estimated Firm-specific Parameters for the Wage Equation

	Mean	Std. Dev.	q1	q5	q25	q50	q75	q95	q99
Part-time	0.0344	0.4644	−0.7983	−0.5281	−0.2270	−0.0583	0.1978	0.9165	1.5304
First period constant (1976–80)	0.0138	0.6353	−1.9025	−1.0553	−0.2934	0.1022	0.3829	0.8695	1.2988
Second period constant (1982–9)	−0.0151	0.7096	−1.9277	−1.6657	−0.3627	0.1234	0.3943	0.9038	1.2977
Third period constant (1991–6)	0.0084	0.7198	−2.2397	−1.3609	−0.3812	0.1683	0.4382	0.9387	1.4081
Tenure (less than two years)	0.0391	1.5918	−0.3602	−0.1425	−0.0243	0.0138	0.0674	0.2227	0.4025
Tenure (2 to 5 years)	0.0014	0.2508	−0.3470	−0.1297	−0.0323	0.0003	0.0318	0.1393	0.3395
Tenure (5 to 10 years)	0.0056	0.2556	−0.2956	−0.0953	−0.0163	0.0032	0.0266	0.1219	0.3143
Tenure (more than 10 years)	0.0189	0.3298	−0.2742	−0.0721	−0.0078	0.0145	0.0338	0.1343	0.4395
Male	0.0212	0.1438	−0.3919	−0.1822	−0.0324	0.0184	0.0771	0.2226	0.4771
Tenure × Male	−0.1544	7.7711	−0.1321	−0.0434	−0.0115	−0.0017	0.0081	0.0448	0.1280
Low general education	−0.0005	0.3524	−0.9861	−0.4865	−0.1353	0.0148	0.1178	0.4481	1.0047
Tenure × Low general education	−0.0211	1.9502	−0.3320	−0.1073	−0.0206	−0.0040	0.0138	0.1113	0.3698
High general education	0.0074	0.4688	−1.4712	−0.6056	−0.1625	0.0059	0.1584	0.6571	1.3720
Tenure × High general education	0.1302	11.9309	−0.4659	−0.1535	−0.0125	0.0082	0.0391	0.1695	0.5464
Correlation between mobility and future wage	−0.3139	1.2977	−4.8606	−3.6833	−0.0743	0.0048	0.0650	0.3359	2.6381
Correlation between mobility and past wage	−0.0260	0.1637	−0.5884	−0.2421	−0.0747	−0.0237	0.0315	0.1661	0.3723
Standard error of the wage shock (in log)	−0.9099	0.8357	−2.4347	−2.0403	−1.4549	−1.0676	−0.5762	0.9425	1.1945

Notes. There is no separate intercept in the equation. Between-firm distribution of the estimated parameters for the wage equation. The model is estimated by maximum likelihood, jointly with the equation in Table 1, separately for each firm. For each firm in the sample, there is a set of estimated parameters used to compute the distribution. Parameters are only estimated for those firms in which there is enough within-firm variability. Number of observations (firms): 3,951. Sources: DADS.

Table 3

Distribution of the Estimated Student-t Statistics for the Firm-specific Effects in the Mobility Equation

	Mean	Std. Dev.	q1	q5	q25	q50	q75	q95	q99
Male	−0.8117	4.6330	−21.0560	−5.1979	−1.0997	0.0320	0.9781	2.6651	4.5846
Part-time	−4.2133	7.7969	−34.6720	−18.2803	−4.6859	−2.0488	−0.6292	1.4457	8.2593
First period constant (1976–80)	−0.6768	8.3624	−22.4660	−13.2404	−3.1356	−0.9161	0.5738	19.7934	29.2288
Second period constant (1982–9)	0.7344	6.4521	−10.0240	−9.2332	−1.4600	−0.0497	1.4135	14.7496	26.3768
Third period constant (1991–6)	1.0553	6.1903	−7.7360	−6.3767	−1.5153	0.0054	1.5007	20.1132	22.0264
Person Effect (from starting wage equation)	−1.7315	3.5785	−11.1400	−9.0843	−2.9645	−1.1439	0.3158	2.5961	7.9233
Experience	0.4688	3.0357	−8.7410	−3.3827	−1.0995	0.0000	1.3716	6.9150	9.8025
$(Experience/10)^2$	−1.0354	4.0817	−15.3340	−11.9717	−1.2908	−0.0230	1.0722	2.9832	4.9620
$(Experience/100)^3$	1.3839	5.0645	−4.3750	−2.5297	−1.0507	−0.0410	1.2761	13.8944	21.0915
$(Experience/1000)^4$	−1.2024	5.2846	−22.1950	−13.3020	−1.1844	0.1353	1.2377	2.9442	4.7710
Tenure (less than two years)	−1.4053	6.9969	−25.4720	−19.3344	−2.2533	−0.2008	1.2807	9.3060	12.9560
Tenure (2 to 5 years)	1.0196	3.9677	−10.1230	−4.3250	−0.5190	0.6255	2.5344	7.2990	10.8282
Tenure (5 to 10 years)	−0.0788	4.3181	−10.5820	−8.0322	−1.6305	0.0210	1.5431	8.5697	11.9530
Tenure (more than 10 years)	3.3559	3.4848	−1.3840	−0.2945	0.9545	2.3665	5.1823	13.0720	14.2948
Tenure × Male	0.0035	1.7796	−4.8060	−2.6630	−1.0286	−0.0119	0.9155	2.4542	5.2250
Tenure × Low general education	−1.7749	2.3458	−7.7430	−7.3194	−2.7654	−1.3170	−0.2316	0.9771	2.1896
Tenure × High general education	−0.6063	1.9080	−6.6810	−3.7927	−1.3293	−0.2803	0.5306	1.8786	2.8912
Low general education	1.3314	2.0393	−2.3000	−1.1336	0.0397	0.9830	2.1132	4.4944	8.9120
High general education	1.5891	2.7773	−2.3580	−1.3711	−0.0076	0.9414	2.5382	9.7713	11.3924
Entry in Q1 of the age at entry distribution	3.0635	4.4751	−3.9560	−1.5673	0.3604	1.7155	4.0094	14.7924	15.1124
Entry in Q2 of the age at entry distribution	2.5399	3.3882	−1.9040	−0.9300	0.3576	1.5074	3.3930	10.8439	11.6351
Entry in Q3 of the age at entry distribution	2.6000	4.1588	−2.0480	−0.9689	0.3372	1.2474	2.8080	14.7258	16.3312
Number of previous jobs	2.4686	2.6466	−2.9220	−0.7917	0.7086	1.9461	3.9635	6.4336	10.5123
Seniority in the previous job	−23.0341	30.8499	−162.1100	−70.0884	−28.7513	−11.0840	−5.6677	−1.8674	0.6994

Notes. Between-firm distribution of the estimated Student-t statistics for the mobility equation. See Table 1 for additional notes.

For instance, a positive coefficient for a_1, as reported in Table 1, means that workers are more likely to move as seniority increases from zero to two years. After two years of seniority, the coefficient a_2 applies and measures the difference (either increasing or decreasing, depending on its sign) from the coefficient a_1. For instance if a_2 is not significantly different from zero, it means that workers with two to five years of seniority have the same probability of leaving the firm as workers with no more than two years of seniority.

Similarly, we implement spline for seniority in the wage equation as follows:

$$\begin{aligned} \mathrm{E}(\ln \mathrm{Wage}|sen) = {} & b_1 sen \times \mathbf{1}(0 \leq anc < 2) \\ & + [2b_1 + b_2(sen - 2)] \times \mathbf{1}(2 \leq anc < 5) \\ & + [2b_1 + 3b_2 + b_3(sen - 5)] \times \mathbf{1}(5 \leq anc < 10) \\ & + [2b_1 + 3b_2 + 5b_3 + b_4(sen - 10)] \times \mathbf{1}(10 \leq anc) \end{aligned}$$

where, once again, a positive b_1 means that at the beginning of their careers in the firm workers' wages increase with seniority; and b_2 picks up where b_1 stops: if b_2 is greater than b_1 returns to seniority increase after two years of seniority and if it is less than b_1 then returns to seniority decrease after two years (these returns can be flat when b_2 is equal to zero or may even decrease if $b_2 < 0$).

3.2.1. *Heterogeneity in the mobility equation*

The mobility equation measures the probability of separating from the employer. A positive coefficient corresponds to a larger propensity to move. Consider first the base propensity of firms to separate workers. Heterogeneity of this propensity is captured by between-firm differences in the intercept of the mobility equation, which we estimate separately for three distinct sub-periods: 1976–80, 1982–9 and 1991–6. Within firms, the estimated sub-period constants are very highly correlated (not shown in Table), therefore, it is reasonable to interpret them as reflecting a long-term component of the firm's mobility policy. A large positive sub-period constant means that the firm has many separations or, equivalently, that it is a high-turnover firm. Interestingly, results in Table 1 show that baseline mobility is very heterogeneous. While this could be due to between-firm differences or within-firm estimation error, Table 3 confirms that 35% of firms are low-turnover firms whereas 10 to 20% of firms are high-turnover firms.

This retention policy of firms also depends on worker types. For instance, we see that the worker's sex appears to matter for less than 30% of firms. But, for roughly 20% of firms, males tend to stay significantly longer periods than females whereas the opposite appears to hold for about 10% of firms. For the rest of firms, there is no measured difference in the mobility behaviour of men and women. By contrast, seniority in the previous job matters for more than 90% of firms, and the effect is always the same: workers with long tenure in their previous job stay longer in their current job.

Tenure in the current firm often has the opposite consequences for mobility; that is, our results show negative duration dependence for job seniority but with substantial heterogeneity. We investigate this more thoroughly below. For the first two years of seniority, increases are associated with less mobility for 30% of the firms and with more mobility for approximately 15% of the firms. Hence, 30% of firms try to stabilise their less-senior workers whereas 15% of the firms exhibit churning since workers with two years of seniority or less tend to move more often than workers with more than two years of seniority. The probability of leaving the firm increases with seniority in most firms for workers with at least ten years of seniority compared to workers with less seniority in the same firm.

Focusing on variables that reflect potential human resource policies of the firm, we consider first the relation between the separation probability and the individual effect from the starting-wage equation (a measure of worker's quality as evaluated by the market). For 30% to 40% of the firms there is a strong negative relation between the starting-wage person effect and the separation probability, as indicated by the distribution of Student-t statistics in the row labelled 'person effect' in Table 3. A much smaller percentage of firms (less than 10%) have a strong positive relation between the starting-wage person effect and the separation probability. Hence, good workers, as evidenced by their labour market valuations, tend to have longer tenures in a large fraction of firms.

An alternative measure of the heterogeneity of human resource policies comes from examining the evidence for internal labour markets (Doeringer and Piore, 1976). In this view, firms can create labour markets within their own organisation. There are privileged ports of entry and the whole career takes place within the firm through moves between positions. To assess this theory, we have created for each firm a distribution of ages at entry. In Tables 1 and 3, this firm-specific distribution of age at entry is summarised by the variables labelled 'Entry in Qn of the age at entry distribution', where n is the first, second or third quartile, respectively (with $n = 4$ as the reference group). If entry at a young age is associated with a career within the organisation, we should see a negative relation between mobility and workers who enter in Q1 or Q2 of this age-at-entry distribution. For instance if a firm hires workers for some jobs on a short-term basis and other workers for core jobs on a long-term basis at a specific age (mostly young), then one expects to see that entry in the first or the second quartile of the age-at-entry distribution is associated with lower separation probabilities. Direct examination of the relevant rows of Tables 1 and 3 shows that there is not much evidence for this interpretation. There is actually a strong positive relation between entry in the first quartile of the age-at-entry distribution and the separation probability for about 50% of the firms and a strong negative relation (as predicted by internal labour markets) for less than 5%.

By contrast, when firms hire workers at various ages and when the worker-firm pair is concerned about the quality of the match, then one expects to see more separations for workers entering in the bottom of the age distribution. As noted above, this is precisely the case. In more than 30% of firms, workers entering in the first three quartiles of the firm-specific age-at-entry distribution move more often than workers entering in the top quartile of the age distribution (unreported results show that these three coefficients are highly positively correlated). Virtually no firm (less than 5%) preferentially retains workers entering in the bottom of the age distribution. These results are largely inconsistent with most versions of the internal labour market theory.

