

Beffy, Magali; Buchinsky, Moshe; Fougère, Denis; Kamionka, Thierry; Kramarz, Francis

Working Paper

The returns to seniority in France (and why are they lower than in the United States?)

IZA Discussion Papers, No. 1935

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Beffy, Magali; Buchinsky, Moshe; Fougère, Denis; Kamionka, Thierry; Kramarz, Francis (2006) : The returns to seniority in France (and why are they lower than in the United States?), IZA Discussion Papers, No. 1935, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/33525>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 1935

**The Returns to Seniority in France
(and Why Are They Lower than in the United States?)**

Magali Beffy
Moshe Buchinsky
Denis Fougère
Thierry Kamionka
Francis Kramarz

January 2006

The Returns to Seniority in France (and Why Are They Lower than in the United States?)

Magali Beffy

CREST-INSEE and IZA Bonn

Moshe Buchinsky

UCLA, CREST and NBER

Denis Fougère

CNRS, CREST and IZA Bonn

Thierry Kamionka

CNRS and CREST

Francis Kramarz

CREST-INSEE, CEPR and IZA Bonn

Discussion Paper No. 1935
January 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Returns to Seniority in France (and Why Are They Lower than in the United States?)*

We estimate a model of the joint participation and mobility along with the individuals' wage formation in France. Our model makes it possible to distinguish between unobserved person heterogeneity and state-dependence. We estimate the model using state of the art bayesian methods employing a long panel (1976-1995) for France. Our results clearly show that returns to seniority are small, and for some education groups are close to zero. The specification here is the same as that used in Buchinsky, Fougère, Kramarz and Tchernis (2002), where the returns to seniority were found to be quite large. This result also holds when using the method employed by Altonji and Williams (1992) for both countries. It turns out that differences between the two countries relate to firm-to-firm mobility. Using a model of Burdett and Coles (2003), we explain the rationale for this phenomenon. Specifically, in a low-mobility country such as France, there is little gain in compensating workers for long tenures because they tend to stay in the firm for most, if not all, of their career. This is true even in cases where individuals clearly possess substantial amount of firm-specific human capital. In contrast, for a high-mobility country such as the United States, high returns to seniority have a clear incentive effect, and firms are induced to pay the premium associated with firm-specific human capital to avoid losing their most productive workers.

JEL Classification: J24, J31, J63

Keywords: participation, wage, job mobility, returns to seniority, returns to experience, individual effects

Corresponding author:

Francis Kramarz
CREST-INSEE
Département de la Recherche
15, bd. Gabriel Péri
92245 Malakoff
France
Email: kramarz@ensae.fr

* We thank Pierre Cahuc, Jean Marc Robin, Alexandru Voicu, and seminar and conference participants at Crest, Laval University (Quebec), Paris-I University, Uppsala University, Warwick University, Royal Holloway (DWP conference), SOLE meetings San Francisco, 2005), JMA (Hammamet, 2005), World Meetings of the Econometric Society (London, 2005) for very helpful comments. In addition, Alexandru Voicu spotted a data error in an early version. Remaining errors are ours.

1 Introduction

In the past two decades, enormous progress has been made in the analysis of the wage structure. However, there is still significant disagreement about wage growth, a key issue in labor economics. In particular, the respective roles of general human capital, as measured by experience, and firm-specific human capital, as measured by tenure, are still generally debated. In general, experience and tenure increase simultaneously except when a worker moves from one firm to another, or becomes unemployed. Hence, studying the nature of participation and firm-to-firm mobility (or in short mobility), which, in turn, determine experience and seniority (or tenure), respectively, should be central to the study of wages. This will allow us to better identify these two components of human capital accumulation, which, in turn, will better serve us in assessing the respective roles of general (transferable) human capital and specific (non-transferable) human capital.

The role and relative importance of job tenure and experience on wage growth has been studied extensively. The results are generally mixed, especially for the U.S. Some authors concluded that experience matters more than seniority in wage growth (e.g. Altonji and Shakotko (1987), Altonji and Williams (1992 and 1997)), while others concluded that both experience and tenure are important factor of wage growth (e.g. Topel (1991), Buchinsky, Fougère, Kramarz and Tchernis (2002) BFKT, hereafter). Indeed, identifying the relative roles of tenure and experience is a somewhat complex issue to study. In fact, it seems that various studies uncovered a number of crucial difficulties and provided varying solutions that potentially affect the ultimate estimates.¹

Empirically, it is generally agreed that there exists a positive correlation between seniority and wages. Several economic theories have offered some explanations for the interdependence between wage growth and job tenure. First, the role of specific job tenure on the dynamic of wages has been studied by various human capital theories, starting with the seminal work of Becker (1964) and Mincer (1974). The central point of this theory is the increase in earnings that stems from individual's investment in human capital. The structure of wages can also be described by job matching theory (Jovanovic (1979), Miller (1984), and Jovanovic (1984)). This theory attempts to provide explanations for both mobility of workers across firms and the observed patterns of significant decreases in the job separation rate as job-tenure increases. The key assumption for these types of models is that there exists a specific productivity level for any worker-occupation match. While the worker's wage depends on this productivity, it is, a priori, unknown. Indeed, the specific human capital investment will be larger when the match is less likely to be terminated (see Jovanovic (1979)). Finally, a job matching model typically predicts an increase in the worker's wage with job seniority.

Alternatively, the dynamic of wage changes can be explained by deferred compensation

¹Few such issues are: The definition of the variables, issues about the errors in measured seniority, the estimation methods that are used, the methods controlling for unobserved heterogeneity components in the model, and the exogeneity assumptions that are made.

theories. The key element of these theories is the existence of a contract between the firm and an employee, which is chosen so that the worker's choice of effort and/or quit decision will be optimal (see Salop and Salop (1976), or Lazear (1979, 1981, 1999)). These theories predict that workers starting in a firm will be paid below their marginal product, whereas workers with long tenure in the firm will be paid above their marginal product.

More recently, equilibrium wage-tenure contracts have been shown to exist within a matching model (see Burdett and Coles (2003) or Postel-Vinay and Robin (2002) in a slightly different context). At the equilibrium, firms post contracts that make wages increase with tenure. Some of these models are able to characterize both workers' mobility and the existing positive relation between wage and tenure. For instance, the Burdett-Coles model, allows the specificities of the wage-tenure contract to depend heavily on workers' preferences, as well as on labor market characteristics such job offer arrival rate.

While the relation between wage growth and mobility (or job tenure) may result from (optimal) choices of the firm and/or the worker, it may also simply stem from spurious duration dependence. Indeed, if there is a correlation between job seniority and a latent variable measuring worker's productivity, and if, in addition, more productive workers have higher wages, then there will be positive correlation between wages and job seniority, even though wages do not directly depend on job tenure (see, for instance Abraham and Farber (1987), Lillard and Willis (1978), and Flinn (1986)). This latter point illustrates the vital importance of being able to control for unobserved heterogeneity components. Furthermore, it highlights the need to control for the endogeneity of the mobility decisions, and consequently of measured job tenure.

BFKT develop a model in which costs that are induced by mobility generate state-dependence in the mobility decision, and similarly for the participation decision (as has already been demonstrated by Hyslop (1999)). It is well-known in the literature that due to the problems raised above the usual Ordinary Least-Squares (OLS) estimates of the returns to seniority will be biased. There are many ways to address this problem. One solution is the use of the instrumental variables framework as is done by Altonji and Shakotko (1987). Alternatively, one can use panel data models that control for fixed effects (e.g. Abowd, Kramarz and Margolis (1999)). Finally, one can take a more direct approach and jointly model the wages outcome, along with the mobility and participation decisions (e.g. BFKT).

In this paper, we adopt the latter approach. More specifically, we control for both state-dependence and (correlated) unobserved individual heterogeneity in the mobility and participation decisions. We also control for correlated unobserved individual heterogeneity in the wage equation. We use a Bayesian framework, similar to that used in BFKT, using Markov Chain Monte Carlo (MCMC) procedure that involve some Gibbs sampling steps combined with the Metropolis-Hastings algorithm. As our model contains limited dependent endogenous variables (i.e., participation and mobility), we also need to use some additional data augmentation steps.

We use data from the match of the French Déclaration Annuelle de Données Sociales (DADS) panel, providing us with observations on wages for the years 1976 through 1995, with the

Echantillon Démographique Permanent (EDP) that provides time-variant and time-invariant personal characteristics. Because we use the exact same specification as in BFKT and relatively similar data sources, we place ourselves in a good position for comparing the returns to seniority in France and in the U.S. The estimates of the returns to seniority appear to be in line with those obtained by Topel (1991), and to a lesser extent with those of Altonji and Shakotko (1987) and Altonji and Williams (1992, 1997), for the U.S. In complete contrast, estimates obtained for France are much smaller than those obtained for the U.S. in any of the studies reported in the literature. In fact, some of the returns to seniority in France are virtually equal to zero. In comparison, the returns to experience are rather large and close to those estimated by BFKT.

We proceed then with an attempt to understand the rationale for the enormous differences between the U.S. and France. For this purpose, we make use of the equilibrium search model with wage-contracts proposed by Burdett and Coles (2003). In this model, contracts differ in the equilibrium rates of returns to tenure, i.e., the slope of the tenure profile. Elements that determine these slopes include: Job arrival rate (and hence workers' propensity to move) and risk aversion. We show that, for all values of the relative risk aversion coefficient, the larger the job arrival rate, the steeper the wage-tenure profile. Indeed, recent estimates show that the job arrival rate for the unemployed is about 1.71 per year in the U.S., while it is only 0.56 per year in France (Jolivet, Postel-Vinay and Robin (2004)). Therefore, the returns to seniority directly reflect the patterns of mobility in the two countries.

The remainder of the paper is organized as follows. Section 2 presents the statistical model. Section 3 explains the crucial parts of the estimation method employed here. Section 4 follows with description of the data sources. Section 5 provides the empirical results obtained for France, while Section 6 contrasts these results with those obtained by BFKT for the U.S. We also provide in this section a theoretical explanation of these differences supported by additional simulations. Finally, Section 7 briefly concludes.

2 The Statistical Model

2.1 Specification of the General Model

The main goal of our study is to examine the returns to experience and seniority. Hence, we need to focus our attention on the obvious endogeneity of the participation and mobility decisions, which, in turn, define experience and tenure on the job. We follow here closely BFKT, extending on Hyslop (1999) (who focuses only on participation). The economic model that supports our approach is a structural dynamic choice model of firm-to-firm worker's mobility, with the mobility costs incurred by the worker. BFKT shows that under a set of plausible assumptions on this cost structure, this model generates first-order state dependence for the participation and mobility processes.² Therefore, the statistical model that we estimate here follows directly from this structural choice model of participation and mobility, whereby the

²For the specific conditions and detailed description see BFKT.

wage equation is estimated jointly with the participation and mobility equations. The equations for participation, mobility, and log are given, respectively, by

$$\begin{aligned} y_{it} &= \mathbf{1}(y_{it}^* > 0), \quad \text{where} \\ y_{it}^* &= \gamma^M m_{it-1} + \gamma^Y y_{it-1} + X_{it}^Y \delta^Y + \theta_i^{Y,I} + v_{it}, \end{aligned} \tag{1}$$

$$\begin{aligned} m_{it} &= y_{it} \mathbf{1}(m_{it}^* > 0), \quad \text{where} \\ m_{it}^* &= \gamma m_{it-1} + X_{it}^M \delta^M + \theta_i^{M,I} + u_{it}, \quad \text{and} \end{aligned} \tag{2}$$

$$w_{it} = y_{it} \left(X_{it}^W \delta^W + J_{it}^W + \theta_i^{W,I} + \epsilon_{it} \right),$$

for all years for which $t > 1$, for $i = 1, \dots, n$. The quantity y_{it}^* is a latent variable measuring the value of participation at time t , while y_{it} is the usual indicator function. Similarly, m_{it}^* is a latent variable measuring worker's i value of moving between t and $t + 1$, and m_{it} is an indicator function, denoting whether or not the individual moved at the end of time t . Note that by definition, observed mobility m_{it} is equal to 0 when the individual does not participate at time t . Finally, note that m_{it} is not observed (censored) whenever a worker participates at date t but does not participate at $t + 1$.

The variable w_{it} denotes the logarithm of the annualized total real labor costs. The variable X_{it} denotes observable time-variant, as well as the time-invariant, characteristics for the individual. The function J_{it}^W summarizes the worker's past career choices at date t (the exact specification is detailed below).

The terms $\theta^{Y,I}$, $\theta^{M,I}$, and $\theta^{W,I}$ denote the random effects specific to the individuals, while u , v and ϵ are idiosyncratic error terms. In principle, there are J firms and N individuals in the panel of length T , but our panel is unbalanced in the sense that we do not observe all individuals in all time periods.

Note that because lagged mobility and lagged participation must be included in the participation and mobility equations, one needs to control for the well-known initial conditions in the first period. We follow Heckman (1981) and add for $t = 1$ the following participation, mobility, and wages equations, respectively

$$y_{i1} = \mathbf{1} \left(X_{i1}^Y \delta_0^Y + \alpha_i^{Y,I} + v_{i1} > 0 \right) \tag{3}$$

$$m_{i1} = y_{i1} \mathbf{1} \left(X_{i1}^M \delta_0^M + \alpha_i^{M,I} + u_{i1} > 0 \right), \quad \text{and} \tag{4}$$

$$w_{i1} = y_{i1} \left(X_{i1}^W \delta^W + \theta_i^{W,I} + \epsilon_{i1} \right). \tag{5}$$

2.2 Stochastic Assumptions

The vector of individual specific effects, including those from the initial condition equations is given by

$$\theta_i^I = \left(\alpha_i^{Y,I}, \alpha_i^{M,I}, \theta_i^{Y,I}, \theta_i^{W,I}, \theta_i^{M,I} \right).$$

We assume that individuals are independent, but their various individual effects may not have the same distribution. Specifically, we assume that

$$\theta_i^I | \Sigma_i^I \sim \text{i.i.d. } \mathcal{N}(0, \Sigma_i^I),$$

where the variance-covariance matrix Σ_i^I takes the following form:

$$\begin{aligned} \Sigma_i^I &= D_i \Delta_\rho D_i, \\ \Delta_\rho &= CC', \quad \text{and} \\ C &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \cos_1 & \sin_1 & 0 & 0 & 0 \\ \cos_2 & \sin_2 \cos_3 & \sin_2 \sin_3 & 0 & 0 \\ \cos_4 & \sin_4 \cos_5 & \sin_4 \sin_5 \cos_6 & \sin_4 \sin_5 \sin_6 & 0 \\ \cos_7 & \sin_7 \cos_8 & \sin_7 \sin_8 \cos_9 & \sin_7 \sin_8 \sin_9 \cos_{10} & \sin_7 \sin_8 \sin_9 \sin_{10} \end{pmatrix}, \\ \cos_i &= \cos(\eta_i), \quad \text{for } \eta_i \in [0; \pi], \\ D_i &= \text{diag}(\sigma_{i1}, \sigma_{i2}, \sigma_{i3}, \sigma_{i4}, \sigma_{i5}) \quad \text{and} \\ \sigma_{ij} &= \exp(x_i^{F'} \gamma_j), \quad \text{for } i = 1, \dots, N; \quad j = 1, \dots, 5. \end{aligned} \tag{6}$$

Note that Σ_i^I is, by construction, positive definite for all values of $\eta' = (\eta_1, \dots, \eta_{10})$. In (6) we use a Cholesky decomposition for the correlation matrix, the matrix C can be expressed using a trigonometric form as shown above. For the diagonal variance matrix, we use a factor decomposition: x_i^F denotes the factors specific to individual i .³

Finally, for the idiosyncratic error terms we assume

$$\begin{pmatrix} v_{it} \\ u_{it} \\ \epsilon_{it} \end{pmatrix} \sim \text{i.i.d.} \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{ym} & \rho_{yw}\sigma \\ \rho_{ym} & 1 & \rho_{wm}\sigma \\ \rho_{yw}\sigma & \sigma\rho_{wm} & \sigma^2 \end{pmatrix} \right).$$

It is worthwhile noting that the specification of the joint distribution of the person specific effects has direct implications for the correlation between the regressors and the corresponding random effects. To see this, consider an individual with seniority level $s_{it} = s$. Note that s_{it} can be written as:

³The γ_j terms are estimated separately in a factor analysis of individual data. The variables that enter this analysis are the sex, the year of birth, the region where the individual lives (Ile de France versus other regions), the number of children, the marital status, the part-time status, and the unemployment rate in the department of work.

$$s_{it} = (s_{it-1} + 1)1(m_{it} = 0, y_{it} = 1).$$

This equation can be expanded by recursion to the individual’s entry year into the firm. Since the seniority level of those currently employed depends on the sequence of past participation and mobility indicators, it must also be correlated with the person-specific effect of the wage equation $\theta_i^{W,I}$. This, in turn, is correlated with $\theta_i^{M,I}$ and $\theta_i^{Y,I}$, the person-specific effects in the mobility and participation equations, respectively. Similarly, experience and the J^W function are also correlated with $\theta_i^{W,I}$. Given that lagged values of participation and mobility, as well as the seniority level appear in both the participation and mobility equations, it follows that the regressors in these two equations are also correlated, albeit in a complex fashion, with the corresponding person-specific effects, namely $\theta_i^{M,I}$ and $\theta_i^{Y,I}$, respectively.

This reasoning also applies to the idiosyncratic error terms. Therefore, the individual specific effect and the idiosyncratic error term in the wage equation are both correlated with the experience and seniority variables through the correlation of the individual effects and idiosyncratic error terms across our system of equations. Putting it differently the system of equation specified here allows for correlated random effects.

3 Estimation

As in BFKT, we adopt here a Bayesian method. The estimates provided are given by the mean of the posterior distribution of the various parameters. We construct the posterior distribution via the use of a Markov Chain Monte Carlo (MCMC) procedure, where in each iteration we draw from the posterior distribution of these parameters, conditional on the data. We do so, because the computation of the analytical form for posterior distribution is intractable. Specifically, we use a combination of the Gibbs Sampling algorithm, augmented by Metropolis-Hastings whenever needed, (for example when drawing the correlation coefficients), to obtain draws from this posterior distribution. Below we briefly explain the implementation of the MCMC for our problem. For a more detailed description see Appendix A.

3.1 Principles of the Gibbs Sampler

Given a parameter set and the data, the Gibbs sampler relies on the recursive and repeated computations of the conditional distribution of each parameter, conditional on all other parameters, and conditional on the data. We thus need to specify a prior density for each parameter. Recall that the conditional posterior distribution for a subset of parameter vector φ satisfies:

$$l(\varphi|\mathcal{P}(\varphi), \text{data}) \propto l(\text{data}|\mathcal{P}(\varphi))\pi(\varphi),$$

where \mathcal{P} denotes the vector containing all parameters of the model, $\mathcal{P}(\varphi)$ denotes all other parameters, except for φ , and $\pi(\varphi)$ is the prior density of φ .

The Gibbs Sampler also allows for an easy treatment of the latent variables through the so-called data augmentation procedure. Therefore, completion of censored observations becomes possible. In particular, we can never observe the latent variables m_{it}^* and y_{it}^* , and the wage w_{it} is observed only if the i th individual works at time t . Censored or unobserved data are simply “augmented”, that is, we compute m_{it}^* and y_{it}^* based on (1)+(2), conditional on all the parameters.

Finally, the Gibbs Sampler procedure does not involve optimization algorithms. Sequential simulations of the conditional densities are the only computations required. There is somewhat of a complication in this procedure when the densities have no conjugate (i.e., when the prior and the posterior do not belong to the same family). In these cases we use the standard Metropolis-Hastings algorithm. Specifically, in this case we cannot directly draw from the true posterior distribution. Hence, we draw the parameters using another distribution (the proposal distribution) and use a rejection method in order to decide whether or not to keep that draw. In our estimation, we need to resort to the Metropolis-Hastings step when drawing elements for the variance-covariance matrix. While for the prior distribution we use an inverse-Wishart prior distribution for the matrix, the posterior distribution does not belong to this family.

3.2 Application to our Problem

In order to use Bayes’ rule, we first need to specify the full conditional likelihood, that is, the density of all variables, observed and augmented, namely y, w, m, m^* and y^* , given all parameters (the parameters of interest as well as the set of augmented parameters). We thus have to properly define the parameter set and to properly “augment” our data.

The full parameter set is given by

$$(\delta_0^Y, \delta_0^M; \delta^Y, \gamma^M, \gamma^Y; \delta^M, \gamma; \delta^W; \sigma^2, \rho_{yw}, \rho_{ym}, \rho_{wm}; \bar{\gamma}; \eta),$$

so that for \mathcal{P} we have

$$\mathcal{P} = (\delta_0^Y, \delta_0^M; \delta^Y, \gamma^M, \gamma^Y; \delta^M, \gamma; \delta^W; \sigma^2, \rho_{yw}, \rho_{ym}, \rho_{wm}; \bar{\gamma}; \eta; \theta^I),$$

where $\bar{\gamma} = (\gamma'_1, \dots, \gamma'_5)'$ and $\eta = (\eta_1, \dots, \eta_{10})'$.

When specifying the relevant set of variables corresponding to each period, special care need to be given for the (censored) mobility variable. There are four cases depending on the values of the couple (y_{it-1}, y_{it}) . For a given individual i we define X_t , the complete set of endogenous variables, conditional on all parameters and random effects, as:

$$X_t = y_t y_{t-1} X_t^{11} + y_{t-1} (1 - y_t) X_t^{10} + y_t (1 - y_{t-1}) X_t^{01} + (1 - y_t) (1 - y_{t-1}) X_t^{00},$$

where

$$\begin{aligned}
X_t^{11} &= (y_t^*, y_t, w_t, m_{t-1}^*, m_{t-1}), \\
X_t^{10} &= (y_t^*, y_t, m_{t-1}^*), \\
X_t^{01} &= (y_t^*, y_t, w_t), \quad \text{and} \\
X_t^{00} &= (y_t^*, y_t).
\end{aligned}$$

For the initial year we similarly define

$$\begin{aligned}
X_1 &= y_1 X_1^1 + (1 - y_1) X_1^0, \\
X_1^1 &= (y_1^*, y_1, w_1), \quad \text{and} \\
X_1^0 &= (y_1^*, y_1).
\end{aligned}$$

The contribution of the i th individual to the conditional likelihood function is given then by⁴

$$L(\underline{X}_T^i | \mathcal{P}) = \left(\prod_{t=2}^T l(X_{it} | \mathcal{P}, \mathcal{F}_{i,t-1}) \right) l(X_{i1}),$$

where $\underline{X}_{i,t} = (X_{i1}, \dots, X_{it})$, $\mathcal{F}_{i,t-1} = \underline{X}_{i,t-1}$, and

$$\begin{aligned}
l(X_{it} | \mathcal{P}, \mathcal{F}_{i,t-1}) &= l(X_{it}^{11} | \mathcal{P}, \mathcal{F}_{i,t-1})^{y_{i,t-1} y_{it}} l(X_{it}^{10} | \mathcal{P}, \mathcal{F}_{i,t-1})^{y_{i,t-1} (1-y_{it})} \\
&\quad l(X_{it}^{01} | \mathcal{P}, \mathcal{F}_{i,t-1})^{(1-y_{i,t-1}) y_{it}} l(X_{it}^{00} | \mathcal{P}, \mathcal{F}_{i,t-1})^{(1-y_{i,t-1}) (1-y_{it})}.
\end{aligned}$$