Finally, approximately 30% of the firms try to keep workers with technical degrees (as opposed to workers with general education, either low or high), as can be seen from the rows labelled 'low general education' and 'high general education' in Tables 1 and 3 since 'technical degrees' is the omitted category. Workers with general education (the coefficient for low and for high are very strongly positively correlated) separate from those firms more often than those with technical degrees.

3.2.2. *Heterogeneity in the wage equation*

We have jointly estimated the firm-specific mobility and wage equations (2). The dependent variable in the firm-specific wage equation is the real log wage rate minus the opportunity wage as measured using the coefficients estimated from the starting-wage equation (1). This opportunity wage is the wage that this worker would receive, in expectation, on the market, at the moment of entry at a random employer. Because this opportunity wage includes $\hat{\theta}_i$, the person effect, unobserved person heterogeneity is (at least partially) controlled. Some variables present in the mobility equation are not included in the wage equation.

Results for the firm-specific wage equation are presented in Table 2 (for the coefficients distribution) and Table 4 (for the Student-t statistic distribution). Many variables are generally statistically different from zero at the 5% level throughout the range of firms. Only seniority in its various guises (as a spline or interacted with education and sex) displays 40% or less of the estimated coefficients that are significant at this level. Even more striking is the almost completely symmetric (around zero) distribution of many wage coefficients.

As with the mobility equation, the best summary of the compensation policy of the firm is captured by the sub-period constants, which are highly positively correlated within firms (not shown in the Tables). A positive sub-period constant corresponds to a high-wage firm and a negative one to a low-wage firm. Roughly 20% of the firms are low-wage firms (at the 5% level) and 30 to 40% of the firms are high-wage firms (again, at the 5% level).

Some additional Table 2 results deserve further comment. That the coefficient for part-time compensation is positive in a substantial fraction of the firms (40%) may appear surprising. However, this variable is also present in the starting wage equation. Hence, what is estimated in the firm-specific parameter is the difference between part-time compensation on the market (included in the opportunity cost of time) and the firm-specific part-time policy. Hence, a positive coefficient on the firm-specific part-time coefficient means that the firm pays its part-time workers better than the market (as compared to a full-time worker with all the same characteristics) and conversely a negative coefficient means that the firm pays its part-time workers worse than the market rate.

It is striking to see that returns to seniority in the first two years of a job – to take a question that has attracted a lot of attention – are significantly negative for 15% to 20% of the firms whereas they are positive for 20% of the firms. These negative returns are still present at higher tenure levels. In fact, the estimated firm-specific seniority coefficients are strongly positively correlated. Hence, around 20% of French firms have negative returns to seniority; 20% have positive returns to seniority and 60% of French firms have returns to seniority that are virtually zero (not significantly different from zero at the 5% level). This result confirms previous findings of AKM or, more recently, of Dostie (2005) using a similar data set but completely different estimation techniques. Notice also that comparing the 5th percentile with the 95th percentile for the male-specific returns to tenure we see that 20% of firms provide higher returns to

Table 4

Distribution of the Estimated Student-t Statistics for the Firm-specific Effects in the Wage Equation

	Mean	Std. Dev.	q1	q5	q25	q50	q75	q95	q99
Part-time	−1.1127	17.0783	−51.0257	−34.3249	−8.0100	−0.9373	4.7356	24.0932	74.0798
First period constant (1976–80)	1.1297	9.4462	−28.1657	−16.3387	−1.6538	1.0256	4.4142	21.0527	23.7489
Second period constant (1982–9)	0.1831	13.8692	−59.8403	−19.6869	−1.6918	0.9407	4.4304	23.3272	25.2641
Third period constant (1991–6)	1.5633	11.7271	−40.2238	−25.1032	−1.6738	1.0696	5.1741	25.1945	27.3398
Tenure (less than two years)	1.1347	4.7704	−6.1039	−4.8529	−0.8270	0.2309	1.6982	13.4474	21.7162
Tenure (2 to 5 years)	0.1226	2.4116	−7.5739	−3.1071	−1.0027	0.0019	1.1426	4.4182	6.1592
Tenure (5 to 10 years)	0.1058	2.6798	−6.1009	−3.8120	−1.0224	0.1469	1.1765	5.0952	7.6454
Tenure (more than 10 years)	1.5590	3.8481	−3.9260	−2.4861	−0.2296	0.6551	1.9420	13.3941	13.5813
Male	0.5180	2.3120	−4.4312	−2.5663	−0.9346	0.3282	1.5950	4.3710	9.3833
Tenure × Male	−0.7141	3.1378	−11.8777	−5.7531	−1.4278	−0.1832	0.8077	2.8793	6.3390
Low general education	−0.0995	2.5634	−9.1418	−3.6319	−0.9258	0.1302	1.1549	3.0289	5.8824
Tenure × Low general education	−0.4878	2.2550	−9.9137	−4.1598	−1.1489	−0.1639	0.6306	2.1636	4.6151
High general education	0.0618	2.1536	−4.4692	−3.6446	−1.2043	0.0510	1.0697	3.1759	8.1428
Tenure × High general education	0.4587	2.1630	−4.9640	−3.4729	−0.5072	0.3880	1.5273	4.3626	6.0300
Correlation between mobility and future wage	−5.2210	22.5500	−95.1290	−78.7770	−1.4760	0.0516	1.1875	5.9166	22.7840
Correlation between mobility and past wage	−0.2850	4.8550	−12.0560	−6.4730	−2.2450	−0.5977	0.7943	8.8196	13.0730
Standard error of the wage shock (in log)	−124.7990	207.9740	−957.1740	−561.5850	−142.9060	−48.2124	−15.6818	73.1256	153.5680

Notes. Between-firm distribution of the estimated Student-t statistics for the wage equation. See notes to Table 2.

tenure to women and 10% to 15% reward male tenure more than female tenure. Results are roughly similar for returns to tenure for our different levels of education. In general, some firms appear to favour low-education workers, other firms appear to favour technical education (the omitted category) and, finally, some firms focus on the high-education group. Those firms that pay low-education workers high wages also pay their high-education workers high wages (see Table 6).

3.2.3. *Heterogeneity of seniority effects in the mobility and wage equations*

Because interpretation of the seniority effects requires (at a minimum) simultaneous assessment of seven coefficients (three sub-period constants and four seniority spline coefficients) in each equation, there is no simple way to disentangle between-firm heterogeneity from within-firm estimation error. To provide a graphical display of the joint pattern of seniority effects in the two equations, we simulated each firm's mobility and wage equation parameters by drawing 1,000 times from their posterior distribution.[12] For each draw from the posterior distribution of the parameters we formed the mobility-seniority profile, estimated using the average constant for the three sub-periods and the four spline coefficients, and the wage-seniority profile (estimated with and without the intercept; when using an intercept, we used the average of the three sub-periods). Averaging the simulations greatly reduces the influence of within-firm sampling error when estimating the quantiles of the distribution of the return to seniority: the simulated distribution of these quantiles is very tight, implying that they have very little sampling variability. Grouping the firms on the basis of this estimated distribution captures between-firm heterogeneity in the returns to seniority.

We grouped the firms into three categories based on the simulated distribution of the return to two years of seniority (excluding the intercept): the lowest quartile, the middle two quartiles, and the highest quartile. Within each of these three groups we averaged the mobility-seniority profile (probit index) and the wage-seniority profile for all simulations of firms in that group. Figure 1 shows the average estimated relation between wage growth and mobility, as a function of seniority, for firms in the lowest quartile of the return-to-seniority distribution. Figure 2 shows the same set of relations for firms in the middle two quartiles of the return-to-seniority distribution. Figure 3 shows the firms in the highest quartile. All three figures are displayed on the same scale.

Dispensing with the middle first, Figure 2 shows that the predominant pattern in French industry is no return to seniority in the wage equation and a very modest increase in the separation rate over the first 10 years of service. Figure 2 also confirms that the dominant policy in this group of firms is a lower base separation rate than in the average firm (negative intercept).[13] Figure 1 shows that firms in the lowest quartile

[12] We sampled from the natural conjugate posterior for the normal regression model for the wage equation and the asymptotic normal posterior for the probit model for the mobility equation. The mode of the posterior distribution in both cases was the maximum likelihood estimate of the firm-specific parameters and the dispersion matrix was the maximum likelihood estimate of the coefficient covariance matrix.

[13] The reference person in this simulation is female, full-time, no experience, zero person effect, technical education, entry in the fourth quartile of the age distribution, zero prior seniority. This affects the location of the intercept, which should be compared to the other two graphs and not interpreted absolutely.

Table 5

Correlation Between Estimated Firm-specific Parameters of the Mobility Equation

	Intercept (first period)	Male	Person Effect	Experience	Tenure (less than two years)	Tenure (more than ten years)	Tenure × Low general ed.	Tenure × High general ed.	Low general education	High general education	Entry in Q1 (age dist.)	Number of previous jobs.	Seniority in the previous job.
Male	0.0151	1.0000	−0.0103	−0.0263	*0.2777*	*0.4171*	*−0.3314*	*−0.2797*	0.0791	*−0.2560*	−0.0203	0.2680	−0.0434
	0.0290		0.0361	0.0471	0.0452	0.0727	0.0745	0.1162	0.0692	0.0946	0.0394	0.2177	0.0310
Person Effect	−0.0118	−0.0103	1.0000	0.1790	*−0.0999*	−0.0202	0.0311	−0.0085	0.0869	−0.0796	*−0.0796*	0.0312	−0.0399
	0.0222	0.0361		0.1492	0.0254	0.0314	0.0383	0.0695	0.0476	0.0579	0.0320	0.0392	0.0220
Experience	−0.0783	−0.0263	0.1790	1.0000	−0.0943	−0.0609	0.1767	0.2524	0.0322	0.0160	−0.0507	0.0463	0.0273
	0.0627	0.0471	0.1492		0.0785	0.0599	0.1463	0.2157	0.0695	0.0701	0.0584	0.0645	0.0357
Tenure (less than	*0.1819*	*0.2777*	*−0.0999*	−0.0943	1.0000	*0.4800*	−0.0483	−0.0764	0.0207	*−0.1279*	*−0.3787*	0.5340	*−0.6249*
two years)	0.0251	0.0452	0.0254	0.0785		0.0355	0.0602	0.0974	0.0478	0.0546	0.0339	0.4125	0.0444
Low general education	*−0.5463*	0.0791	0.0869	0.0322	0.0207	*0.1920*	*−0.7060*	*−0.5792*	1.0000	*0.5269*	*0.5242*	0.2611	*0.3166*
	0.0395	0.0692	0.0476	0.0695	0.0478	0.0590	0.0856	0.1517		0.1357	0.0625	0.2085	0.0408
High general education	*−0.5540*	*−0.2560*	−0.0796	0.0160	*−0.1279*	−0.0989	*−0.3987*	*−0.6859*	*0.5269*	1.0000	*0.6113*	0.2424	*0.4529*
	0.0453	0.0946	0.0579	0.0701	0.0546	0.0687	0.0974	0.1861	0.1357		0.0685	0.1934	0.0489
Entry in Q1	*−0.7446*	−0.0203	*−0.0796*	−0.0507	*−0.3787*	*−0.3247*	*−0.2313*	*−0.1865*	*0.5242*	*0.6113*	1.0000	0.0133	0.6071
(age dist.)	0.0280	0.0394	0.0320	0.0584	0.0339	0.0374	0.0579	0.0850	0.0625	0.0685		0.0347	0.0505
Number of	−0.2234	0.2680	0.0312	0.0463	0.5340	0.2740	−0.0659	−0.0641	0.2611	0.2424	0.0133	1.0000	−0.0657
previous jobs	0.1702	0.2177	0.0392	0.0645	0.4125	0.2189	0.0942	0.1140	0.2085	0.1934	0.0347		0.0557
Seniority in the	*−0.4358*	−0.0434	−0.0399	0.0273	*−0.6249*	*−0.2649*	*−0.3374*	*−0.3468*	*0.3166*	*0.4529*	*0.6071*	−0.0657	1.0000
previous job	0.0391	0.0310	0.0220	0.0357	0.0444	0.0432	0.0538	0.0771	0.0408	0.0489	0.0505	0.0557	

Notes. Correlations are computed using the firm-specific estimates of the parameters of the mobility and wage equations. The mobility and the wage equations are jointly estimated by maximum likelihood. For each firm in the sample, there is a set of estimated parameters used to compute the correlation. Parameters are only estimated for those firms in which there is enough within-firm variability. The estimated correlations are corrected for the estimated error in the firm-specific parameters (see Appendix C). Italic indicates that the correlation is significant at the 5% level or less (standard errors are given below the correlation). Number of observations (firms): 2,507. Source: DADS.