Thus, the full conditional likelihood is given by

$$\begin{aligned}
L(\underline{X}_T | \mathcal{P}) &= \left(\frac{1}{V^w} \right)^{\frac{\sum_{i=1}^N \sum_{t=1}^T y_{it}}{2}} \left(\frac{1}{V^m} \right)^{\frac{\sum_{i=1}^N \sum_{t=1}^{T-1} y_{it}}{2}} \\
&\times \prod_{i=1}^N (\mathbf{1}(y_{i1}^* > 0))^{y_{i1}} (\mathbf{1}(y_{i1}^* \leq 0))^{1-y_{i1}} \exp \left\{ -\frac{1}{2} (y_{i1}^* - m_{y_{i1}^*})^2 \right\} \exp \left\{ -\frac{y_{i1}}{2V^w} (w_{i1} - M_{i1}^w)^2 \right\} \\
&\times \prod_{t=2}^T (\mathbf{1}(y_{it}^* > 0))^{y_{it}} (\mathbf{1}(y_{it}^* \leq 0))^{1-y_{it}} \exp \left\{ -\frac{1}{2} (y_{it}^* - m_{y_{it}^*})^2 \right\} \exp \left\{ -\frac{y_{it}}{2V^w} (w_{it}^* - M_{it}^w)^2 \right\} \\
&\times \left((\mathbf{1}(m_{i,t-1}^* > 0))^{m_{i,t-1}} (\mathbf{1}(m_{i,t-1}^* \leq 0))^{1-m_{i,t-1}} \right)^{y_{i,t-1} y_{it}} \exp \left\{ -\frac{y_{i,t-1}}{2V^m} (m_{i,t-1}^* - M_{i,t-1}^m)^2 \right\},
\end{aligned}$$

⁴Even though our notations do not make this explicit, all our computations allow for an individual-specific entry and exit date in the panel, that is, an unbalanced panel.

where

$$\begin{aligned}
V^w &= \sigma^2(1 - \rho_{yw}^2), \\
V^m &= \frac{1 - \rho_{yw}^2 - \rho_{ym}^2 - \rho_{wm}^2 + 2\rho_{yw}\rho_{ym}\rho_{wm}}{1 - \rho_{yw}^2}, \\
M_{it}^m &= m_{m_{it}^*} + \underbrace{\frac{\rho_{y,m} - \rho_{w,m}\rho_{y,w}}{1 - \rho_{y,w}^2}}_a (y_{it}^* - m_{y_{it}^*}) + \underbrace{\frac{\rho_{w,m} - \rho_{y,m}\rho_{y,w}}{\sigma(1 - \rho_{y,w}^2)}}_b (w_{it} - m_{w_{it}}), \\
M_{it}^w &= m_{w_{it}} + \sigma\rho_{y,w}(y_{it}^* - m_{y_{it}^*}),
\end{aligned}$$

and the residuals' correlations are parameterized by:

$$\begin{aligned}
\theta &= (\theta_{yw}, \theta_{ym}, \theta_{wm})', \quad \rho_{yw} = \cos(\theta_{yw}), \quad \rho_{ym} = \cos(\theta_{ym}), \quad \text{and} \\
\rho_{wm} &= \cos(\theta_{yw})\cos(\theta_{ym}) - \sin(\theta_{yw})\sin(\theta_{ym})\cos(\theta_{wm}).
\end{aligned}$$

Finally, we define the various prior distributions as follows:

$$\begin{aligned}
\delta_0^Y &\sim \mathcal{N}(m_{\delta_0^Y}, v_{\delta_0^Y}), \quad \delta_0^M \sim \mathcal{N}(m_{\delta_0^M}, v_{\delta_0^M}), \quad \delta^Y \sim \mathcal{N}(m_{\delta^Y}, v_{\delta^Y}), \quad \delta^W \sim \mathcal{N}(m_{\delta^W}, v_{\delta^W}), \\
\delta^M &\sim \mathcal{N}(m_{\delta^M}, v_{\delta^M}), \quad \gamma^Y \sim \mathcal{N}(m_{\gamma^Y}, v_{\gamma^Y}), \quad \gamma^M \sim \mathcal{N}(m_{\gamma^M}, v_{\gamma^M}), \quad \gamma \sim \mathcal{N}(m_{\gamma}, v_{\gamma}), \\
\sigma^2 &\sim \text{Inverse gamma}\left(\frac{v}{2}, \frac{d}{2}\right), \\
\theta &\sim \text{iid}\mathcal{U}[0, \pi] \\
\eta_j &\sim \text{iid}\mathcal{U}[0, \pi] \quad \text{for } j = 1\dots 10, \quad \text{and} \\
\gamma_j &\sim \text{iid}\mathcal{N}(m_{\gamma_j}, v_{\gamma_j}) \quad \text{for } j = 1\dots 5.
\end{aligned}$$

Based on these prior distributions and the full conditional likelihood, all posterior densities can be evaluated (for a more detailed description see Appendix A).

4 The Data

The data on workers come from two sources, the Déclarations Annuelles de Données Sociales (DADS) and the Echantillon Démographique Permanent (EDP) that are matched together. Our first source, the DADS, is an administrative file based on mandatory reports of employees' earnings by French employers to the Fiscal Administration. Hence, it matches information on workers and on their employing firm. This data set is longitudinal and covers the period 1976-1995 for all workers employed in the private and semi-public sector who were born in October of an even year. Finally, for all workers born in the first four days of October of an even year, information from the EDP is also available. The EDP comprises various censuses and demographic information. These sources are presented in more detail in the following paragraphs.

The DADS Data Set:

Our main data source is the DADS, a large collection of matched employer-employee information collected by the Institut National de la Statistique et des Etudes Economiques (INSEE) and maintained in the Division des Revenus. The data are based upon mandatory employer reports of the gross earnings of each employee subject to French payroll taxes. These taxes apply to all “declared” employees and to all self-employed persons, essentially all employed persons in the economy.

The Division des Revenus prepares an extract of the DADS for scientific analysis, covering all individuals employed in French enterprises who were born in October of even-numbered years, with civil servants excluded.⁵ Our extract covers the period from 1976 through 1995, with 1981, 1983, and 1990 excluded because the underlying administrative data were not sampled in those years. Starting in 1976, the Division des Revenus kept information on the employing firm using the newly created SIREN number from the SIRENE system⁶. However, before this date, there was no available identifier of the employing firm. Each observation of the initial data set corresponds to a unique individual-year-establishment combination. Each observation in this initial DADS file includes an identifier that corresponds to the employee (called ID below), an identifier that corresponds to the establishment (SIRET), and an identifier that corresponds to the parent enterprise of the establishment (SIREN). For each individual, we have information on the number of days during the calendar year the individual worked in the establishment and the full-time/part-time status of the employee. In addition we also have information on the individual’s sex, date and place of birth, occupation, total net nominal earnings during the year and annualized net nominal earnings during the year, as well as the location and industry of the employing establishment. The resulting data set has 13,770,082 observations.

The Echantillon Démographique Permanent:

The Division of Etudes Démographiques at INSEE maintains a large longitudinal data set containing information on many socio-demographic variables of French individuals. All individuals born in the first four days of the month of October of an even year are included in this sample. All questionnaires for these individuals from the 1968, 1975, 1982, and 1990 censuses are gathered into the EDP. The exhaustive long-forms of the various censuses were entered under electronic form only for this fraction of the population living in France (1/4 or 1/5 of the population, depending on the date). The Division des Etudes Démographiques had to find all the censuses questionnaires for these individuals. The INSEE regional agencies were in charge of this task. The usual socio-demographic variables are available in the EDP.⁷

For every individual, education, measured as the highest diploma, and the age at the end of school are collected. Since the categories differ in the three censuses, we first created eight edu-

⁵Individuals employed in the civil service move almost exclusively to other positions within the civil service. Thus the exclusion of civil servants should not affect our estimation of a worker’s market wage equation. For greater details see Abowd, Kramarz, and Margolis (1999).

⁶The SIRENE system is a directory identifying all French firms and their corresponding establishments.

⁷It is important to note that no earnings or income variables have ever been asked in the French censuses.

cation groups (identical to those used in Abowd, Kramarz, and Margolis (1999), namely: (1) No terminal degree; (2) Elementary School; (3) Junior High School; (4) High School; (5) Vocational-Technical School (basic); (6) Vocational Technical School (advanced); (7) Technical College and Undergraduate University; and (8) Graduate School and Other Post-Secondary Education. Other variables that are regularly collected are: nationality (including possible naturalization to French citizenship), country of birth, year of arrival in France, marital status, number of children, employment status (wage-earner in the private sector, civil servant, self-employed, unemployed, inactive, apprentice), spouse’s employment status, information on the equipment of the house or apartment, type of city, location of the residence (region and département).⁸ In some of the censuses, data on parents’ education and social status were collected as well.

In addition to the Census information, all French town-halls in charge of Civil Status registers and ceremonies transmit information to INSEE for the same individuals. This information includes any birth, death, wedding, and divorce involving an individual of the EDP. For each of the above events, additional information on the dates, as well as the occupation of the persons concerned, are collected. Finally, both censuses and civil status information contain the person identifier (ID) of the individual, so the two sources of data can be merged.

Creation of the Matched Data File:

Based on the person identifier, identical in the two datasets (EDP and DADS), it is possible to create a file containing approximately one tenth of the original 1/25th of the population born in October of an even year, i.e., those born in the first four days of that month. Notice that we do not have wages of the civil-servants (even though the census information allows us to determine whether a person is a civil-servant), or the income of self-employed individuals. The individual-level information also contains the employing firm identifier, the so-called SIREN number, that allows us to follow workers from firm to firm and compute the seniority variable. This final data set has approximately 1.5 million observations.

5 The Empirical Results

5.1 Specification and Identification

First, we describe the variables included in each equation. The wage equation is standard for most of its components and includes, in particular, a quadratic function of experience and seniority.⁹ It also includes the following individual characteristics: Sex, marital status, and if unmarried an indicator for living in couple, an indicator for living in the Ile de France region (the Paris region), the département (roughly a U.S. county) unemployment rate, an indicator for French nationality for the person as well as for his/her parents, and cohort effects. We also

⁸ A French “département” corresponds roughly to a county in the U.S. Several départements form a region which is an administrative division.

⁹BFKT also presents estimates with a quartic specification in both experience and seniority. This issue is addressed further below.

include information on the job characteristics: An indicator function for part-time work, and 14 indicators for the industry of the employing firm. In addition we include a complete set of year dummy variables.

Finally, following the specification adopted in BFKT, we include a function, denoted J_{it}^W , that captures the sum of all wage changes that resulted from job changes (i.e., moves between one firm and another) prior to the current date t . This term allows us to control for discontinuous jumps in one's wage when he/she changes jobs. The jumps are allowed to differ depending on the level of seniority and total labor market experience at the point in time when the individual changes jobs. Specifically,

$$J_{it}^W = (\phi_0^s + \phi_0^e e_{i0}) d_{i1} + \sum_{l=1}^{M_{it}} \left[\sum_{j=1}^4 (\phi_{j0} + \phi_j^s s_{t_l-1} + \phi_j^e e_{t_l-1}) d_{jit_l} \right]. \quad (7)$$

Suppressing the i subscript, the variable d_{1t_l} equals 1 if the l th job lasted less than a year, and equals 0 otherwise. Similarly, $d_{2t_l} = 1$ if the l th job lasted between 1 and 5 years, and equals 0 otherwise, $d_{3t_l} = 1$ if the l th job lasted between 5 and 10 years, and equals 0 otherwise, $d_{4t_l} = 1$ if the l th job lasted more than 10 years and equals 0 otherwise. The quantity M_{it} denotes the number of job changes by the i th individual, up to time t (not including the individual's first sample year). If an individual changed jobs in his/her first sample then $d_{i1} = 1$, otherwise $d_{i1} = 0$. The quantities e_t and s_t denote the experience and seniority in year t , respectively.¹⁰ Hence, at the start of a new job, two individuals with identical characteristics, but with different career paths enter their new job with potentially different starting wages.

Turning now to the mobility equation, most of the variables that are included in the wage equation are also present in the mobility equation with the exclusion of the J_{it}^W function. However, an indicator for the lagged mobility decision and indicators for having children between the ages of 0 and 3, and for having children between the ages of 3 and 6 are included in the mobility equation, but are not present in the wage equation.

The specification of the participation equation is very similar to that of the mobility equation. Nevertheless, because job-specific variables cannot be defined for workers who have no job, seniority, the part-time status, and the employing industry, all present in the mobility equation but are excluded from the participation equation. Lagged participation decision (or employment status) is included in the participation equation, whereas this variable is meaningless (i.e., cannot be identified) in the mobility equation since mobility implies participation in both year t and year $t - 1$, as discussed above in Section 2.

Finally, the initial mobility and participation equations are simplified versions of these equa-

¹⁰Note that this specification for the term J_{it}^W produces thirteen different regressors in the wage equation. These regressors are: a dummy for job change in year 1, experience in year 0, the numbers of switches of jobs that lasted less than one year, between 2 and 5 years, between 6 and 10 years, or more than 10 years, seniority at last job change that lasted between 2 and 5 years, between 6 and 10 years, or more than 10 years, and experience at last job change that lasted less than one year, between 2 and 5 years, between 6 and 10 years, or more than 10 years.

tions, that is, all the variables that appear in the general participation and mobility equation are also included in the corresponding initial conditions, except for the lagged dependent variables.

Note the specification above introduces multiple exclusion restrictions. For instance, the industry affiliation is included in both wage and mobility equations but is excluded (for obvious reasons) from the participation equation. Conversely, the children variables are not present in the wage equation but are included in the two other equations. Furthermore, the J_{it}^W function is included in the wage equation but not in the participation and mobility equations. Unfortunately, there appears to be no good exclusion that would guarantee convincing identification of the initial conditions equations, except functional form (i.e., the normality assumptions).

Below we present the estimation results, which are organized as follows. Table 1 presents the estimation results for the wage equation for each of the four education groups. Table 2 presents the estimation results for the participation equation, by education groups, while Table 3 does the same for the inter-firm mobility equation. Table 4 presents the estimates of the variance-covariance matrices for the individual-specific effects (across the five equations) and for the idiosyncratic terms (across the three main equations).¹¹

Tables 5 through 8 in the next section provide a detailed comparison between the results obtained for the U.S. and those obtained for France. Since we have essentially estimated the same model as was estimated by BFKT, we are able to compare parameter estimates for high school dropouts and college graduates in both countries. Table 5 presents estimates for the college graduates in the U.S. and France. Table 6 presents similar estimates for high school dropouts. Table 7 compares the marginal and cumulative returns to experience and seniority at various points of the life cycle for these two groups. Finally, Table 8 presents estimates using two other methods that have been previously used in the literature—a simple OLS and IV method—of the returns to seniority and of the cumulative returns to seniority, for the two groups and in the two countries.

5.2 Certificat d’Etudes Primaires Holders (High School Dropouts)

In France, apart from those quitting the education system without any diploma, the Certificat d’Etudes Primaires (CEP, hereafter) holders are those leaving the system with the lowest possible level of education.¹² They are essentially comparable to High School dropouts in the United States.

Wage Equation:

The results for this group are presented in the first four columns of Table 1. In line 4 of the table we clearly see that the return to seniority is small, only less than 0.3% per year in the first few years (when the linear term is the dominating term). In contrast, the return to experience is almost twenty times larger than the return to tenure.

¹¹Descriptive Statistics are presented in Appendix B, Table B.1.

¹²There is a possibility of some confusion, although relatively small, between not having education and missing response to the question about education in the various censuses.

However, the results of Table 1 also show that the timing of mobility in one’s career matters. First, time spent in a firm makes a significant difference as is indicated from the coefficient estimates in lines 33–36. Moves after relatively short spells are rewarded. There is a 5% increase for change that takes place in less than one year on a job and 10% if the job last between two to five years. However, the part induced by the level of seniority (lines 37–39) is negative. In particular, a move after a two-year spell on a job is better compensated than moves after a five-year spell. The overall jump after 5 years is essentially zero. In comparison, a job spell of between 6 to 10 years of seniority carries neither a penalty nor a reward. For jobs that last more than ten years, workers lose almost 2% per year of seniority (line 39). Finally, there is also the component of wage jump due to experience (lines 40–43). Moves early in one’s work life have a small negative impact on wage gains. In contrast, moves that occur later in one’s career, e.g. after 10 years of experience, add 0.6% for every year of experience, for a total of more than 6% for experience above 10 years. Overall, the loss of earning can be substantial, especially for people who spend long spells in one firm. For example, displaced workers who spent their entire career, of say 20 years, in a single firm face on average a wage loss of about 15% in their new firm (i.e., from lines 36, 39, and 43 we have $100 [1 - \exp \{ .054 - .0167 \cdot 20 + .006 \cdot 20 \}] = 85.2\%$). It is important to note though, that mobility is very low in France, as Table B.1 clearly indicates: An average CEP worker moves only once over the entire period. However, mobility across firms is not evenly distributed over the population. Hence, benefits of voluntary mobility, as well as difficulties that stem from involuntary moves, are confined to a relatively very small fraction of the workers’ population.

Table 1 provides additional facts worthy of notice. Confirming results by Abowd, Kramarz, Lengermann, and Roux (2004), inter-industry wage differences for CEP workers are compressed relatively to the groups with higher levels of education. This is clearly one of the consequences of minimum wage policies in France, because many workers in this group are at the very bottom of the overall wage distribution.

Participation and Mobility Equations:

The first four columns of Tables 2 and 3 present the estimates for the participation and mobility equations, respectively, for the CEP workers. Most of the results are not surprising and are on the order of magnitude that one might expect. For example, having young children lowers the probability of participation significantly, as well as the probability of a move. Also, experience and seniority have an enormous effect on the mobility decision. Individuals with higher experience and tenure are a lot less likely to move.

Also of major interest are the coefficients on the lagged mobility and lagged participation. In contrast with most previous analyses (Altonji and Shakotko (1987), Altonji and Williams (1992 and 1997), and Topel (1991)), we are able to distinguish between state-dependence and unobserved heterogeneity. Not surprisingly, past participation and past mobility favors participation. However, lagged mobility has virtually no impact on the mobility decision. The results in BFKT for the mobility decision imply that a move is optimal only every few years. Hence, a

move in the previous period is associated with lower mobility in the period immediately after (see also the France-U.S. comparison of Table 6). This lack of state-dependence in the current results is obviously a reflection of the French labor market institutions where some workers often go from one short-term contract to another, especially for the CEP group. Unfortunately, as already mentioned above, our data sources provide no information on the nature of the contract, so we cannot examine this point any further.

Stochastic Components:

The first four columns of Table 4 present estimates of variance-covariance components for the CEP group. The individual specific effects for the five-equation model are presented in the first panel, while the results for the residual terms of the three main equations are provided in the second panel. The results clearly show that in terms of the individual specific effects, those who participate more also tend to be high-wage workers. Non-participation (non-employment) and mobility are negatively correlated in terms of individual effects but the corresponding idiosyncratic components are positively correlated. Consequently, we find that high mobility workers tend to be low-employment workers. However, temporary positive shock on mobility (as measured by large draw of the idiosyncratic term) “enhances” participation. Finally, both the idiosyncratic terms and the individual specific effects in the mobility and wage equations are negatively correlated. This implies that high-wage workers tend to be relatively immobile.

It is important to note that most parameters in Table 4 are quite large in absolute terms and are highly significant. This exemplifies the need for the joint estimation adopted here. Joint estimation of these equations clearly has a strong effect on the estimated returns to seniority and experience. Neglecting to control for the joint simultaneous effects would therefore lead to severe bias in the estimated returns to seniority and experience.

5.3 CAP-BEP Holders (Vocational Technical School, basic)

One element that distinguishes the education systems in Continental Europe from that in the U.S., and especially in France and Germany, is the existence of a well-developed apprenticeship training. Indeed, this feature is well-known for Germany but it is also quite important in France. Students who qualify for the Certificat d’Aptitude Professionnelle (CAP) or the Brevet d’Enseignement Professionnel (BEP) have to spend part of their education in firms, and the rest within schools where they are taught both general and vocational subjects. There is no real analog to this system in the U.S.

Wage Equation:

The returns to seniority coefficient, presented in Table 1 (columns 5-8), for workers with vocational/technical education are slightly negative and barely significant. The estimates for the parameters that correspond to the J^W function (lines 33-43) are somewhat different from those obtained for the high school dropouts. Focusing on the two components related to seniority, in lines 33-36, we see that a move after one year in a job brings with it an average increase in wages

of about 3%. Between 2 and 5 years this increase amounts to almost 20%, while moves after 6 to 10 years on a job correspond to an average increase of over 16%. However, the coefficients on the seniority level are all negative (lines 37-39), so the overall jump is much smaller. For example, an individual with 8 years of seniority has an increase of $20\% - 8 \cdot .2\% = 4\%$. Furthermore, for those who have long tenure on the job, say 15 years, there is a significant loss associated with moving from one firm to another, which amounts to approximately $-6\% - 15 \cdot 1.8\% = -33\%$. However, this severe decline is compensated somewhat because of an increase that stems from having more experience. Therefore, for an individual who also has 15 years of experience the overall change in wage is $-6\% - 15 \cdot 1.8\% + 15 \cdot 0.9\% = -19.5\%$. Clearly, relative to the changes observed for the high school dropouts, the losses incurred by the CAP-BEP workers are larger and more significant.

In terms of the return to experience, the experience profile for the CAP-BEP workers is steeper than that observed for the CEP workers, as is apparent from the results reported in lines 2-3 of the table. That is, the CAP-BEP workers tend to accumulate human capital on the job, which is more general and transferable across firms, while the CEP workers gain more firm-specific human capital. Nevertheless, for both groups general human capital is far more important than firm-specific human capital as is demonstrated by the magnitude of the coefficient that relates to experience and seniority.

Participation and Mobility Equations:

As one might expect, the estimated coefficients for the participation and the mobility equations are quite similar to those obtained for the CEP group. In particular, lagged mobility has no significant effect on current mobility. While for the CAP-BEP group longer experience has a more significant effect on the participation decision (see lines 2-3), and lagged participation has a smaller effect. That is, state-dependence is stronger for the less skilled group. The effect of seniority on mobility is virtually the same for the two groups. That is, longer seniority on the job significantly reduces the likelihood of a move.

Stochastic Components:

As Table 4 indicates, the results for main effects for the CAP-BEP group are again quite similar to those obtained for the high school dropout group. In particular, as for the previous group, high-wage workers tend to participate more and they are also a lot less likely to move.

5.4 Baccalauréat Holders (High-School Graduates)

In order to qualify for a high school diploma in France a student has to pass a national exam, called the Baccalauréat. It is a “passport” to higher education, even though not all holders of the Baccalauréat choose to pursue post-secondary education. There are some individuals who choose to attend a university, but never complete the requirement toward a specific degree. We include all these individual in the Baccalauréat Holders group, that is, this group includes all workers who received their Baccalauréat and, in addition, may have had some college education.

The results for this group are presented in columns 9-12 of Tables 1 through 4.

Wage Equation:

The results for this group, presented again in Table 1, display some substantial differences relative to the other two groups discussed above. First, the return to experience is much larger, and, in fact, it is the largest of all groups. However, the return to seniority is, essentially, zero. The estimates for the parameters embedded in the J^W function are quite similar to those previously observed, especially for workers in the the CAP-BEP group. While moves after short spells seems to induce wage increase (see line 33), they also carry some losses of 3.4% per year of seniority (line 37). The overall average change for a worker with 3 years of seniority is hence $12.8\% - 3 \cdot 3.4\% = 2.4\%$. Moves after relatively long employment spells in firms entail large wage losses. For example moves that lasted more than 5 years carry with them a loss of about 1.5% per year of seniority. It can also be seen that the level of experience has very little effect on the initial jump when moving to a new firm, and the effect is usually negative (lines 40-43).