Table 6

Correlation Between Estimated Firm-specific Parameters of the Mobility and Wage Equations

Wage eq.	Intercept (First Period)	Tenure (less than two years)	Tenure (two to five years)	Male	Low general education	High general education	Tenure × High general ed.
Mobility eq.							
Intercept	*−0.2606*	*0.4047*	0.1645	*−0.0809*	*−0.3003*	*−0.1897*	*0.4428*
(First Period)	0.0219	0.1226	0.0897	0.0352	0.0424	0.0359	0.1762
Male	0.0008	−0.0205	−0.0419	−0.0698	0.0212	0.0353	−0.0713
	0.0265	0.0373	0.0754	0.0517	0.0541	0.0589	0.1870
Person Effect	*0.2909*	*−0.0691*	0.1026	0.0540	*0.2106*	*0.1695*	*−0.4002*
	0.0424	0.0322	0.0648	0.0361	0.0527	0.0475	0.1820
Experience	0.0180	−0.0531	−0.0177	−0.0216	0.0012	−0.0364	0.2296
	0.0273	0.0582	0.0634	0.0432	0.0497	0.0596	0.2592
Tenure (less than two years)	*−0.0913*	*0.0968*	0.1091	*−0.0588*	−0.0280	−0.0232	*0.3704*
	0.0151	0.0434	0.0655	0.0265	0.0292	0.0298	0.1545
Low general education	*0.1368*	−0.0321	0.1210	0.0497	0.0842	0.1307	*−0.8366*
	0.0373	0.0537	0.1154	0.0643	0.0942	0.0863	0.3617
High general education	0.0746	0.0061	0.2761	0.1074	0.0699	0.0416	−0.6996
	0.0430	0.0588	0.1611	0.0793	0.0957	0.1066	0.4179
Entry in Q1 (age dist.)	*0.1474*	*−0.2271*	−0.1321	*0.1056*	*0.2926*	*0.1780*	*−0.4369*
	0.0244	0.0826	0.0844	0.0433	0.0514	0.0482	0.2144
Number of previous jobs	0.0677	−0.1384	−0.1958	0.0683	0.0628	−0.0209	0.0966
	0.0582	0.1228	0.2044	0.0629	0.0654	0.0427	0.1442
Seniority in the previous job	*0.0916*	*−0.1390*	*−0.2486*	*0.0833*	0.0438	*−0.1024*	*−0.4428*
	0.0137	0.0517	0.1158	0.0288	0.0246	0.0245	0.1936
Wage eq.							
Tenure (less than two years)	*−0.1855*	1.0000					
	0.0565						
Tenure (two to five years)	−0.1457	−0.3956	1.0000				
	0.0951	0.3074					
Male	*0.1450*	*−0.2282*	−0.2538	1.0000			
	0.0524	0.1085	0.1800				
Low general education	0.0874	*−0.6339*	*−1.1195*	0.1135	1.0000		
	0.0582	0.2229	0.5334	0.0759			
High general education	0.1070	*−0.4235*	*−0.8532*	0.1447	*0.3131*	1.0000	
	0.0509	0.1550	0.4084	0.0801	0.1335		

Notes. Correlations are computed using the firm-specific estimates of the parameters of the mobility and wage equations. The mobility and the wage equations are jointly estimated by maximum likelihood. For each firm in the sample, there is a set of estimated parameters used to compute the correlation. Parameters are only estimated for those firms in which there is enough within-firm variability. The estimated correlations are corrected for the estimated error in the firm-specific parameters (see Appendix C). Italic indicates that the correlation is significant at the 5% level or less (standard errors are given below the correlation). Number of observations (firms): 2,507. Source: DADS.

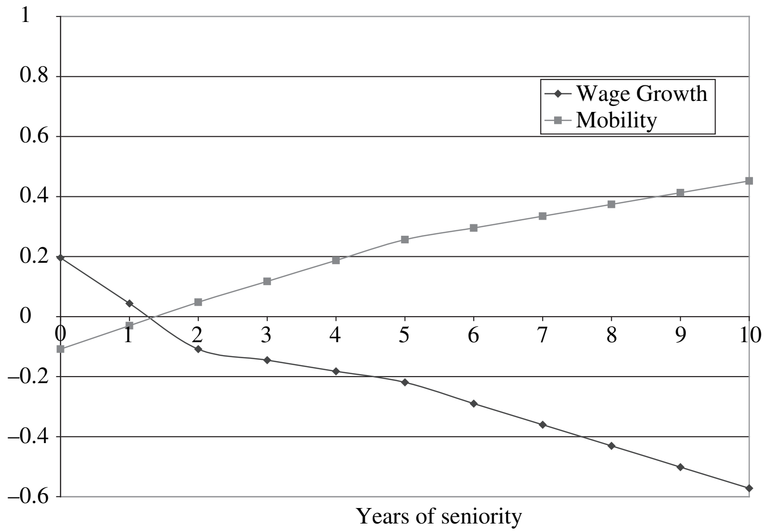

Fig. 1. *Wage Growth and Mobility (Lowest Quartile for Wage Growth)*

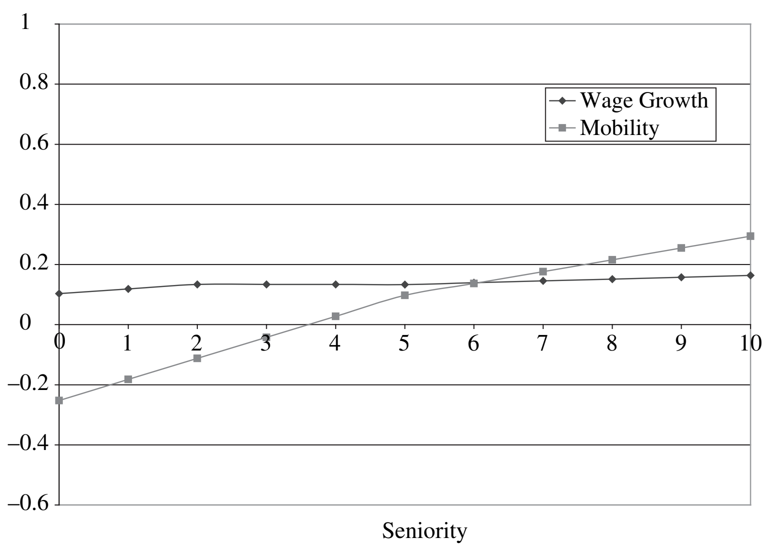

Fig. 2. *Wage Growth and Mobility (Second and Third Quartiles of Wage Growth)*

of the return to seniority distribution have a negative return to seniority over the first 10 years, a higher base separation rate (as compared to the middle group), and about the same rate of increase in the separation rate as the middle group. Figure 3 shows that the highest quartile of the return to seniority distribution has a very substantial positive return to seniority, a higher base rate of separation (compared to the middle group) and a much stronger increase in separation probabilities over the first 10 years of seniority.

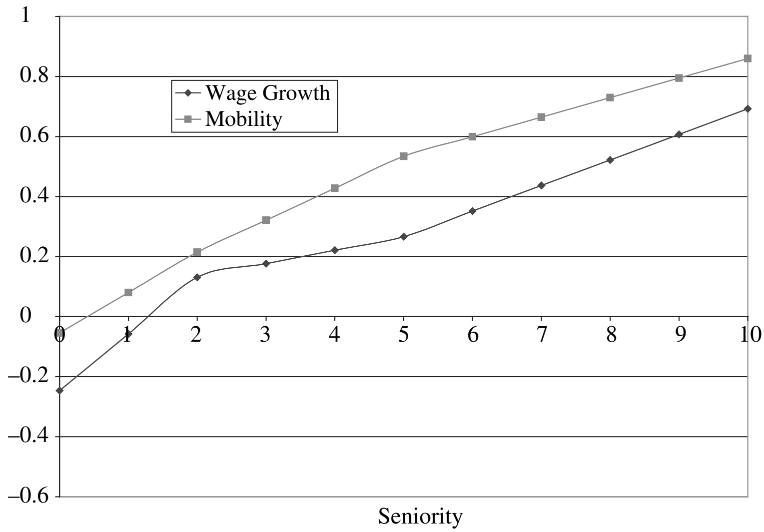

Fig. 3. *Wage Growth and Mobility (Top Quartile of Wage Growth)*

3.2.4. *Some initial evidence for the existence of a mobility policy*

Although we have not developed a full structural model of the relations among the various parameters of a firm's mobility policy, one way to begin examining this question is to consider the between-firm correlation across the various estimated coefficients of the mobility equation. These correlations are presented in Table 5. We correct these correlations to reduce the influence of within-firm estimation error in the firm-specific coefficients.[14] In Table 5, as well as in those that follow, the most significant relations are italicised. For instance, we see that high-turnover firms (based on the first sub-period intercept) are also firms where the separation rate of workers increases as they go from zero to two years of seniority. High-turnover firms (again, based on the first sub-period intercept) tend to keep workers entering at relatively younger ages (based on the coefficient of entry in the first quartile of the age-at-entry distribution). High-turnover firms tend to separate workers with high or low general education (hence, separate workers with a technical education). Similarly, firms in which males are mobile are also firms in which separation probabilities increase with tenure. But, they are also those firms in which workers with general education (as opposed to technical education) are least mobile and tend to have decreasing mobility with tenure. Few firms have a mobility policy that differs for men and women.

Interestingly, the person effect, which measures worker's external quality as estimated from the entry wage equation, is only mildly related to other elements of the

[14] The derivations are shown in Appendix C. Related methods are described in Fuller (1987) and Deaton (1988). Similarly, our performance analysis in a later subsection takes account of the sampling variance induced by the firm-specific estimation, again as described in Appendix C. This correction sometimes entails a loss of observations or even of variables. For instance, when a coefficient is almost never significant in any firm, the sampling variance of the estimated coefficients is always large and, therefore, the true variance of the population parameter (obtained as difference between the estimated total variance of the coefficients and the estimated sampling variation of the coefficients; see Appendix C) becomes negative. For such variables, we do not present corrected correlations.

retention policy of the firm. By contrast, education (general versus technical) appears to be an important element of firms' mobility policies. Firms where workers with a general education separate more are also firms where workers who enter at a young age (in Q1 of the firm-specific age-at-entry distribution) move more often. They are also firms in which workers with high seniority in their previous job separate more often. Finally, in these firms, even though workers with a general education separate more often than those with a technical education, this tendency is decreasing with seniority.

3.2.5. *First evidence for the existence of a pay policy*
Next we examine the correlation between the various estimated coefficients of the wage equation. These results are presented in the bottom panel of Table 6. As AKM found, there is a negative correlation between the firm-specific intercept and returns to seniority in the wage equation. Firms that pay high wages tend to have low returns to seniority, evidence of a 'zero profit' condition. Focusing on the male coefficient, it appears that firms that pay males higher wages than females, all other things equal, are the high-wage firms with low returns to seniority. Finally, firms that pay high wages to workers with a technical education (i.e., not general) also have larger returns to seniority for all workers.