Participation and Mobility Equations:

The results obtained for this group (presented in Tables 2 and 3) are largely consistent with those obtained for the lower education groups discussed above. Nevertheless, the dependence of mobility on lagged mobility becomes marginally negative. This result is consistent with results previously obtained for the U.S., which are discussed below. Also, workers are less mobile the larger their experience and seniority levels, and more pronouncedly so for the high school graduate than for the two lower education groups. Moreover, longer experience has a more significant effect on the participation decision (see lines 2-3) than for the two lower education groups, while lagged participation has a smaller effect. That is, more highly educated workers have a stronger attachment to the labor force. However, state-dependence is a much less significant factor than for the less educated workers.

Stochastic Components:

As was previously observed for the lower education groups, we see that high-wage workers also tend to be high-participation workers. Nevertheless, in contrast to the other two groups, high-wage workers are only marginally low-mobility workers. Indeed, Table B.1 in Appendix B shows that mobility for Baccalauréat holders is the highest among all four education groups, whereas the levels of tenure and experience are the lowest. Part of the reason for these results is that this group contains a disproportionately large number of relatively young individuals. In contrast, the CEP group, for example, includes a relatively large fraction of mature individuals who have, on average, significantly lower levels of education.

5.5 University and Grandes Ecoles Graduates

An important feature that distinguishes the French education system from other European education systems, as well as from the American system, is the existence of a very selective set of educational institutions, known as Grandes Ecoles, that work in parallel with Universities.

The system intends to provide master degrees, mostly in engineering and in business. Unlike the regular university system, the Grandes Ecoles system is very selective and only a relatively small fraction of the relevant population is admitted to the various programs. We include in this group all graduates from both regular universities, as well as graduate from the Grandes Ecoles system, and, for simplicity, we refer to this group as the college graduate group. The results for this education group are provided in the last four columns of Tables 1 through 4.

Wage Equation:

Interestingly, the results for the group of graduates stand in sharp contrast with those obtained for all other education groups. The returns to experience are quite large as for the other groups. However, there is a striking difference in the return to seniority, which is large and significant for the college graduate group. The return to seniority is about 2.6% per year of tenure, with very little curvature. Nevertheless, the return to seniority is small relative to the return experience, namely only one half that of the return to experience.

Lines 33–43 also indicate that the J^W function for the college graduates is quite different from that for the other groups. In general, moves are associated with some loss that is attributed to general experience (see lines 40–43), but it is compensated by a large positive contribution of seniority. For example, a move after a very short employment spell in a firm (up to one year) is associated with an average increase in wages at the new firm of over 18% (line 33). A job change after a spell of 2 to 5 years carries with it an average increase of 5.3%, with an additional increase of over 3% per year of seniority at the time of the move. For a spell lasting between 6 and 10 years the wage increase is even larger and amounts to 17.3%, with an additional increase of close to 1% per year of seniority. A sizeable increase is also evident for larger spells of over 10 years. Clearly, for this highly educated group, seniority is well compensated for.

There are some other additional results worth noting for this group. First, working part-time entails much bigger losses for individuals in this group than for individuals in all other groups. Furthermore, there are sizeable and significant inter-industry wage differences (see lines 30–40). These results are largely consistent with those obtained by Abowd, Kramarz, Lengermann, and Roux (2004). In France minimum wages compress the bottom part of the wage distribution, and wage inequality is confined mostly to the upper part of the distribution, relevant indeed for this college-educated group. Finally, in contrast to all other education groups, it seems that foreign born are being discriminated against relative to their French counterparts (see line 12). While it might be true, in general, that having a higher education allows one to find a job more easily, the wages paid to college graduates who are foreign born is on average 7.5% lower than that paid to French born individuals.

Participation and Mobility Equations:

Similar to the results for the other education groups, lagged mobility seems to have no effect on current mobility. Moreover, experience has virtually no effect on the likelihood of a move. Less experienced workers are no more likely to move than their more experienced counterparts. One possible explanation for this is that while the individuals in this group are not compensated

for their level of experience (see the results for the J^W function) they are well compensated for their seniority, which, in turn, has a significant, yet very small, negative effect on mobility. These results seem to indicate that career paths for engineers and other professionals entail job changes at all ages. Furthermore, in contrast to all other groups, participation choices are only marginally affected by having young children. This might simply indicate that individuals in this group have a stronger attachment to the labor force and hence they choose to acquire much higher levels of education for which they are well compensated.

Stochastic Components:

Similar to what has already been found above for the other educational groups, we find that high-participation individuals are also high-wage individuals. In addition, for the most highly educated workers, we also find that high-wage workers are also low-mobility workers. Also, individuals that are faced with a positive idiosyncratic wage shock tend to be faced with a negative mobility shock. These correlations are similar to those obtained for the other groups. However, this is the only group for which a move entails a jump in wages, compensating them for their seniority.

6 A Comparison with the United States

In this section, we compare our results with those obtained previously by BFKT for the U.S. using very similar model specification. The model was estimated for three education groups: High school dropouts, high school graduates, and college graduates, using an extract from the Panel Study of Income Dynamics (PSID) for the years 1976-1992. Some variables included in BFKT were not available in the panel that we use here (e.g. race), but for the most part very similar definitions were used, especially for the main variables of interest, namely seniority and experience.

6.1 Comparison of Selected Parameters

We present here a comparison of the estimates for a subset of the parameters that are most important. Estimates for the college graduate group are presented in Table 5, while estimates for high school dropouts are presented in Table 6. In each table, the first four columns provide the results for the U.S., while the last four columns report the results for France.

The first and most significant difference between the two studies is in the estimated returns to seniority. They are large and significant in the U.S.: The linear component is around 5% per year for both low and high-education groups. In contrast, these returns are insignificant for all the lower educational groups in France. For the college graduates they are around 2.6%, a lot smaller than for their American counterparts. We also see that the returns to experience are larger in France than they are in the U.S. for the high school dropouts. For the college-educated workers the return to experience are very similar in both countries. Overall, the

combined returns to experiences and seniority are much larger in the U.S. than in France for both groups.

To meaningfully compare wage changes that are associated with a firm-to-firm move we concentrate our discussion on the estimated components of the J^W functions. Some major differences stand out. First, the coefficient on the number of job to job switches appears to indicate that for both education groups, job changes are better compensated in the U.S. than they are in France. For example, for the college-educated workers, who move to a new job from jobs that lasted more than 10 years receive a 60% $((\exp\{.4717\} - 1) * 100)$ increase in wages in their next job. In comparison, the equivalent premium in France is only 6%. This phenomenon is even more pronounced for the high school dropouts: While French workers lose a substantial fraction of their wage after a long tenure in a firm, their American counterparts gain a substantial amount.

Other results on wages are worth noting as well. Inter-industry wage differentials are very small in France for the less educated individuals, but are somewhat more spread for the college graduates. In contrast, the U.S. inter-industry wage differentials are quite large for all education groups (see Table 1 for France, and Table 2 of BFKT for the United States).

There are also significant differences in the mobility processes for the two countries. The mobility process in the U.S. exhibits negative lagged dependence. That is, a worker who just moved is less likely to move in the next period. For France, lagged mobility has virtually no effect at all on current mobility. However, in the U.S., as well as in France, workers tend to move early in a job, as is demonstrated by the negative coefficient on seniority in the mobility equation.

Finally, the comparison of the variance-covariance matrices of individual effects and of the variance-covariance matrices of idiosyncratic effects across the two countries confirms previous findings. First, the U.S. data source (the PSID), because it is a survey, captures initial conditions much better than the French data source (the DADS-EDP), which is largely based on administrative data. More precisely, since individuals are directly interviewed in the PSID, much better data on personal characteristics can be obtained. In France, because the data is administrative, some variables are not available and personal characteristics are likely to be measured with some error. For instance, civil-status and nationality variables come from different sources that can be sometimes contradictory, even though the wage measures and seniority measures are clearly of much better quality in the DADS. In addition, no measure of family income, and very little information on the spouse characteristics, are available in France. Also, imputations of seniority have to be performed in year 1976 for the French data.¹³ Consequently, the correlations between the random terms of the initial condition equations and the other equations are generally weaker for France.

Second, concentrating on the correlation between individual specific effects in the three main

¹³In practice, the conditional expectation of seniority is obtained using the “structure des salaires” survey, see Abowd, Kramarz, and Margolis (1999) for more details.

equations, several facts stand out. In both countries high-wage workers also tend to be high-participation workers. Moreover, high-mobility individuals tend also to be low-wage workers in both countries, but with a much stronger effect in the U.S. (especially for the college graduate workers). This signifies the different roles played by mobility in the two countries as far as wage growth is concerned. France is a country with very low mobility, while mobility across firms is quite common in the U.S. Finally, high-participation workers also tend to be low-mobility, and here again the effect is much stronger in the United States.

6.2 The Returns to Experience and Seniority

To summarize the overall impact of the results presented above, Table 7 presents the estimated cumulative and marginal returns to experience, as well as the cumulative and marginal returns to seniority in the U.S. and France at various points in the life cycle.¹⁴ The cumulative returns to experience are large for both countries, with larger returns in France for both education groups. In complete contrast, the cumulative returns to tenure are much larger in the U.S. For high school dropouts we see that there is absolutely no return to seniority in France. This is somewhat different for the college graduates, even though the cumulative returns to seniority in France are less than half what they are in the U.S.

Note that for the high school dropouts the cumulative return to seniority at 5 years of seniority is almost 27% in the U.S., while in France it is statistically not different from zero. At 15 years of seniority the cumulative return in the U.S. rises to 39%, while in France it remains unchanged. That is, the marginal returns to seniority, at all levels of seniority, are virtually zero in France and hence the cumulative returns remain zeros.

The situation for the college graduate is less pronounced, but as indicated above, the cumulative returns in the U.S. are more than twice as large as those in France. For example, at 5 years of experience the return in the U.S. is almost 24%, while in France it is barely 12%. Similarly, at 15 years of seniority the return in the U.S. is 64%, while in France it is a mere 31%. These differences are mitigated somewhat when one takes into account the return to experiences, but the overall growth of wages in the U.S. that stems from both experience and seniority is still much larger in the U.S. than it is in France.

6.3 Robustness and Specification Checks

We tested various specifications to assess the robustness of the results obtained here. In particular, we examined whether the differences between the U.S. and France stem from inherent differences in the data extracts used for the two countries.

Particular attention was given to investigating to what extent the results obtained here are

¹⁴In all panels, the specifications include a quadratic function of experience and seniority in the wage equation. For the U.S., BFKT compares the estimates with those obtained based on a quartic specification of experience and education. Estimates of the cumulative returns to experience and seniority are very similar. Hence, we resort here to a comparison based on the quadratic version only.

induced by the specific method employed in this paper. Specifically we examine how the results changed relative to those previously obtained in the literature. To do that we use the methods used by Altonji and Williams (1992). We first estimated a wage regression using a simple OLS regression. Then, we estimated the exact same equation using Altonji and Williams’s methodology (specifically the method they label as IV1). The estimation was carried out for both data sets—the PSID for the U.S. and the DADS-EDP for France. The results are reported in Table 8. The top panel of the table presents the IV estimates, while the bottom panel presents the OLS estimates.

First, the OLS estimates of the returns to seniority in France are somewhat smaller than those obtained from our model for the college graduate group. For the high school dropouts they are essentially the same as those obtained by our model, namely zero. For France, the IV method yields point estimates of the returns to seniority that are lower and insignificant. That is, all estimation methods indicate that the returns to seniority in France are quite small, and in most cases are not significantly different from zero. The result for the returns to seniority in the U.S. are strikingly different. The returns to seniority are larger in all specifications than those obtained for France. For both levels of education, the IV method yields the lowest returns to seniority (see the linear tenure effect, but most importantly, the cumulative returns). The OLS estimates for the linear term are slightly larger than those estimated by the IV method. The cumulative returns to seniority have a clear order: The IV method yields the lowest returns. Our estimation method, based on a system of equations, yields the largest, while the OLS estimates are exactly in between, for both groups and both countries.¹⁵

To summarize, all tests show that the returns to seniority and experience are biased when endogeneity is not accounted for. Irrespective of the method used to correct for this endogeneity, the returns to seniority are much larger in the U.S. than in France, and more so for the least educated individuals, who are also most likely to face higher unemployment rates.

6.4 Are the Returns to Seniority an “Incentive Device”?

A natural question arising from the above comparison of the U.S. and France can be formulated as follows: Are the different features that seem to prevail in the two countries related? Do these features lead to lower returns to seniority in France than in the U.S.?

The results presented above indicate that there is relatively low job-to-job mobility in France, while there is relatively high job-to-job mobility in the U.S. In addition the risk of unemployment in France is significantly higher than that faced by the American workers. Can these institutional differences lead to the observed differences in the returns to seniority in the two countries?

In this section, we show that these features are indeed part of a global system and are,

¹⁵BFKT also presents estimates of the returns to experience and seniority without introducing the J^W function. Cumulative returns most often decrease when J^W is excluded. Similar estimates for France for high school dropouts and college graduates (not included here for brevity) show a similar pattern: The linear component of the return to seniority is roughly equal to zero for the former group and equal to about 1% for the latter group.

hence, tightly connected. We use an equilibrium search model with wage-tenure contracts to examine this question. The properties of the wage profiles implied by the model at the stationary equilibrium are contracted using the respective differential characteristics of the two labor markets, namely the U.S. and France.

It turns out that the labor market characteristics that have significant explanatory power are (1) the unemployment rate; and (2) the job arrival rate. The unemployment rate in France has always been larger than that in the U.S. For example, OECD data indicate that the unemployment rates in March of 2004 were 9.4% and 5.7%, for France and the U.S., respectively. It is also been estimated in the literature that the job offer arrival rate in the U.S. is larger than that in most European countries. For example, Jolivet, Postel-Vinay and Robin (2004) use a job search model to provide such estimates. Using the PSID for the U.S. they estimate the job arrival rate to be 1.71 per annum. Similar estimates for several European countries (using the ECHP, for 1994-2001) provide an estimate of 0.56 per annum.

The job search model we employ here has been introduced recently by Burdett and Coles (2003). The most important feature for us in this model is the fact that it generates a unique equilibrium wage-tenure contract. We show below that this wage-tenure contract is such that the slope of tenure in the wage function is an increasing function of the job offer arrival rate. That is, the return to seniority increases when the mobility rate of workers in the economy increases.

We start by summarizing the important aspects of the model. The model is a continuous time model in which individuals are risk adverse. Let λ denote the job offers arrival rate and let δ be the arrival rate of new workers into the labor force and the outflow rate of workers from the labor market. Let p denote the instantaneous revenue received by firms for each worker employed and let b be the instantaneous benefit received by each unemployed worker ($p > b > 0$). Let $u(\cdot)$ denote the instantaneous utility, which is assumed to be strictly increasing and concave. A firm is assumed to offer the same wage contract to all new workers. Also, there is no recall of workers.

Burdett and Coles show that under certain assumptions the implied equilibrium is unique. Also, the optimal wage-tenure contract selected by the firm offering the lower starting wage satisfies

$$\frac{dw}{dt} = \frac{\delta}{\sqrt{p-w_2}} \frac{p-w}{u'(w)} \int_w^{w_2} \frac{u'(s)}{\sqrt{p-s}} ds, \quad (8)$$

with the initial condition $w(0) = w_1$ and where w_1, w_2 are such that

$$\left(\frac{\delta}{\lambda + \delta} \right)^2 = \frac{p-w_2}{p-w_1}, \quad (9)$$

$$u(w_1) = u(b) - \frac{\sqrt{p-w_1}}{2} \int_{w_1}^{w_2} \frac{u'(s)}{\sqrt{p-s}} ds, \quad (10)$$

where $[w_1, w_2]$ is the support of the distribution of wages paid by the firms ($w_1 < b$ and $w_2 < p$).

Assume now that the utility function is constant relative risk aversion (CRRA) of the form $u(w) = w^{1-\sigma}/(1-\sigma)$, where $\sigma > 0$. Burdett and Coles (2003) show that the optimal wage-tenure contract, namely the baseline salary contract, is such that there exists a tenure level such that from that level onward the baseline salary contract is identical to the contract offered by a high-wage firm with a higher entry wage. That is,

$$\frac{d^2 w}{dt^2} = \left(\frac{dw}{dt}\right)^2 \frac{1}{p-w} \left[\frac{\sigma p}{w} - (\sigma + 1)\right] - \delta \frac{\sqrt{p-w}}{\sqrt{p-w_2}} \frac{dw}{dt}, \quad (11)$$

with the initial conditions $w(0) = w_1$ and

$$\frac{dw(0)}{dt} = \frac{\delta}{\sqrt{p-w_2}} \frac{p-w_1}{u'(w_1)} \int_{w_1}^{w_2} \frac{u'(s)}{\sqrt{p-s}} ds. \quad (12)$$

The differential equation (11) is highly non-linear and has to be solved numerically. This can be done by assigning some values for the parameter vector $(\lambda, \delta, \sigma, p)$, and solving the model numerically (e.g. using the procedure `NDSolve` of Mathematica).

In order to study the shape of the wage-tenure contract curve and its sensitivity with respect to the values of the job offers arrival rate, we used the same parameter values as Burdett and Coles (see section 5.2 in their paper). We, set $p = 5$, $\delta/\lambda = 0.1$ and $b = 4.6$. For each value of the relative risk aversion coefficient ($\sigma = 0.2, 0.4, 0.8, 1.4$), we solve the system of equations (11)–(12) numerically for a set a values of the job offer arrival rates. The results are depicted in Figure 1a for $\sigma = 0.2$, in Figure 1b for $\sigma = 0.4$, in Figure 1c for $\sigma = 0.8$, and in Figure 1d for $\sigma = 1.4$. The figures present the wage contract profiles for the first 10 years of seniority. For all values of the relative risk aversion coefficient, we note that the wage profiles are steeper, especially in the first year, when the job offers arrival rates is larger.

The values of the job offer arrival rates (per year) estimated by Jolivet, Postel-Vinay and Robin (2004) correspond to the values $\lambda = 0.005$ for the U.S., and $\lambda = 0.001$ for France. Regardless of the particular value of the relative risk aversion parameter, the equilibrium wage-tenure contract curves are such that there are larger returns to seniority for the high-mobility country, namely the U.S., than the low mobility country, namely France.

Two points are worth noting. First, we take—as firms appear to be doing—institutions that affect mobility as given. For example, the housing market in the U.S. is a lot more developed than in France (because of strong regulations and transaction costs in the latter country). Also, subsidies and government interventions preventing firms from going bankrupt seem to be more prevalent in France, dampening the forces of “creative destruction” in this country. Consequently, French firms face a workforce that is mostly stable with little incentives to move, even after an involuntary separation. Second, as a recent paper by Wasmer (2003) argues, it is more likely that French firms will invest in firm-specific human capital for this exact reason. In contrast, American firms face a workforce that is very mobile. Therefore, these firms should rely on general human capital.

Does this mean that the return to seniority should be larger in France than in the U.S.? Or, putting it differently: Should French firms pay for something they get “by construction” due to strong institutional forces? It seems that there is somewhat of a misconception that has plagued some of the research in this area in recent years. The model discussed here provides a useful tool for the empirical results we obtained in this paper. That is, the optimal return to tenure when individuals are mobile should be larger than when there are not.

7 Conclusions

A central tenet of many theories in labor economics states that compensation should rise with seniority. Nevertheless, there has been much disagreement about the empirical support for this general claim, especially in papers that use data from the U.S. Part of the reason that the empirical research has not lead to a conclusive answer is because of the vast disagreement about the proper method for assessing these theories (see, among others, Altonji and Shakotko (1987), and Topel (1991) for the United States and Abowd, Kramarz, and Margolis (1999) for France).

In this paper we reinvestigate the relations between wages, participation, and firm-to-firm mobility in France. We contrast the result with those obtained in the BFKT analysis that re-examined the return to seniority in the U.S., using the same data source as that used by Topel (1991), Altonji and Shakotko (1987), and Altonji and Shakotko (1992).

We start with a structural model and estimate the return to seniority in a model in which participation, mobility and wages are jointly determined. We include both state-dependence parameters, as well as unobserved correlated individual specific effects in all the model’s equations. To estimate this complex structure, we use a state of the art Bayesian MCMC technique. The model is estimated using French longitudinal data sources for the period 1976-1995 for four separate education groups.

The results indicate that the returns to seniority are virtually zero, and potentially negative for the low education groups. In contrast, the return for the college graduates group is positive and significant, i.e., 2.5% per year of seniority.

We provide a detailed comparison with results previously obtained for the U.S. in BFKT, using the exact same specification used here, and an identical estimation method. The comparison shows that while the returns to seniority are much lower in France than in the U.S., the returns to experience are very close. Furthermore, we find that in both countries there is a significant impact on the estimated returns to seniority when one controls for wage changes when switching from one firm to another (as is summarized by the J^W function introduced here). Hence, we conclude that there is strong evidence that controlling for the individual’s career path and past mobility are essential for proper estimation of the return to seniority.

Additional results show that OLS estimates of the cumulative returns to seniority are lower than those obtained from the system of equations. Furthermore, the same results demonstrate that instrumental variables estimation, following Altonji and Williams (1992), give the smallest

cumulative returns to seniority among all methods used. This is true for both the U.S. and for France. Finally, a comparison between the two countries shows that regardless of the technique used the returns to seniority are lower in France—a low firm-to-firm mobility country—than in the United States—a high firm-to-firm mobility country.

One interpretation of these results is that the returns to seniority are directly related to patterns of mobility. We discuss this aspect using a theoretical framework borrowed from Burdett and Coles (2003). The model clearly indicates that rewarding seniority is likely to play the role of an incentive device designed to counter excessive mobility.

Consequently, the modeling approach adopted here, of jointly estimating the participation, mobility decision along with the wage outcome has non-trivial consequences that may vary across countries. In particular, labor market institutions, state regulations, and other state factors affecting the local economy are likely to have far-reaching effects on the participation, and most importantly the mobility process. This, in turn, is very likely to affect the estimated return to seniority.

8 References

- Abowd J. M., Kramarz F. and D. N. Margolis (1999), "High Wage Workers and High Wage Firms", *Econometrica*, 67, 251-334.
- Abraham K.G. and H.S. Farber (1987), "Job Duration, Seniority, and Earnings", *American Economic Review*, vol. 77, 3, 278-297.
- Altonji J. G. and R. A. Shakotko (1987), "Do Wages Rise With Job Seniority ?", *Review of Economic Studies*, LIV, 437-459.
- Altonji J. G. and N. Williams (1992), "The Effects of Labor Market Experience, Job Seniority and Job Mobility on Wage Growth", NBER Working Paper Series 4133.
- Altonji J. G. and N. Williams (1997), "Do Wages Rise With Job Seniority ? A Reassessment", NBER Working Paper Series 6010.
- Becker G.S. (1964), *Human Capital*, Columbia Press.
- Buchinsky M., Fougère D., Kramarz F. and R. Tchernis (2002), "Interfirm Mobility, Wages, and the Returns to Seniority and Experience in the U.S.", CREST Working Papers 2002-29, Paris.
- Burdett K. and M. Coles (2003), "Equilibrium Wage-Tenure Contracts", *Econometrica*, vol. 71, 5, 1377-1404.
- Flinn C. J. (1986), "Wages and Job Mobility of Young Workers", *Journal of Political Economy*, vol. 94, S88-S110.
- Heckman J. J. (1981), "Heterogeneity and State Dependence", in *Studies in Labor Market*, Rosen S. ed., University of Chicago Press.
- Hyslop D. R. (1999), "State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women", *Econometrica*, 67, 1255-1294.
- Jolivet G., Postel-Vinay F. and J-M. Robin (2004), "The Empirical Content of the Job Search Model: Labor Mobility and Wage Distributions in Europe and the US", Mimeo CREST, Paris.
- Jovanovic B. (1979), "Firm-specific Capital and Turnover", *Journal of Political Economy*, vol. 87, 6, 1246-1260.
- Jovanovic B. (1984), "Matching, Turnover, and Unemployment", *Journal of Political Economy*, vol. 92, 1, 108-122.
- Lazear E. P. (1979), "Why is There Mandatory Retirement?", *Journal of Political Economy*, vol. 87, 6, 1261-1284.
- Lazear E. P. (1981), "Agency, Earnings Profiles, Productivity and Hours Restrictions", *American Economic Review*, vol. 71, 4, 606-620.
- Lazear E. P. (1999), "Personnel Economics: Past Lessons and Future Directions. Presidential Address to the Society of Labor Economics, San Francisco, May 1, 1998.", *Journal of Labor Economics*, vol. 17, 2, 199-236.