3.2.6. *A second analysis of the mobility and pay relation*
Figures 1 to 3 were our first attempt to interpret the evidence on the relation between mobility and pay. The between-firm correlations between the estimated coefficients of the mobility equation and the estimated coefficients of the wage equation, presented in the top panel of Table 6, provide us with additional evidence. These results allow us to examine more precisely the initial evidence for joint mobility and pay policies. Very few firm-specific coefficients of the wage equation are correlated with firm-specific coefficients in the mobility equation. The central correlation to examine is that between the constant in the wage equation and the constant in the mobility equation. Here, the message is very clear: firms that pay high wages also have a low turnover rate and firms that have large returns to seniority are also those where workers are more mobile. These results reconfirm what we showed in Figure 3. Therefore, returns to seniority should not be viewed as a compensation device for workers with long tenures in firms with a stable workforce but as an incentive device that tries to counteract the potential adverse effects on capital accumulation of high inter-firm worker mobility. Furthermore, high-wage firms are firms where the most able workers (large person effects) tend to separate most frequently. But they are also the firms where mobility decreases with seniority, where workers with technical education tend to stay longer, and where workers who entered at an early age (in Q1 of the age-at-entry distribution) tend to move more often.

In Tables 2 and 4, we present the estimated coefficients and estimated Student-t statistics for the correlation between the mobility error term with the wage error term (future, ρ_1 and past, ρ_2). Here again, the heterogeneity is daunting: 20% of firms have a significant (at the 5% level) and negative ρ_1 whereas slightly less than 20% of firms have a significant and positive ρ_1. Similarly, more than 30% of firms have a significant and negative ρ_2 whereas 15% have a significant and positive ρ_2. A larger (more positive) ρ_1 means that workers who face a positive shock to mobility also face a positive shock to their future wage. This is a potential reflection of firms trying to counteract workers' decisions to accept outside offers. Alternatively, the result may mean that workers who have a tendency to move face very good prospects in their origin firm. Conversely a low ρ_1 means that workers who face a positive shock to their mobility also face a negative shock to their future wage. Once again, the mobility decision is an equilibrium outcome in which workers who will not get promoted may decide to move. The question of who initiates the potential separation is virtually impossible to resolve. Now, ρ_2 captures the correlation between the past shock on wages and the mobility decision. When positive, workers move after an unexpected wage increase. Apparently, this move should be induced by the worker's decision. When negative, workers stay after an unexpected wage increase, potentially resulting from a joint decision. As already mentioned, the latter case (joint optimisation) is much more common: an unexpected wage hike is associated with workers staying an additional year in most firms.

In Tables 7 and 8 respectively, we examine the between-firm correlation of ρ_1 (mobility shock and future wage shock), ρ_2 (mobility shock and past wage shock), and σ (standard error of the wage shock) with the estimated coefficients of the mobility and wage equations, respectively. Firms with a high ρ_1 are low mobility firms, as can be seen from the negative correlation between the intercept (first sub-period) and ρ_1. Furthermore, these low mobility firms also have relatively small σs, as evidenced by the positive correlation between these parameters, which might be due to smaller firm size or more constraints induced by collective agreements and compensation rules in such firms. As could be expected given the positive association between high wages (as measured by the intercept in the within-firm wage equation) and low mobility (as measured by the intercept in the within-firm mobility equation), high-wage firms tend to have large positive ρ_1s, negative ρ_2s and relatively small σs. Recall that a positive ρ_1 means that, in such firms, workers with a positive shock on future wages are more likely to be mobile. Hence, in these high-paying, low-mobility firms, mobility seems to be due to quits, potentially for better jobs. An (unreported) between-firm regression of σ on observed characteristics of the firm, such as employment, the industry, the skill level of the workforce and the labour capital ratio shows that, not surprisingly, large firms have large σ. It also shows that firms with large σ employ high wage workers and have relatively low capital-labour ratios. A similar regression shows that firms with large ρ_1 are in fact firms that are highly capital intensive (large capital labour ratio), but that size of the firm *per se* is not related to ρ_1. By contrast, firms with large ρ_2 are firms with relatively little capital and, once again, size of the firm *per se* is not related to ρ_1. All of these results are consistent with the result that ρ_1 and ρ_2 are strongly negatively correlated and that ρ_1 and σ are also negatively correlated.

3.2.7. *A summary of human resource policies*

To gain a better understanding of potential relations among these various firm-specific effects beyond the simple correlations that we just presented, we performed a principal component analysis of all these estimated coefficients, i.e. those that characterise the mobility policy, those that characterise the pay policy, and those that characterise their relations. Factor estimates, using the covariance matrix corrected for the within-firm parameter estimation error (see Appendix C), are presented in Tables 9 and 10. Table 9 shows the eigenvalues and Table 10 shows the factor loadings for the first four

Table 7

Correlation Between Estimated Firm-specific Parameters of the Mobility Equation and Firm-specific Error Term Parameters

Mobility equation	Intercept (first period)	Male	Person effect	Experience	Tenure (<=2 years)	Tenure (>10 years)	Tenure × Low general ed.	Tenure × High general ed.	Low general education	High gen. education	Entry in Q1 (age dist.)	No. previous jobs	Seniority previous job
Error term parameters													
Correlation between mobility and future wage (ρ_1)	*−0.4304*	*0.0899*	*0.2032*	−0.0497	*−0.0844*	*−0.0816*	*−0.0608*	*0.0959*	*0.1417*	−0.0295	*0.3464*	0.1675	*0.0984*
	0.0305	0.0181	0.0273	0.0419	0.0103	0.0140	0.0148	0.0286	0.0137	0.0208	0.0235	0.1320	0.0148
Correlation between mobility and past wage (ρ_2)	0.0556	*−0.2112*	*−0.3071*	−0.0254	*0.0448*	−0.0218	0.0349	0.0405	−0.0553	*0.2188*	−0.0316	−0.0311	0.0147
	0.0404	0.0685	0.0867	0.0321	0.0189	0.0229	0.0321	0.0536	0.0370	0.0607	0.0272	0.0336	0.0206
Standard error of the wage shock (in logs)	*0.5404*	*0.0654*	*0.0828*	0.1393	0.0023	*0.1484*	*0.0963*	*−0.1165*	*−0.2447*	*−0.2326*	*−0.4798*	−0.2257	*−0.1163*
	0.0238	0.0209	0.0172	0.1138	0.0143	0.0214	0.0248	0.0384	0.0278	0.0307	0.0259	0.1772	0.0231

Notes. Correlations are computed using the firm-specific estimates of the parameters of the mobility and wage equations. The mobility and the wage equations are jointly estimated by maximum likelihood. For each firm in the sample, there is a set of estimated parameters used to compute the correlation. Parameters are only estimated for those firms in which there is enough within-firm variability. The estimated correlations are corrected for the estimated error in the firm-specific parameters (see Appendix C). Italic indicates that the correlation is significant at the 5% level or less (standard errors are given below the correlation). Number of observations (firms): 2,507. Source: DADS.

Table 8

Correlation Between Estimated Firm-specific Parameters of the Wage Equations and Firm-specific Error Term Parameters

Wage equation	Intercept (First Period)	Tenure (less than two years)	Tenure (two to five years)	Male	Low general education	High general education	Tenure × High general ed.
Error term parameters							
Correlation between mobility and future wage (ρ_1)	*0.4933*	*−0.6957*	*−0.4571*	*0.1979*	*0.4108*	*0.3115*	*−0.4068*
	0.0205	0.2050	0.2081	0.0518	0.0341	0.0242	0.1482
Correlation between mobility and past wage (ρ_2)	*−0.3372*	*0.5062*	*0.3144*	*−0.1013*	*−0.1849*	*−0.1574*	0.1906
	0.0809	0.2004	0.1832	0.0488	0.0636	0.0545	0.2105
Standard error of the wage shock (in logs)	*−0.4003*	*0.3888*	*0.3871*	*−0.2317*	*−0.3282*	*−0.3089*	*0.5822*
	0.0203	0.1208	0.1788	0.0624	0.0377	0.0311	0.1868

Notes. Correlations are computed using the estimates of the firm-specific parameters of the mobility and wage equations. The mobility and wage equations are jointly estimated by maximum likelihood. For each firm in the sample, there is a set of estimated parameters used to compute the correlation. Parameters are only estimated for those firms in which there is enough within-firm variability. The estimated correlation is corrected for the estimation error of the firm-specific parameters. Italic indicates that the correlation is significant at the 5% level or less (standard errors are given below the correlation). Number of observations (firms): 2,507. Source: DADS.

Table 9

Factor Analysis of the Firm-specific Parameters of the Mobility and Wage Equations: Eigenvalues

	Eigenvalue	Difference	Proportion	Cumulative
1	8.5768	3.8185	0.2382	0.2382
2	4.7583	0.2521	0.1322	0.3704
3	4.5062	0.3948	0.1252	0.4956
4	4.1114	1.4632	0.1142	0.6098
5	2.6482	0.6009	0.0736	0.6834
6	2.0473	0.2041	0.0569	0.7402
7	1.8432	0.1915	0.0512	0.7914
8	1.6517	0.3102	0.0459	0.8373
9	1.3416	0.2582	0.0373	0.8746
10	1.0833	0.1433	0.0301	0.9047

axes. The results can be summarised as follows. The first four axes capture 61% of the variance. These four dimensions are built on the following linear combinations. The reader is cautioned to remember that the factors being constructed are linear combinations of the firm-specific effects of the indicated variables on the indicated outcome and not linear combinations of the variables themselves.

The first axis contrasts high-wage and low-mobility firms with those that pay low wages and are high-mobility firms. It may be helpful to completely explain this interpretation. The column labelled 'Factor1' in Table 10 shows three large positive loadings for the three sub-period constants in the mobility equation. When these constants are positive, the firm is a high mobility firm (the probability of separation is higher regardless of the characteristics of the individual). The same column also shows three large negative loadings on the sub-period constants in the wage equation. When

Table 10

Factor Analysis of the Firm-specific Parameters of the Mobility and Wage Equations: Factor Loadings

	Factor1	Factor2	Factor3	Factor4
Mobility equation				
Male	0.0546	0.3818	0.1798	0.1933
Full-Time	0.1944	0.1237	0.0593	0.0606
First period constant (1976–80)	0.7950	0.1154	−0.1732	0.0732
Second period constant (1982–9)	0.7877	0.1230	−0.1547	0.1381
Third period constant (1991–6)	0.6613	0.1476	−0.0601	0.0940
Person effect (from starting wage equation)	−0.1142	−0.1159	−0.2308	0.2744
Experience	0.0215	−0.6514	−0.3178	−0.1150
Experience^2	−0.0582	0.7565	0.4127	0.1286
Experience^3	0.0371	−0.7738	−0.4240	−0.1725
Experience^4	−0.0739	0.7764	0.4373	0.2043
Tenure (less than two years)	0.4400	0.3540	0.1326	0.2981
Tenure (2 to 5 years)	0.3606	0.1688	0.0463	0.0662
Tenure (5 to 10 years)	0.5099	0.2869	−0.0822	0.2645
Tenure (more than 10 years)	0.2875	0.5369	0.0964	0.1514
Tenure × Low general education	0.3427	−0.5515	−0.3015	0.1436
Tenure × High general education	0.1976	−0.4691	−0.4085	0.2759
Low general education	−0.5644	0.3350	0.1004	−0.2020
High general education	−0.6033	0.1447	0.0996	−0.4848
Entered in Q1 of age at entry distribution	−0.8451	−0.0473	0.2928	−0.2711
Entered in Q2 of age at entry distribution	−0.7938	−0.0238	0.2961	−0.3535
Entered in Q3 of age at entry distribution	−0.8217	−0.1516	0.2051	−0.4010
Number of previous jobs	−0.0513	0.2263	0.1599	0.2406
Seniority in the previous job	−0.6476	−0.0622	0.0794	−0.3952
Wage equation				
Full-Time	0.1383	−0.0747	0.3109	−0.3358
First period constant (1976–80)	−0.5328	0.1030	−0.4236	0.5723
Second period constant (1982–9)	−0.6220	0.0493	−0.3499	0.5999
Third period constant (1991–6)	−0.5988	0.0019	−0.3360	0.5161
Tenure (less than two years)	0.4467	0.1651	−0.0335	−0.3241
Tenure (2 to 5 years)	0.3198	0.2129	−0.1430	−0.5459
Male	−0.1853	−0.0463	0.0622	0.2192
Low general education	−0.1580	−0.5198	0.8486	0.4522
High general education	−0.0217	−0.4907	0.8542	0.4446
Tenure × High general education	0.8412	−0.6744	1.0749	0.0678
Correlation between mobility and future wage	−0.5646	−0.1012	0.0574	0.6084
Correlation between mobility and past wage	0.2295	0.0496	0.0356	−0.5463
Standard error of the wage shock (in logs)	0.6217	−0.0305	−0.0993	−0.3681

these constants are negative the firm is a low-wage firm regardless of the characteristics of the worker. For this reason we interpret this factor as increasing in the cluster (high-mobility, low-wage) or decreasing in the cluster (low-mobility, high-wage). The high-wage firms also hire relatively older workers whereas the high-mobility firms mostly hire workers at younger ages, as measured by the firm-specific age-at-entry coefficients. This configuration of policies is clearly consistent with the existence of many short-term formal contracts. Among high-mobility firms, because workers are more likely to move as seniority increases, returns to seniority are high in the first five years, potentially as a device designed to keep some carefully selected workers.

The second factor loading axis combines experience and education. The factor groups firms in which mobility is increasing with experience and workers with technical education are better paid in contrast with firms in which stability is decreasing with experience and workers with general education are better paid.

The third axis requires more subtle interpretation. The factor contrasts low-wage firms (negative sub-period constants in the wage equation) that are also high-wage firms for workers with a general education (positive coefficients on the two general education characteristics in the wage equation) with high-wage firms that do not favour one type of education over another. Interestingly, the fourth axis also relates to pay choices for the different education types. It contrasts high-wage firms with relatively low wages for the technically educated and low returns to seniority with low-wage firms with high wages for the technically educated and large returns to seniority. Such firms are neither low nor high-turnover.

Of course, such bundles of characteristics do not exist to help the analyst describe firms. These characteristics should stem from firms' choices that have to be related to their underlying productivity, employment, or capital choices. We try to say more about this in the next subsection. We note in summary for this subsection that a firm-level regression of the four factors on industry indicators demonstrates that these policies are not associated with any specific industry. Hence, these contrasts are mostly a within-industry phenomenon.

3.3. *The Performance Equations*

Our performance equations are presented in Table 11. Each column has the same format: a firm-level performance outcome is related to the four factor-analytic axes that best summarise the estimated parameters, the firm's skill structure, industry indicators, the capital-labour ratio, and employment (for the analysis of value added).[15] The table presents results for log employment, the stability index (defined in the data section), log (capital/labour), log value-added per employee, and operating profit per unit of capital. All these equations are very much in the spirit of those presented in AKM but they improve in many dimensions over them. First, the mobility policy is present whereas it was absent from AKM. Second, the pay policy comprises many more parameters, which are estimated on almost twice as many observations and controlling for the endogeneity of workers' mobility (which was assumed to be exogenous, conditional on the person and firm effects in AKM). Notice however that matching the DADS data with external sources such as the BRN reduces the number of available observations from 2,507 for the correlations to approximately 1,800 for these performance equations.

We first discuss the results for log employment. Although nothing structural should be inferred from our estimates, we still expect to capture descriptively important elements of human resource management policies. Large firms tend to hire workers with low person effects, a result potentially surprising given their frequent use of human resource departments. However, screening on a large scale maybe difficult. Considering the personnel policy as measured by the factor-analytic axes, we see that (very) large firms are often relatively low-wage firms and high-mobility firms.[16] Some high-wage firms that

[15] All estimation results correct for the estimated nature of the explanatory variables as described in Appendix C.

[16] Remember that most of our firms are large since we restrict attention to those with at least 200 individual observations to estimate the firm-specific effects in our model.

Table 11

Relating Firm Characteristics and Performance to the Hiring, Retention and Compensation Policies

	Employment (in logs)		Stability index		Capital-labour ratio (in logs)		Value-added per employee (in logs)		Operating profit per unit of Capital	
	Coefficient	Std Err	Coefficient	Std Err	Coefficient	Std Err	Coefficient	Std Err	Coefficient	Std Err
First factor	0.0256	0.0113	−0.0029	0.0008	−0.0552	0.0120	−0.0044	0.0048	0.0042	0.0010
Second factor	−0.0834	0.0276	−0.0081	0.0015	−0.0290	0.0301	−0.0065	0.0109	0.0014	0.0028
Third factor	−0.0931	0.0340	0.0089	0.0022	0.1158	0.0306	0.0060	0.0128	−0.0005	0.0028
Fourth factor	−0.0369	0.0264	0.0147	0.0019	0.0828	0.0257	0.0117	0.0118	−0.0040	0.0024
Intercept	8.2271	0.2734	0.5066	0.0196	1.9273	0.2654	4.3822	0.1680	0.1856	0.0272
Person effect in the starting wage within the firm (θ, Q1)	3.0688	0.4465	−0.0014	0.0308	−1.2161	0.4506	0.3393	0.2029	−0.1525	0.0467
Person effect in the starting wage within the firm (θ, Q2)	−0.6158	0.7054	0.1860	0.0533	−0.2349	0.7682	0.1421	0.3451	0.2167	0.0710
Person effect in the starting wage within the firm (θ, Q3)	−0.9573	0.3522	−0.0104	0.0335	2.5667	0.4483	0.7975	0.2087	−0.1172	0.0383
Characteristic effect in the starting wage within the firm ($\mathbf{x\beta}$, Q1)	−0.0313	0.3259	0.1137	0.0230	−0.1539	0.3396	0.5052	0.1230	−0.0088	0.0295
Characteristic effect in the starting wage within the firm ($\mathbf{x\beta}$, Q2)	−0.9090	0.5010	0.0496	0.0371	1.0599	0.5320	−0.4358	0.2153	0.0922	0.0470
Characteristic effect in the starting wage within the firm ($\mathbf{x\beta}$, Q3)	0.6054	0.3885	0.0629	0.0279	1.0366	0.4008	−0.1230	0.1723	−0.2211	0.0375
Total Assets (in logs)							0.2637	0.0104		
Employment (in logs)							−0.3049	0.0146		
Number of Observations:	1,812		1,827		1,804		1,746		1,782	

Sources. DADS, BRN. Estimation is by ordinary least squares correcting for the estimated nature of the regressors (the factors). Each regression also includes a full set of industry indicators.

compensate their workers with technical education well also tend to be large (third factor). Similarly, firms that keep their more experienced workers tend to be large firms.

Turning now to the capital–labour ratio, firms with large values of K/L resemble firms that have relatively small employment. High-wage firms that are also high-wage firms for workers with general education have relatively large capital–labour ratios. In addition, skills (as measured by observables or unobservables) are positively associated with a high capital–labour ratio.

Some firms are able to reduce the turnover of their workforce whereas others are not, as reflected in the stability index. Factors related to low separation rates matter in each of our four estimated factor-analytic axes. There are a variety of ways of achieving low separation rates and these are related to the different outcome variables through the estimated factor loadings rather than directly. Hence, combinations of policies are associated with our firm-level outcome variables as Ichniowski *et al.* (1997) found for high-performance workplaces in the steel industry .

A more natural measure of firm performance is value-added per employee. Here, the message is striking: no factor-analytic axis is associated with an effect on productivity. Therefore, compensation and/or mobility strategy do not seem to be correlated with productivity. Put differently, given any such strategy, some adopting firms are more productive while others are not. Should we be surprised? Not necessarily. First, these effects control for firms' skill structure and capital-labour ratio. Our left-hand-side variable is a measure of total factor productivity rather than labour productivity, more likely to be affected by human resources policies.[17] Second, firms are productive for reasons that go beyond human resource management strategies. However, even though compensation and retention strategies do not appear to have an impact on productivity, the quality of the workforce matters. This is confirmed by the positive relation between unobserved but compensated quality as measured at entry, in particular at the top of the distribution of starting-wage person effects. By contrast, observable quality at entry affects productivity at the bottom of the skill distribution. Therefore, hiring policies appear to matter for productivity even though other human resource policies appear to have a smaller effect, if any.

Finally, the estimates for the profit variable shows that the main contrast is between high and low-turnover firms, even though these are not associated with measurable productivity differences. Potentially, low turnover firms, because they are also high-wage firms, make lower profits even though they are as productive as high-turnover firms. High wages may therefore reflect relatively strong union power and significant rents accruing to workers.

4. Conclusion

In this article we have tried to show the benefits of using longitudinally linked worker-firm data to investigate issues that have been central to labour economics and human resource management for years. To do so, we set up a descriptive estimating framework to help us think about the relation between mobility and

[17] The direct correlation between labor productivity and wages is clearly positive in our data.

wages for an individual, both from the worker's own perspective as well as from the employer's perspective. The data sources were based on a very large, longitudinal employer-employee data set for France, the DADS. The system of equations was estimated with many firm-specific effects, very much relying upon the perspective adopted by authors such as Baker *et al.* (1994*a*, *b*) with the distinctive feature that we capture elements of the outside labour market, at entry through an entry wage equation with both person and firm effects as well as at exit by explicitly modelling the joint mobility and wage processes, whereas these authors could not. The results are destructive of the homogeneous view of the labour market in which firms adopt very similar strategies. In fact, the amount of heterogeneity in the policies adopted by the firms is daunting. After documenting this heterogeneity, we tried to characterise what compensation and retention strategies could be in such a world. To do so, we used a simple factor analysis that was able to guide us and show that four factors gave a useful summary view of the heterogeneity. We focus here on the first factor, which appears to give a very simple and clear-cut overview of our results. The main contrast between high-wage, low-mobility firms where returns to seniority are low (even negative) and low-wage, high-mobility firms where returns to seniority are relatively high (in a country where the average returns to seniority are lower than in the US, even compared with Altonji's results) gives a good first-order approximation of the French landscape. Recent job search and matching models (Postel-Vinay and Robin, 2002; Woodcock, 2003) with person and firm heterogeneity appear to be able to generate exactly this type of effect. Other dimensions contrast firms that favour general education with firms that favour more technical education. We show that these firm-level factors appear to be related to inputs, more precisely capital and labour, that the firm uses to produce. Finally, all such choices appear to be unrelated to value-added per worker, showing that there are multiple routes to productivity enhancement.

On the methodology side, this article uses some newer, recently developed, techniques for analysing the matched employer-employee data. It also contains a non trivial number of methodological firsts. To name but a few, the firm-by-firm (maximum likelihood) estimation strategy, the correction for the estimated nature of the parameters characterising the firm policies, the joint modelling of wages and mobility at the firm level, and the identification strategy relying on exclusion restrictions based on variables that can only be constructed using the matched worker-firm aspect of the data (for instance the age at entry within the firm-specific age distribution). In addition, all techniques and models presented here can be used almost identically in other fields of applied research such as health or education, as advocated in the introduction.

We believe that the analysis presented here opens more avenues of research than it closes doors and solves problems. But, we see it as an important next step in understanding the substance as well as the methods to use when analysing firms' hiring, retention, compensation, or more generally human resource management policies. New methods should also be developed that would allow us to perform an analysis of workers' firm to firm movements.