- Lillard L. A. and R. J. Willis (1978), “Dynamic Aspects of Earnings Mobility”, *Econometrica*, 46, 985-1012.
- Miller R.A. (1984), “Job Matching and Occupational Choice”, *Journal of Political Economy*, vol. 92, 6, 1086-1120.
- Mincer J. (1974), “Progress in Human Capital Analysis of the Distribution of Earnings”, NBER Working Paper 53.
- Postel-Vinay F. and J-M. Robin (2002), “Equilibrium Wage Dispersion with Worker and Employer Heterogeneity”, *Econometrica*, 70, 2295-2350.
- Salop J. and S. Salop (1976), “Self-Selection and Turnover in the Labor Market”, *Quarterly Journal of Economics*, 90, 619-627.
- Topel R. H. (1991), “Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority”, *Journal of Political Economy*, 99, 145-175.
- Wasmer E. (2003), “Interpreting Europe and US Labor Markets Differences : The Specificity of Human Capital Investments”, CEPR Discussion Paper 3780.

Appendix A

A.1 Mobility equation

Parameter γ : The parameter γ enters $m_{m_{it}^*}$ for $t = 2, \dots, T-1$

$$m_{m_{it}^*} = \gamma m_{it-1} + X_{it}^M \delta^M + \Omega_i^I \theta^{M,I}.$$

The term in the full conditional likelihood that contains this parameter is given by

$$\begin{aligned} & \prod_{i=1}^N \prod_{t=2}^{T-1} \exp\left(-\frac{y_{it}}{2V^m} (m_{it}^* - M_{it}^m)^2\right) \\ &= \exp\left(-\frac{1}{2V^m} \sum_{i=1}^N (\widetilde{m}_i^{*2,T-1} - \widetilde{M}_i^{m2,T-1})' (\widetilde{m}_i^{*2,T-1} - \widetilde{M}_i^{m2,T-1})\right) \\ &= \exp\left(-\frac{1}{2V^m} \sum_{i=1}^N (\widetilde{A}_i^{2,T-1} - \gamma \widetilde{Lm}_i^{2,T-1})' (\widetilde{A}_i^{2,T-1} - \gamma \widetilde{Lm}_i^{2,T-1})\right) \end{aligned}$$

where

$$\begin{aligned} M_{it}^m &= m_{m_{it}^*} + \underbrace{\frac{\rho_{y,m} - \rho_{w,m}\rho_{y,w}}{1 - \rho_{y,w}^2}}_a (y_{it}^* - m_{y_{it}^*}) + \underbrace{\frac{\rho_{w,m} - \rho_{y,m}\rho_{y,w}}{\sigma(1 - \rho_{y,w}^2)}}_b (w_{it} - m_{w_{it}}), \\ \widetilde{m}_i^{*2,T-1} &= (y_{i2}m_{i2}^*, \dots, y_{iT-1}m_{iT-1}^*)', \\ \widetilde{M}_i^{m2,T-1} &= (y_{i2}M_{i2}^m, \dots, y_{iT-1}M_{iT-1}^m)', \quad \text{and} \\ A_{it} &= m_{it}^* - M_{it}^m + \gamma m_{it-1} \\ &= m_{it}^* - X_{it}^M \delta^M - \Omega_i^I \theta^{I,M} - a(y_{it}^* - m_{y_{it}^*}) - b(w_{it} - m_{w_{it}}). \end{aligned}$$

Collecting the squared and crossed terms gives

$$\begin{aligned} V_\gamma^{post,-1} &= V_\gamma^{prior,-1} + \frac{1}{V^m} \sum_{i=1}^N \left(\widetilde{Lm}_i^{2,T-1} \right)' \widetilde{Lm}_i^{2,T-1}, \quad \text{and} \\ V_\gamma^{post,-1} M_\gamma^{post} &= V_\gamma^{prior,-1} M_\gamma^{prior} + \frac{1}{V^m} \sum_{i=1}^N \left(\widetilde{Lm}_i^{2,T-1} \right)' \widetilde{A}_i^{2,T-1}. \end{aligned}$$

Parameter δ^M : We proceed in the same way as for the parameter γ discussed above, to get

$$\begin{aligned} V_{\delta^M}^{post,-1} &= V_{\delta^M}^{prior,-1} + \frac{1}{V^m} \sum_{i=1}^N \left(\widetilde{X}_i^{M2,T-1} \right)' \widetilde{X}_i^{M2,T-1}, \quad \text{and} \\ V_{\delta^M}^{post,-1} M_{\delta^M}^{post} &= V_{\delta^M}^{prior,-1} M_{\delta^M}^{prior} + \frac{1}{V^m} \sum_{i=1}^N \left(\widetilde{X}_i^{M2,T-1} \right)' \widetilde{A}_i^{2,T-1}, \end{aligned}$$

where

$$\begin{aligned} A_{it} &= m_{it}^* - M_{it}^m + \delta^M X_{it}^M \\ &= m_{it}^* - \gamma m_{it-1} - \Omega_i^I \theta^{I,M} - a(y_{it}^* - m_{y_{it}^*}) - b(w_{it} - m_{w_{it}}). \end{aligned}$$

A.2 Wage equation

Parameter δ^W : Note that we have to take into account that the parameter δ^W enters both $m_{w_{it}}$, for $t = 1 \dots T$, as well as M_{it}^m , for $t = 1 \dots T - 1$. The corresponding terms in the full conditional likelihood are given by

$$\begin{aligned} &\prod_{i=1}^N \exp\left(-\frac{1}{2V^w} \sum_{t=1}^T y_{it}(w_{it} - M_{it}^w)^2\right) \exp\left(-\frac{1}{2V^m} \sum_{t=1}^{T-1} y_{it}(m_{it}^* - M_{it}^m)^2\right) \\ &= \prod_{i=1}^N \exp\left(-\frac{1}{2V^w} \sum_{t=1}^T y_{it}(A_{it} - X_{it}^W \delta^W)^2\right) \exp\left(-\frac{1}{2V^m} \sum_{t=1}^{T-1} y_{it}(B_{it} + bX_{it}^W \delta^W)^2\right), \end{aligned}$$

where

$$\begin{aligned} w_{it} - M_{it}^w &= A_{it} - X_{it}^W \delta^W, \quad \text{and} \\ m_{it}^* - M_{it}^m &= B_{it} + bX_{it}^W \delta^W, \end{aligned}$$

which is equivalent to

$$\begin{aligned} A_{it} &= w_{it} - \Omega_i^I \theta^{I,W} - \rho_{y,w} \sigma(y_{it}^* - m_{y_{it}^*}), \quad \text{and} \\ B_{it} &= m_{it}^* - m_{m_{it}^*} - a(y_{it}^* - m_{y_{it}^*}) - b(w_{it} - \Omega_i^I \theta^{I,W}). \end{aligned}$$

Using analogous notation to the ones used above we have

$$\begin{aligned} V_{\delta^W}^{post,-1} &= V_{\delta^W}^{prior,-1} + \frac{1}{V^w} \sum_{i=1}^N \left(\widetilde{\underline{X}}_i^{W,1,T} \right)' \widetilde{\underline{X}}_i^{W,1,T} + \frac{b^2}{V^m} \sum_{i=1}^N \left(\widetilde{\underline{X}}_i^{W,1,T-1} \right)' \widetilde{\underline{X}}_i^{W,1,T-1}, \quad \text{and} \\ V_{\delta^W}^{post,-1} M_{\delta^W}^{post} &= V_{\delta^W}^{prior,-1} M_{\delta^W}^{prior} + \frac{1}{V^w} \sum_{i=1}^N \left(\widetilde{\underline{X}}_i^{W,1,T} \right)' \widetilde{\underline{A}}_i - \frac{b}{V^m} \sum_{i=1}^N \left(\widetilde{\underline{X}}_i^{W,1,T-1} \right)' \widetilde{\underline{B}}_i. \end{aligned}$$

A.3 Participation equation

Parameter γ^Y : We have to take into account that γ^Y enters both $m_{y_{it}^*}$ for $t = 2 \dots T$, M_{it}^w for $t = 2 \dots T$ and M_{it}^m for $t = 2 \dots T - 1$. The corresponding terms in the full conditional likelihood are

$$\begin{aligned} &\prod_{i=1}^N \exp\left(-\frac{1}{2} \sum_{t=2}^T (y_{it}^* - m_{y_{it}^*})^2 - \frac{1}{2V^w} \sum_{t=2}^T y_{it}(w_{it} - M_{it}^w)^2 - \frac{1}{2V^m} \sum_{t=2}^{T-1} y_{it}(m_{it}^* - M_{it}^m)^2\right) \\ &= \prod_{i=1}^N \exp\left(-\frac{1}{2} \sum_{t=2}^T (A_{it} - \gamma^Y L y_{it})^2 - \frac{1}{2V^w} \sum_{t=2}^T y_{it}(B_{it} + \rho_{y,w} \sigma \gamma^Y L y_{it})^2 - \frac{1}{2V^m} \sum_{t=2}^{T-1} y_{it}(C_{it} + a \gamma^Y L y_{it})^2\right), \end{aligned}$$

where

$$\begin{aligned} y_{it}^* - m_{y_{it}}^* &= A_{it} - \gamma^Y Ly_{it}, \\ w_{it} - M_{it}^w &= B_{it} + \rho_{y,w} \sigma \gamma^Y Ly_{it}, \quad \text{and} \\ m_{it}^* - M_{it}^m &= C_{it} + a \gamma^Y Ly_{it}, \end{aligned}$$

which is equivalent to

$$\begin{aligned} A_{it} &= y_{it}^* - \gamma^M Lm_{it} - X_{it}^Y \delta^Y - \Omega_i^I \theta^{I,Y}, \\ B_{it} &= w_{it} - m_{w_{it}} - \rho_{y,w} \sigma A_{it}, \quad \text{and} \\ C_{it} &= m_{it}^* - m_{m_{it}}^* - b(w_{it} - m_{w_{it}}) - a A_{it}. \end{aligned}$$

Using analogous notation to the notations used above, we have

$$\begin{aligned} V_{\gamma^Y}^{post,-1} &= V_{\gamma^Y}^{prior,-1} + \sum_{i=1}^N \left(\underline{Ly}_i^{2,T} \right)' \underline{Ly}_i^{2,T} + \frac{\rho_{y,w}^2 \sigma^2}{V^w} \sum_{i=1}^N \left(\widetilde{\underline{Ly}}_i^{2,T} \right)' \widetilde{\underline{Ly}}_i^{2,T} \\ &+ \frac{a^2}{V^m} \sum_{i=1}^N \left(\widetilde{\underline{Ly}}_i^{2,T-1} \right)' \widetilde{\underline{Ly}}_i^{2,T-1}, \quad \text{and} \\ V_{\gamma^Y}^{post,-1} M_{\gamma^Y}^{post} &= V_{\gamma^Y}^{prior,-1} M_{\gamma^Y}^{prior} + \sum_{i=1}^N \left(\underline{Ly}_i^{2,T} \right)' \underline{A}_i^{2,T} - \frac{\rho_{y,w} \sigma}{V^w} \sum_{i=1}^N \left(\widetilde{\underline{Ly}}_i^{2,T} \right)' \widetilde{\underline{B}}_i^{2,T} \\ &- \frac{a}{V^m} \sum_{i=1}^N \left(\widetilde{\underline{Ly}}_i^{2,T-1} \right)' \widetilde{\underline{C}}_i^{2,T-1}. \end{aligned}$$