Cornell University, U.S. Census Bureau, CREST, NBER, and IZA
CREST, INSEE, CEPR and IZA
INSEE and CREST

Appendix A: The Likelihood Function for the Firm-specific Model of Wages and Mobility

Consider the starting-wage equation (1) and the firm-specific wage and mobility equations (2) in the text and all definitions associated with those equations. We derive the likelihood for the firm-specific model of wages and mobility in this Appendix. After entry in firm j, and at each value of seniority $s(i,\ t)$, the worker and firm mutually decide to continue or terminate the employment relation. The latent variable $R^{*}_{itj} = \mathbf{Q}^{s(i,t)}_{itj}\boldsymbol{\gamma}_j + v^{s(i,t)}_{ij}$ corresponds to the observation R_{itj} whether the job goes on at date t. A wage rate is observed for $s > 0$ if and only if the employment relation continues. At date t for a worker with seniority s, (after subtracting the effect of the market variables as measured by $\mathbf{X}_{it}\hat{\boldsymbol{\beta}}$), the mobility process can be expressed by (2) in the text. Consider the $s = 2$ example from (3) From this structure of correlation, multivariate normality implies that:

$$\left.\begin{matrix}\eta^{11}_{ij}\\ \eta^{12}_{ij}\\ \eta^{21}_{ij}\\ \eta^{22}_{ij}\\ \eta^{31}_{ij}\end{matrix}\right\} = \begin{cases} v^{1}_{ij} - (\rho_{1j}/\sigma_j)\varepsilon_{it_0(i)+1}\\ \varepsilon_{it_0(i)+1}\\ v^{2}_{ij} - (\rho_{2j}/\sigma_j)\varepsilon_{it_0(i)+1} - (\rho_{1j}/\sigma_j)\varepsilon_{it_0(i)+2}\\ \varepsilon_{it_0(i)+2}\\ v^{3}_{ij} - (\rho_{2j}/\sigma_j)\varepsilon_{it_0(i)+2}\end{cases}$$

$$\rightsquigarrow N\left(0, \begin{bmatrix} 1-(\rho^2_{1j}/\sigma^2_j) & 0 & 0 & 0 & 0\\ 0 & \sigma^2_j & 0 & 0 & 0\\ 0 & 0 & 1-(\rho^2_{2j}/\sigma^2_j)-(\rho^2_{1j}/\sigma^2_j) & 0 & 0\\ 0 & 0 & 0 & \sigma^2_j & 0\\ 0 & 0 & 0 & 0 & 1-(\rho^2_{2j}/\sigma^2_j)\end{bmatrix}\right)$$

where the vector $\boldsymbol{\eta}$ has components with subscripts ij denoting the individual-employer match and superscripts $s(i,\ t)e$ denoting seniority and the equation number ($e = 1$ for the mobility equation; $e = 2$ for the wage equation). This last result is useful for estimation since the likelihood does not involve multiple integration of the normal distribution as shown by

$$R^{*}_{it_0(i)+1j} = \mathbf{Q}^{1}_{it_0(i)+1j}\boldsymbol{\gamma}_j + \frac{\rho_{1j}}{\sigma_j}\varepsilon_{it_0(i)+1} + \eta^{11}_{ij} < 0$$

$$\log w_{it_0(i)+1} - \mathbf{X}_{it_0(i)+1}\hat{\boldsymbol{\beta}} - \hat{\theta}_i - \hat{\psi}_j = \mathbf{Z}^{1}_{it_0(i)+1j}\boldsymbol{\beta}_j + \varepsilon_{it_0(i)+1}$$

$$R^{*}_{it_0(i)+2j} = \mathbf{Q}^{2}_{it_0(i)+2j}\boldsymbol{\gamma}_j + \frac{\rho_{2j}}{\sigma_j}\varepsilon_{it_0(i)+1} + \frac{\rho_{1j}}{\sigma_j}\varepsilon_{it_0(i)+2} + \eta^{21}_{ij} < 0$$

$$\log w_{it_0(i)+2} - \mathbf{X}_{it_0(i)+2}\hat{\boldsymbol{\beta}} - \hat{\theta}_i - \hat{\psi}_j = \mathbf{Z}^{2}_{it_0(i)+2j}\boldsymbol{\beta}_j + \varepsilon_{it_0(i)+2}$$

$$R^{*}_{it_0(i)+3j} = \mathbf{Q}^{3}_{it_0(i)+3j}\boldsymbol{\gamma}_j + \frac{\rho_{2j}}{\sigma_j}\varepsilon_{it_0(i)+2} + \eta^{31}_{ij} > 0.$$

The contribution to the log likelihood of this sequence of observations is

$$\log L(R_{it_0(i)+1j}, w_{it_0(i)+1}, R_{it_0(i)+2j}, w_{it_0(i)+2j}, R_{it_0(i)+3j}) =$$
$$\log\Phi\left[-\mathbf{Q}^1_{it_0(i)+1j}\gamma_j - \frac{\rho_{1j}}{\sigma_j}\left(\frac{\log w_{it_0(i)+1} - \mathbf{X}_{it_0(i)+1}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^1_{it_0(i)+1j}\boldsymbol\beta_j}{\sigma_j}\right)\right]$$
$$+\log\varphi\left(\frac{\log w_{it_0(i)+1} - \mathbf{X}_{it_0(i)+1}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - Z^1_{it_0(i)+1j}\beta_j}{\sigma_j}\right)$$
$$+\log\Phi\left[\begin{array}{c} -\mathbf{Q}^2_{it_0(i)+2j}\gamma_j - \dfrac{\rho_{2j}}{\sigma_j}\left(\dfrac{\log w_{it_0(i)+1} - \mathbf{X}_{it_0(i)+1}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^1_{it_0(i)+1j}\boldsymbol\beta_j}{\sigma_j}\right) \\ -\dfrac{\rho_{1j}}{\sigma_j}\left(\dfrac{\log w_{it_0(i)+2} - \mathbf{X}_{it_0(i)+2}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^2_{it_0(i)+2j}\boldsymbol\beta_j}{\sigma_j}\right)\end{array}\right]$$
$$+\log\varphi(\log w_{it_0(i)+2} - \mathbf{X}_{it_0(i)+2}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^2_{it_0(i)+2j}\boldsymbol\beta_j)$$
$$+\log\left\{1-\Phi\left[\mathbf{Q}^3_{it_0(i)+3j}\gamma_j + \frac{\rho_{2j}}{\sigma_j}\left(\frac{\log w_{it_0(i)+2} - \mathbf{X}_{it_0(i)+2}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^2_{it_0(i)+2j}\boldsymbol\beta_j}{\sigma_j}\right)\right]\right\}$$

where $R_{it_0(i)+sj} = 1$ when $R^*_{it_0(i)+sj} > 0$. More generally, the log likelihood for person i who arrived at date $t_0(i)$ in firm j and stayed exactly S periods (i.e., with one entry wage and $S-1$ observed wages in firm j after this initial date) is:

$$\log L(R_{it_0(i)+1j}, w_{it_0(i)+1}, \ldots, R_{it_0(i)+Sj}) =$$
$$R_{it_0(i)+1j}\log\Phi\left[-\mathbf{Q}^1_{it_0(i)+1j}\gamma_j - \frac{\rho_{1j}}{\sigma_j^2}(\log w_{it_0(i)+1} - \mathbf{X}_{it_0(i)+1}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - Z^1_{it_0(i)+1j}\boldsymbol\beta_j)\right]$$
$$+(1-R_{it_0(i)+1j})\log\Phi\left[-\mathbf{Q}^1_{it_0(i)+1j}\gamma_j - \frac{\rho_{1j}}{\sigma_j^2}(\log w_{it_0(i)+1} - \mathbf{X}_{it_0(i)+1}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^1_{it_0(i)+1j}\boldsymbol\beta_j)\right]$$
$$+\sum_{s=1}^{S-1}\left\{\log\varphi\left(\frac{\log w_{it_0(i)+s} - \mathbf{X}_{it_0(i)+s}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^s_{it_0(i)+sj}\boldsymbol\beta_j}{\sigma_j}\right)\right.$$
$$\left. + R_{it_0(i)+s+1j}\log\Phi\left[\begin{array}{c} -\mathbf{Q}^s_{it_0(i)+s+1j}\gamma_j \\ -\dfrac{\rho_{1j}}{\sigma_j^2}(\log w_{it_0(i)+s+1} - \mathbf{X}_{it_0(i)+s+1}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^{s+1}_{it_0(i)+s+1j}\boldsymbol\beta_j) \\ -\dfrac{\rho_{2j}}{\sigma_j^2}(\log w_{it_0(i)+s} - \mathbf{X}_{it_0(i)+s}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^s_{it_0(i)+sj}\boldsymbol\beta_j)\end{array}\right]\right\}$$
$$+(1-R_{it_0(i)+Sj})\log\bar\Phi\left[\mathbf{Q}^S_{it_0(i)+Sj}\gamma_j + \frac{\rho_{2j}}{\sigma_j^2}(\log w_{it_0(i)+S} - \mathbf{X}_{it_0(i)+S}\hat{\boldsymbol\beta} - \hat\theta_i - \hat\psi_j - \mathbf{Z}^S_{it_0(i)+Sj}\boldsymbol\beta_j)\right].$$

Appendix B: Starting-wage Equation Estimates

Table B1 presents the results for the starting wage equation. We present only the coefficients and not the standard errors, which are not directly delivered by the Abowd *et al.* (2003) estimation technique. Standard errors could be obtained by subtracting the estimated person and firm effects from the wage and rerunning the regression on the observed characteristics contained in this table. In general, the coefficients in this Table are at least 100 times their standard errors.

Appendix C: Correction of Estimation Errors

We start with a simple linear performance equation

Table B1

Entry Wage Equation

	Coefficients	
	Male	Female
Experience	0.1024	0.0667
Experience^2	−0.5612	−0.3347
Experience^3	0.1436	0.0891
Experience^4	−0.0138	−0.0090
Year = 1977	−0.1545	−0.1381
Year = 1978	−0.1293	−0.0976
Year = 1979	−0.1454	−0.1201
Year = 1980	−0.1742	−0.1670
Year = 1982	−0.1987	−0.1693
Year = 1984	−0.1658	−0.1476
Year = 1985	−0.1823	−0.1668
Year = 1986	−0.1915	−0.2040
Year = 1987	−0.1967	−0.1989
Year = 1988	−0.2230	−0.2296
Year = 1989	−0.1955	−0.2135
Year = 1991	−0.1581	−0.1952
Year = 1992	−0.1715	−0.2175
Year = 1993	−0.2309	−0.2543
Year = 1994	−0.3225	−0.3222
Year = 1995	−0.3448	−0.3404
Year = 1996	−0.3691	−0.3859
Region = Ile de France	0.0583	0.0656
Full-time = yes	0.7893	0.7526
First Job	0.0987	0.0780
Second Job	0.1464	0.1268
Third Job	0.1732	0.1619
Fourth Job or More	0.2038	0.2036

Notes. DADS. 4,616,093 observations. The regression also includes a person and a firm effect. Estimated by a conjugate gradient algorithm. See Abowd *et al.* (2003).

$$Y = \mathbf{X}\boldsymbol{\beta} + \varepsilon$$

with $\varepsilon \rightarrow \mathrm{N}(0, \sigma_\varepsilon^2)$ i.i.d. and where $\mathbf{X}$ comes from a first step equation and, therefore, is measured with error following

$$\hat{\mathbf{X}} = \mathbf{X} + \boldsymbol{v}$$

in which $\boldsymbol{v}_i \rightarrow \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_i)$. We know the probability distribution of $\boldsymbol{v}$ for each observation since the first-step estimation delivered a variance-covariance matrix for each firm (set of parameters).

We derive the estimating formula for β together with its associated variance-covariance matrix. Rewriting the above equations, we have

$$Y = \hat{\mathbf{X}}\boldsymbol{\beta} + \varepsilon - \boldsymbol{v}\boldsymbol{\beta}.$$

Assuming that ε is not correlated with $\boldsymbol{v}$ gives

$$V(\varepsilon_i - \boldsymbol{v}_i\boldsymbol{\beta}) = \sigma_\varepsilon^2 + \boldsymbol{\beta}'\boldsymbol{\Sigma}_i\boldsymbol{\beta}$$

Furthermore,

$$\begin{aligned}\hat{\mathbf{X}}_i'\mathbf{Y}_i &= \hat{\mathbf{X}}_i'(\hat{\mathbf{X}}\boldsymbol{\beta} + \varepsilon - \boldsymbol{v}\boldsymbol{\beta})_i \\ &= \hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\boldsymbol{\beta} + \hat{\mathbf{X}}_i'\boldsymbol{\varepsilon}_i - (\mathbf{X}_i' + \boldsymbol{v}_i')\boldsymbol{v}_i\boldsymbol{\beta}.\end{aligned}$$

By taking the average over the observations, the above implies:

$$\mathbf{M}_{XY} = \frac{1}{N}\sum \hat{\mathbf{X}}_i'\mathbf{Y}_i = \frac{1}{N}\sum \hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\boldsymbol{\beta} + \hat{\mathbf{X}}_i'\boldsymbol{\varepsilon}_i - (\mathbf{X}_i' + \boldsymbol{v}_i')\boldsymbol{v}_i\boldsymbol{\beta}$$
$$= \left(\mathbf{M}_{XX} - \frac{1}{N}\sum \boldsymbol{v}_i'\boldsymbol{v}_i\right)\boldsymbol{\beta} + \frac{1}{N}\sum(\mathbf{X}_i' + \boldsymbol{v}_i')\varepsilon_i - \mathbf{X}_i'\boldsymbol{v}_i\boldsymbol{\beta}.$$

Then, by noting that ε_i and v_i are uncorrelated among themselves as well as with $\mathbf{X}_i$, we see that the second and third components of the above equality tend to zero. An empirical counterpart for the first component is needed. Even though we do not know v_i, we know its probability distribution. We estimate the mean of the variance of the residuals by its empirical counterpart:

$$\frac{1}{N}\sum \boldsymbol{v}_i'\boldsymbol{v}_i \to \frac{1}{N}\sum \boldsymbol{\Sigma}_i.$$

Hence, an estimator of $\boldsymbol{\beta}$ is:

$$\hat{\boldsymbol{\beta}}^{(1)} = \left(\mathbf{M}_{XX} - \frac{1}{N}\sum \boldsymbol{\Sigma}_i\right)^{-1}\mathbf{M}_{\mathbf{X}Y}.$$

The difference between $\boldsymbol{\beta}$ and its estimator is given by:

$$\left(\mathbf{M}_{XX} - \frac{1}{N}\sum \boldsymbol{\Sigma}_i\right)(\hat{\boldsymbol{\beta}}^{(1)} - \boldsymbol{\beta}) = \frac{1}{N}\sum(\boldsymbol{\Sigma}_i - \boldsymbol{v}_i'\boldsymbol{v}_i)\boldsymbol{\beta} + (\mathbf{X}_i' + \boldsymbol{v}_i')\boldsymbol{\varepsilon}_i - \mathbf{X}_i'\boldsymbol{v}_i\boldsymbol{\beta}.$$

Now consider the variance of the following random variable of dimension $(k, 1)$

$$\xi_i = (\boldsymbol{\Sigma}_i - \boldsymbol{v}_i'\boldsymbol{v}_i)\boldsymbol{\beta} + \mathbf{X}_i'(-\boldsymbol{v}_i\boldsymbol{\beta} + \varepsilon_i) + \boldsymbol{v}_i'\varepsilon_i.$$

By the central limit theorem, the variance of $(\hat{\boldsymbol{\beta}}^{(1)} - \boldsymbol{\beta})$ can be deduced from the variance-covariance matrix of ξ_i. We have

$$V(\xi_i) = V(\boldsymbol{\Sigma}_i\boldsymbol{\beta} - \boldsymbol{v}_i'\boldsymbol{v}_i\boldsymbol{\beta}) + \sigma_\varepsilon^2\boldsymbol{\Sigma}_i + \mathbf{X}_i'\mathbf{X}_i[\mathrm{E}(\boldsymbol{v}_i\boldsymbol{\beta})^2 + \sigma_\varepsilon^2] - 2\mathbf{X}_i'\mathrm{E}(\boldsymbol{v}_i\boldsymbol{\beta})(\boldsymbol{\Sigma}_i - \boldsymbol{v}_i'\boldsymbol{v}_i)\boldsymbol{\beta}.$$

Focusing on the first element, we see that

$$V[(\boldsymbol{\Sigma}_i\boldsymbol{\beta} - \boldsymbol{v}_i'\boldsymbol{v}_i\boldsymbol{\beta})] = \mathrm{E}[(\boldsymbol{\Sigma}_i\boldsymbol{\beta} - \boldsymbol{v}_i'\boldsymbol{v}_i\boldsymbol{\beta})(\boldsymbol{\beta}'\boldsymbol{\Sigma}_i - \boldsymbol{\beta}'\boldsymbol{v}_i'\boldsymbol{v}_i)]$$
$$= \mathrm{E}(\boldsymbol{v}_i'\boldsymbol{v}_i\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{v}_i'\boldsymbol{v}_i) - \boldsymbol{\Sigma}_i'\boldsymbol{\beta}\boldsymbol{\beta}\boldsymbol{\Sigma}_i.$$

Rewriting $\boldsymbol{v}_i = \boldsymbol{v}_i\boldsymbol{\Sigma}_i^{\frac{1}{2}}$ where $\boldsymbol{v}_i'$ is a normal vector with unit variance, we have

$$\mathrm{E}(\boldsymbol{v}_i'\boldsymbol{v}_i\boldsymbol{\beta}\boldsymbol{\beta}\boldsymbol{v}_i'\boldsymbol{v}_i) = \boldsymbol{\Sigma}_i^{\frac{1}{2}}\mathrm{E}(\boldsymbol{v}_i'\boldsymbol{v}_i\boldsymbol{\Sigma}_i^{\frac{1}{2}}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{\Sigma}_i^{\frac{1}{2}}\boldsymbol{v}_i'\boldsymbol{v}_i)\boldsymbol{\Sigma}_i^{\frac{1}{2}}$$
$$= \boldsymbol{\Sigma}_i^{\frac{1}{2}}\mathrm{E}[(\boldsymbol{v}_i\boldsymbol{\Sigma}_i^{\frac{1}{2}}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{\Sigma}_i^{\frac{1}{2}}\boldsymbol{v}_i')\boldsymbol{v}_i'\boldsymbol{v}_i]\boldsymbol{\Sigma}_i^{\frac{1}{2}}$$
$$= \boldsymbol{\Sigma}_i^{\frac{1}{2}}\mathrm{E}[(\boldsymbol{v}_i\boldsymbol{\Omega}\boldsymbol{v}_i')\boldsymbol{v}_i'\boldsymbol{v}_i]\boldsymbol{\Sigma}_i^{\frac{1}{2}}$$

with

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma}_i^{\frac{1}{2}}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{\Sigma}_i^{\frac{1}{2}}.$$

Hence,

$$\boldsymbol{v}_i\boldsymbol{\Omega}\boldsymbol{v}_i' = \sum_{j=1,l=1}^{K} v_{i,j}v_{i,l}\Omega_{jl}$$

and

$$\mathrm{E}[(\mathbf{v}_i\boldsymbol{\Omega}\mathbf{v}_i')\boldsymbol{v}_i'\boldsymbol{v}_i]_{kk'} = v_{i,k}v_{i,k'}\sum_{j=1,l=1}^{K} v_{i,j}v_{i,l}\Omega_{jl}.$$

The variance-covariance matrix is rewritten as a function of k and k' as follows. If $k \neq k'$, then

$$\mathrm{E}[\mathbf{v}_i'(\mathbf{v}_i\boldsymbol{\Omega}\mathbf{v}_i')\mathbf{v}_i]_{kk'} = \mathrm{E}v_{i,k}^2 v_{i,k'}^2(\Omega_{kk'} + \Omega_{k'k}) = 2\Omega_{kk'}.$$

If $k = k'$, then

$$\mathrm{E}[\mathbf{v}_i'(\mathbf{v}_i\boldsymbol{\Omega}\mathbf{v}_i')\mathbf{v}_i]_{kk} = \mathrm{E}v_{i,k}^4\Omega_{kk} = 3\Omega_{kk}.$$

Denoting $d\boldsymbol{\Omega}$ the diagonal of $\boldsymbol{\Omega}$, one obtains

$$\begin{aligned}\mathrm{E}[\mathbf{v}_i'(\mathbf{v}_i\boldsymbol{\Omega}\mathbf{v}_i')\mathbf{v}_i] &= 2\boldsymbol{\Omega} + d\boldsymbol{\Omega}\\ &= 2\boldsymbol{\Sigma}_i^{\frac{1}{2}}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{\Sigma}_i^{\frac{1}{2}} + d\boldsymbol{\Omega}.\end{aligned}$$

Hence,

$$V[(\boldsymbol{\Sigma}_i\boldsymbol{\beta} - \boldsymbol{v}_i'\boldsymbol{v}_i\boldsymbol{\beta})] = \boldsymbol{\Sigma}_i^{\frac{1}{2}}(\boldsymbol{\Omega} + d\boldsymbol{\Omega})\boldsymbol{\Sigma}_i^{\frac{1}{2}}.$$

Furthermore,

$$\mathrm{E}(\boldsymbol{v}_i\boldsymbol{\beta})(\boldsymbol{\Sigma}_i - \boldsymbol{v}_i'\boldsymbol{v}_i)\boldsymbol{\beta} = -\mathrm{E}(\boldsymbol{v}_i\boldsymbol{\beta})^2\boldsymbol{v}_i' = 0$$

since it is an odd moment. Taken together, this implies that

$$V(\xi_i) = \boldsymbol{\Sigma}_i^{\frac{1}{2}}(\boldsymbol{\Omega} + d\boldsymbol{\Omega})\boldsymbol{\Sigma}_i^{\frac{1}{2}} + \sigma_\varepsilon^2\boldsymbol{\Sigma}_i + \mathbf{X}_i'\mathbf{X}_i\boldsymbol{\beta}'\Sigma_i\boldsymbol{\beta} + \mathbf{X}_i'\mathbf{X}_i\sigma_\varepsilon^2.$$

Now, with this estimation of $V(\xi_i)$ we obtain the variance of $\hat{\boldsymbol{\beta}}^{(1)}$

$$\begin{aligned}V(\hat{\boldsymbol{\beta}}^{(1)}) = \frac{1}{N}(\mathbf{M}_{XX} - \frac{1}{N}\sum\boldsymbol{\Sigma}_i)^{-1} \times \left\{\frac{1}{N}\sum[2\boldsymbol{\Sigma}_i^{\frac{1}{2}}d\boldsymbol{\Omega}\boldsymbol{\Sigma}_i^{\frac{1}{2}} + \sigma_\varepsilon^2\boldsymbol{\Sigma}_i + \mathbf{X}_i\mathbf{X}_i'(\boldsymbol{\beta}'\boldsymbol{\Sigma}_i\boldsymbol{\beta} + \sigma_\varepsilon^2)]\right\}\\ \times\left(\mathbf{M}_{XX} - \frac{1}{N}\sum\boldsymbol{\Sigma}_i\right)^{-1}.\end{aligned}$$

To estimate σ_ε^2 we use a consistent estimate of β to obtain

$$V(\hat{u}_i) - \frac{1}{N}\sum\hat{\boldsymbol{\beta}}'\boldsymbol{\Sigma}_i\hat{\boldsymbol{\beta}} = \sigma_\varepsilon^2$$

where $\hat{u}_i$ denotes the residual of the equation.

This framework is easily adapted to one where some of the **X**s would be measured without error, as in our problem. In this situation, $\boldsymbol{\Sigma}_i$ could fail to be invertible. We rewrite the problem distinguishing between $\mathbf{X}_e$, the variables measured with error and $\mathbf{X}_s$, those variables measured without error.

Then,

$$\hat{\boldsymbol{\beta}}^{(3)} = \left(\frac{1}{N}\sum\begin{matrix}\boldsymbol{\Sigma}_i^{-1}\hat{\mathbf{X}}_{ei}'\hat{\mathbf{X}}_{ei} - I_{K_e} & \boldsymbol{\Sigma}_i^{-1}\hat{\mathbf{X}}_{ei}'\mathbf{X}_{si}\\ \mathbf{X}_{si}'\hat{\mathbf{X}}_{ei} & \mathbf{X}_{si}'\mathbf{X}_{si}\end{matrix}\right)^{-1}\left(\frac{1}{N}\sum\begin{matrix}\boldsymbol{\Sigma}_i^{-1}\hat{\mathbf{X}}_{ei}'\mathbf{Y}_i\\ \mathbf{X}_{si}'\mathbf{Y}_i\end{matrix}\right)$$

and

$$\begin{aligned}M_{\Sigma_i^{-1}X_e,X_s} &= \frac{1}{N}\sum\begin{pmatrix}\boldsymbol{\Sigma}_i^{-1}\hat{\mathbf{X}}_{ei}'\mathbf{Y}_i\\ \mathbf{X}_{si}'\mathbf{Y}_i\end{pmatrix} = \frac{1}{N}\sum\begin{pmatrix}\boldsymbol{\Sigma}_i^{-1}\hat{\mathbf{X}}_{ei}'\\ \mathbf{X}_{si}'\end{pmatrix}(\hat{\mathbf{X}}_{ei}\boldsymbol{\beta}_e + \mathbf{X}_{si}\boldsymbol{\beta}_s - \boldsymbol{v}_{ei}\boldsymbol{\beta}_e + \varepsilon_i)\\ &= \frac{1}{N}\sum\begin{pmatrix}\boldsymbol{\Sigma}_i^{-1}\hat{\mathbf{X}}_{ei}'\hat{\mathbf{X}}_{ei} - \boldsymbol{\Sigma}_i^{-1}\boldsymbol{v}_{ei}'\boldsymbol{v}_{ei} & \boldsymbol{\Sigma}_i^{-1}\hat{\mathbf{X}}_{ei}'\mathbf{X}_{si}\\ \mathbf{X}_{si}'\hat{\mathbf{X}}_{ei} & \mathbf{X}_{si}'\mathbf{X}_{si}\end{pmatrix}\begin{pmatrix}\boldsymbol{\beta}_e\\ \boldsymbol{\beta}_s\end{pmatrix}\\ &\quad + \frac{1}{N}\sum\begin{bmatrix}\boldsymbol{\Sigma}_i^{-1}(\mathbf{X}_{ei}' + \boldsymbol{v}_{ei})\varepsilon_i - \boldsymbol{\Sigma}_i^{-1}\mathbf{X}_{ei}'\boldsymbol{v}_{ei}\boldsymbol{\beta}_e\\ -\mathbf{X}_{si}'\boldsymbol{v}_{ei}\boldsymbol{\beta}_e + \mathbf{X}_{si}'\boldsymbol{\varepsilon}_i\end{bmatrix}.\end{aligned}$$

So that $\hat{\boldsymbol{\beta}}^{(3)}$ is a consistent estimator.

Unfortunately, the above estimators pose practical problems because some estimates of $\hat{\mathbf{X}}$ are too imprecise. Those $\hat{\mathbf{X}}$ that are the least precise make $\hat{\boldsymbol{\beta}}^{(3)}$ almost impossible to interpret. However, one way to address this difficulty is by weighting the estimator presented above by the inverse of the variance of the $\hat{\mathbf{X}}_i$. In practice, it is much easier to use the trace of the variance covariance matrix estimated at the firm-level. Therefore, noting that

$$\begin{aligned}\frac{1}{tr\boldsymbol{\Sigma}_i}\hat{\mathbf{X}}_i'\mathbf{Y}_i &= \frac{1}{tr\boldsymbol{\Sigma}_i}\hat{\mathbf{X}}_i'(\hat{\mathbf{X}}\boldsymbol{\beta}+\varepsilon-\boldsymbol{v}\boldsymbol{\beta})_i\\ &= \frac{1}{tr\boldsymbol{\Sigma}_i}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\boldsymbol{\beta}+\frac{1}{tr\boldsymbol{\Sigma}_i}\hat{\mathbf{X}}_i'\boldsymbol{\varepsilon}_i-\frac{1}{tr\boldsymbol{\Sigma}_i}\left(\mathbf{X}_i'+\boldsymbol{v}_i'\right)\boldsymbol{v}_i\boldsymbol{\beta}\end{aligned}$$

we have

$$\begin{aligned}M_{\frac{1}{tr\Sigma}XY} &= \frac{1}{N}\sum\frac{1}{tr\boldsymbol{\Sigma}_i}\hat{\mathbf{X}}_i'\mathbf{Y}_i = \frac{1}{N}\sum\frac{1}{tr\boldsymbol{\Sigma}_i}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\boldsymbol{\beta}+\frac{1}{tr\boldsymbol{\Sigma}_i}\hat{\mathbf{X}}_i'\boldsymbol{\varepsilon}_i-\frac{1}{tr\boldsymbol{\Sigma}_i}\left(\mathbf{X}_i'+\boldsymbol{v}_i'\right)\boldsymbol{v}_i\boldsymbol{\beta}\\ &= \left(M_{\frac{1}{tr\Sigma}XX}-\frac{1}{N}\sum\frac{1}{tr\boldsymbol{\Sigma}_i}\boldsymbol{v}_i'\boldsymbol{v}_i\right)\boldsymbol{\beta}+\frac{1}{N}\sum\frac{1}{tr\boldsymbol{\Sigma}_i}(\mathbf{X}_i'+\boldsymbol{v}_i')\boldsymbol{\varepsilon}_i-\frac{1}{tr\boldsymbol{\Sigma}_i}\mathbf{X}_i'\boldsymbol{v}_i\boldsymbol{\beta}.\end{aligned}$$

Since $\boldsymbol{\varepsilon}_i$ and $\boldsymbol{v}_i$ are independent and uncorrelated with $\mathbf{X}_i$, the second term of the previous equation tends to 0. To obtain an estimate of the first term, we do the following. The mean of the variance of the residuals can be estimated using the data as

$$\frac{1}{N}\sum\frac{1}{tr\boldsymbol{\Sigma}_i}\boldsymbol{v}_i'\boldsymbol{v}_i \to \frac{1}{N}\sum\frac{\boldsymbol{\Sigma}_i}{tr\boldsymbol{\Sigma}_i}.$$

An estimator of β can be written as

$$\hat{\boldsymbol{\beta}}^{(4)} = \left(M_{\frac{1}{tr\Sigma}XX}-\frac{1}{N}\sum\frac{\boldsymbol{\Sigma}_i}{tr\boldsymbol{\Sigma}_i}\right)^{-1}M_{\frac{1}{tr\Sigma}XY}.$$

The difference between β and its estimator is given by

$$\left(M_{\frac{1}{tr\Sigma}XX}-\frac{1}{N}\sum\frac{\boldsymbol{\Sigma}_i}{tr\boldsymbol{\Sigma}_i}\right)(\hat{\boldsymbol{\beta}}^{(4)}-\boldsymbol{\beta}) = \left(\frac{1}{N}\sum\frac{(\boldsymbol{\Sigma}_i-\boldsymbol{v}_i'\boldsymbol{v}_i)}{tr\boldsymbol{\Sigma}_i}\boldsymbol{\beta}+\frac{1}{tr\boldsymbol{\Sigma}_i}(\mathbf{X}_i'+\boldsymbol{v}_i')\boldsymbol{\varepsilon}_i-\frac{1}{tr\boldsymbol{\Sigma}_i}\mathbf{X}_i'\boldsymbol{v}_i\boldsymbol{\beta}\right)$$

and computations similar to those presented above allow us to compute the variance matrix of the estimator.

References

Abowd J. M., Creecy R. and Kramarz F. (2003). 'Computing person and firm effects using linked longitudinal employer-employee data', Cornell University Working Paper, revised.

Abowd J. M. and Kramarz F. (1999). 'The analysis of labor markets using matched employer-employee data', in (O. Ashenfelter and D. Card, eds.), *Handbook of Labor Economics*, ch. 26, vol. 3B, pp. 2629–710, Amsterdam: North-Holland.

Abowd J. M., Kramarz F. and Margolis D. N. (1999). 'High wage workers and high wage firms', *Econometrica*, vol. 67, pp. 251–334.

Abowd J. M. and Vilhuber L. (2005). 'The sensitivity of economic statistics to coding errors in personal identifiers', *Journal of Business and Economics Statistics*, vol. 23(2) (April), pp. 133–52, JBES Joint Statistical Meetings invited paper with discussion and 'Rejoinder' (April) pp. 162–5.

Abraham K.J. and Farber H.S. (1987). 'Job duration, seniority and earnings', *American Economic Review*, vol. 77, pp. 278–97.

Altonji J. G. and Shakotko R. A. (1987). 'Do wages rise with job seniority?', *Review of Economic Studies*, vol. 54, pp. 437–59.

Altonji J. G. and Williams N. (1992). 'The effects of labor market experience, job seniority and job mobility on wage growth', NBER Working Paper Series, no. 4133.

Altonji J. G. and Williams N. (1997). 'Do wages rise with job seniority? A reassessment', NBER Working Paper Series, no. 6010.

Baker G., Gibbs M. and Holmstrom B. (1994*a*). 'The internal economics of the firm: evidence from personnel data', *Quarterly Journal of Economics*, vol. 109 (November), pp. 881–919.

Baker G., Gibbs M. and Holmstrom B. (1994*b*). 'The wage policy of a firm', *Quarterly Journal of Economics*, vol. 109 (November), pp. 921–55.

Blumen I., Kogan M. and McCarthy P. J. (1955). *The Industrial Mobility of Labor as a Probability Process*, Cornell Studies of Industrial and Labor Relations, vol. 6, Ithaca, NY: Cornell University Press.

Brown J. M. (1989). 'Why do wages increase with tenure? On-the-job training and life-cycle wage growth observed within firms', *American Economic Review*, vol. 79, pp. 971–91.

Buchinsky M., Fougère D., Kramarz F. and Tchernis R. (2002). 'Interfirm mobility, wages, and the returns to seniority and experience in the U.S', CREST Working Paper no. 2002–29.

Deaton A. (1988). 'Quality, quantity, and spatial variation of price', *American Economic Review*, vol. 78, pp. 418–30.

Doeringer P. B. and Piore M. J. (1976). *Internal Labor Markets and Manpower Analysis*, Lexington: CD Heath.

Dostie B. (2005). 'Job turnover and the returns to seniority', *Journal of Business and Economic Statistics*, vol. 23(2) (April) pp. 192–9.

Dustman C. and Meghir C. (2003). 'Wages, experience and seniority', Working Paper, no. 01/01, The Institute for Fiscal Studies.

Flinn C. J. (1986). 'Wages and job mobility of young workers', *Journal of Political Economy*, vol. 94(3), part 1 (June), pp. S88–110.

Fuller W. A. (1987). *Measurement Error Models*, NewYork: Wiley.

Heckman J. J. (1981). 'Heterogeneity and state dependence', in (S. Rosen, ed.), *Studies in Labor Market*, pp. 91–139, Chicago: University of Chicago Press.

Ichniowski C., Shaw K. and Prennushi G. (1997). 'The effects of human resource management practices on productivity: a study of steel finishing lines', *American Economic Review*, vol. 87, pp. 291–313.

Jovanovic B. (1979). 'Job matching and the theory of turnover', *Journal of Political Economy*, vol. 87, pp. 972–90.

Lazear E. P. (1995). *Personnel Economics*, Cambridge, MA: MIT Press.

Lillard L. A. and Willis R. J. (1978). 'Dynamic aspects of earnings mobility', *Econometrica*, vol. 46, pp. 985–1012.

Margolis D. N. (1996). 'Cohort effects and the return to seniority in France', *Annales d'Economie et Statistique*, vol. 41/42, pp. 443–64.

Mincer J. and Jovanovic B. (1981). 'Labor mobility and wages. Heterogeneity and state dependence', in (S. Rosen, ed.), *Studies in Labor Market*, Chicago: University of Chicago Press.

Postel-Vinay F. and Robin J-M. (2002). 'Equilibrium wage dispersion with worker and employer heterogeneity', *Econometrica*, vol. 70, pp. 2295–350.

Topel R.H. (1991). 'Specific capital, mobility, and wages: wages rise with job seniority', *Journal of Political Economy*, vol. 99, pp. 145–75.

Topel R. H. and Ward M. P. (1992). 'Job mobility and the careers of young men', *Quarterly Journal of Economics*, vol. 107(2), pp. 439–79.

Woodcock S. D. (2003). 'Agent heterogeneity and learning: an application to labor markets', Chapter 1 of Ph.D thesis, Cornell University.