Parameter γ^M : We proceed in a similar fashion with the parameter γ^M , to get

$$\begin{aligned} V_{\gamma^M}^{post,-1} &= V_{\gamma^M}^{prior,-1} + \sum_{i=1}^N \left(\underline{Lm}_i^{2,T} \right)' \underline{Lm}_i^{2,T} \\ &+ \frac{\rho_{y,w}^2 \sigma^2}{V^w} \sum_{i=1}^N \left(\widetilde{\underline{Lm}}_i^{2,T} \right)' \widetilde{\underline{Lm}}_i^{2,T} + \frac{a^2}{V^m} \sum_{i=1}^N \left(\widetilde{\underline{Lm}}_i^{2,T-1} \right)' \widetilde{\underline{Lm}}_i^{2,T-1}, \\ V_{\gamma^M}^{post,-1} M_{\gamma^M}^{post} &= V_{\gamma^M}^{prior,-1} M_{\gamma^M}^{prior} + \sum_{i=1}^N \left(\underline{Lm}_i^{2,T} \right)' \underline{A}_i^{2,T}, \quad \text{and} \\ &- \frac{\rho_{y,w} \sigma}{V^w} \sum_{i=1}^N \left(\widetilde{\underline{Lm}}_i^{2,T} \right)' \widetilde{\underline{B}}_i^{2,T} - \frac{a}{V^m} \sum_{i=1}^N \left(\widetilde{\underline{Lm}}_i^{2,T-1} \right)' \widetilde{\underline{C}}_i^{2,T-1}, \end{aligned}$$

where

$$\begin{aligned} A_{it} &= y_{it}^* - \gamma^Y Ly_{it} - X_{it}^Y \delta^Y - \Omega_i^I \theta^{I,Y}, \\ B_{it} &= w_{it} - m_{w_{it}} - \rho_{y,w} \sigma (A_{it}), \quad \text{and} \\ C_{it} &= m_{it}^* - m_{m_{it}}^* - b(w_{it} - m_{w_{it}}) - a (A_{it}). \end{aligned}$$

Parameter δ^Y : We proceed the same way and we get

$$\begin{aligned}
V_{\delta^Y}^{post,-1} &= V_{\delta^Y}^{prior,-1} + \sum_{i=1}^N \left(\underline{X}_i^{Y2,T} \right)' \underline{X}_i^{Y2,T} + \frac{\rho_{y,w}^2 \sigma^2}{Vw} \sum_{i=1}^N \left(\widetilde{X}_i^{Y2,T} \right)' \widetilde{X}_i^{Y2,T} \\
&\quad + \frac{a^2}{Vm} \sum_{i=1}^N \left(\widetilde{X}_i^{Y2,T-1} \right)' \widetilde{X}_i^{Y2,T-1} \\
V_{\delta^Y}^{post,-1} M_{\delta^Y}^{post} &= V_{\delta^Y}^{prior,-1} M_{\delta^Y}^{prior} + \sum_{i=1}^N \left(\underline{X}_i^{Y2,T} \right)' \underline{A}_i^{2,T} \\
&\quad - \frac{\rho_{y,w} \sigma}{Vw} \sum_{i=1}^N \left(\widetilde{X}_i^{Y2,T} \right)' \widetilde{B}_i^{2,T} - \frac{a}{Vm} \sum_{i=1}^N \left(\widetilde{X}_i^{Y2,T-1} \right)' \widetilde{C}_i^{2,T-1}
\end{aligned}$$

where

$$\begin{aligned}
A_{it} &= y_{it}^* - \gamma^Y L y_{it} - \gamma^M L m_{it} - \Omega_i^I \theta^{I,Y}, \\
B_{it} &= w_{it} - m_{w_{it}} - \rho_{y,w} \sigma (A_{it}), \\
C_{it} &= m_{it}^* - m_{m_{it}} - b(w_{it} - m_{w_{it}}) - a(A_{it}).
\end{aligned}$$

A.4 Initial equations

Parameter δ_0^M : Similarly, for the parameter δ_0^M , which enters only m_{i1}^* we have

$$\begin{aligned}
V_{\delta_0^M}^{post,-1} &= V_{\delta_0^M}^{prior,-1} + \frac{1}{Vm} \sum_{i=1}^N \left(\widetilde{X}_{i1}^M \right)' \widetilde{X}_{i1}^M, \quad \text{and} \\
V_{\delta_0^M}^{post,-1} M_{\delta_0^M}^{post} &= V_{\delta_0^M}^{prior,-1} M_{\delta_0^M}^{prior} + \frac{1}{Vm} \sum_{i=1}^N \left(\widetilde{X}_{i1}^M \right)' \widetilde{A}_{i1},
\end{aligned}$$

where

$$A_{i1} = m_{i1}^* - \Omega_i^I \alpha^{I,M} - a(y_{i1}^* - m_{y_{i1}}) - b(w_{i1} - m_{w_{i1}}).$$

Parameter δ_0^Y : We proceed the same way for δ_0^Y and get

$$\begin{aligned}
V_{\delta_0^Y}^{post,-1} &= V_{\delta_0^Y}^{prior,-1} + \sum_{i=1}^N X_{i1}^{Y'} X_{i1}^Y + \left(\frac{\rho_{y,w}^2 \sigma^2}{Vw} + \frac{a^2}{Vm} \right) \sum_{i=1}^N \widetilde{X}_{i1}^{Y'} \widetilde{X}_{i1}^Y, \quad \text{and} \\
V_{\delta_0^Y}^{post,-1} M_{\delta_0^Y}^{post} &= V_{\delta_0^Y}^{prior,-1} M_{\delta_0^Y}^{prior} + \sum_{i=1}^N X_{i1}^{Y'} A_i - \sum_{i=1}^N \widetilde{X}_{i1}^{Y'} \left(\frac{\rho_{y,w} \sigma}{Vw} \widetilde{B}_i + \frac{a}{Vm} \widetilde{C}_i \right),
\end{aligned}$$

where

$$\begin{aligned}
A_i &= y_{i1}^* - \Omega_i^E \alpha^{I,Y}, \\
B_i &= w_{i1} - m_{w_{i1}} - \rho_{y,w} \sigma A_i, \quad \text{and} \\
C_i &= m_{i1}^* - m_{m_{i1}} - b(w_{i1} - m_{w_{i1}}) - a A_i.
\end{aligned}$$

A.5 Latent variables

Latent participation y_{it}^* : Several cases are possible here for y_{it}^* , depending on whether $y_{it} = 1$ or $y_{it} = 0$.

1. For $t = 1 \dots T - 1$ we have the following:

- (a.) If $y_{it} = 1$, then

$$\begin{aligned} y_{it}^* &\sim \mathcal{NT}_{\mathbb{R}^+}(M^{Apost}, V^{Apost}), \\ V^{Apost, -1} M^{Apost} &= \left(\frac{\sigma \rho_{v,\epsilon}}{Vw} - \frac{ab}{Vm}\right)(w_{it} - m_{w_{it}}) + \frac{a}{Vm}(m_{it}^* - m_{m_{it}^*}) + \left(\frac{\sigma^2 \rho_{v,\epsilon}^2}{Vw} + \frac{a^2}{Vm} + 1\right)m_{y_{it}^*}, \\ V^{Apost} &= \frac{1}{1 + \frac{a^2}{Vm} + \frac{\sigma^2 \rho_{v,\epsilon}^2}{Vw}}. \end{aligned}$$

- (b.) If $y_{it} = 0$, then

$$y_{it}^* \sim \mathcal{NT}_{\mathbb{R}^-}(m_{y_{it}^*}, 1).$$

2. For $t = T$ we have the following:

- (a.) If $y_{iT} = 1$, then

$$\begin{aligned} y_{iT}^* &\sim \mathcal{NT}_{\mathbb{R}^+}(M^{Apost}, 1 - \rho_{v,\epsilon}^2), \\ M^{Apost} &= (1 - \rho_{v,\epsilon}^2) \left(m_{y_{iT}^*} \left(1 + \frac{\sigma^2 \rho_{v,\epsilon}^2}{Vw}\right) + \frac{\sigma \rho_{v,\epsilon}}{Vw} (w_{iT} - m_{w_{iT}}) \right). \end{aligned}$$

- (b.) If $y_{iT} = 0$, then

$$y_{iT}^* \sim \mathcal{NT}_{\mathbb{R}^-}(m_{y_{iT}^*}, 1).$$

Latent mobility m_{it}^* : Two conditions must be checked: First, $t = 1 \dots T - 1$, and second, it must be that $y_{it} = 1$. When these two conditions are met, we distinguish between several different cases:

1. If $y_{it+1} = 0$, then

$$m_{it}^* \sim \mathcal{N}(M_{it}^m, V^m) \quad \text{and} \quad m_{it} = \mathbb{I}(m_{it}^* > 0).$$

2. If $y_{it+1} = 1$, then: (a.) If $m_{it} = 1$, then

$$m_{it}^* \sim \mathcal{NT}_{\mathbb{R}^+}(M_{it}^m, V^m).$$

- (b.) If $m_{it} = 0$, then

$$m_{it}^* \sim \mathcal{NT}_{\mathbb{R}^-}(M_{it}^m, V^m).$$

A.6 Variance-Covariance Matrix of Residuals

Because the prior distribution is not conjugate (i.e., the posterior distribution does not belong to the same family of distributions as the prior), we have to resort to the Metropolis-Hastings algorithm.

Variance-Covariance Matrices of Individual Effects $\Sigma_i^I(\dots); z; y, w$:

The parameters $\eta_j, j = 1 \dots 10$ and $\gamma_j, j = 1 \dots 5$ do not enter the full conditional likelihood. They only enter the prior distributions. Let us denote by p the parameter we are interested in among the $\eta_j, j = 1 \dots 10$ and $\gamma_j, j = 1 \dots 5$. Then,

$$\begin{aligned}
l(p|(-p), \theta^I) &= l(\theta^I|p)\pi^0(p) \\
&= \pi^0(p) \prod_{i=1}^N l(\theta_i^I|\Sigma_i^I(p)) \\
&\propto \pi^0(p) \prod_{i=1}^N \frac{1}{\sqrt{\det(\Sigma_i^I(p))}} \exp\left(-\frac{1}{2}\theta_i^{I'}\Sigma_i^{I,-1}(p)\theta_i^I\right).
\end{aligned}$$

We face non-conjugate distributions therefore we use the independent Metropolis-Hastings algorithm with the prior distribution as the instrumental distribution.

A.7 Individual effects

The likelihood terms that include θ^I are given by

$$\begin{aligned}
&\prod_{i=1}^N \exp\left(-\frac{1}{2}(y_{i1}^* - m_{y_{i1}^*})^2\right) \exp\left(-\frac{y_{i1}}{2V^w}(w_{i1} - M_{i1}^w)^2\right) \\
&\times \prod_{t=2}^T \exp\left(-\frac{1}{2}(y_{it}^* - m_{y_{it}^*})^2\right) \exp\left(-\frac{y_{it}}{2V^w}(w_{it} - M_{it}^w)^2\right) \exp\left(-\frac{y_{it-1}}{2V^m}(m_{it-1}^* - M_{it-1}^m)^2\right),
\end{aligned}$$

where

$$\begin{aligned}
M_{it}^m &= m_{m_{it}^*} + a(y_{it}^* - m_{y_{it}^*}) + b(w_{it} - m_{w_{it}}), \quad \text{and} \\
M_{it}^w &= m_{w_{it}} + \sigma\rho_{v,\varepsilon}(y_{it}^* - m_{y_{it}^*}).
\end{aligned}$$

The following notations are useful:

1. In the first term:

$$\begin{aligned}
(y_{i1}^* - m_{y_{i1}^*})^2 &= (A_{i1} - \Omega_i^I \alpha^{I,Y})^2, \\
A_{i1} &= y_{i1}^* - XY_{i1} \delta_0^Y.
\end{aligned}$$

2. In the second term:

$$\begin{aligned}
y_{i1}(w_{i1} - M_{w_{i1}})^2 &= y_{i1}(B_{i1} - \Omega_i^I \theta^{I,W} + \rho_{v,\varepsilon} \sigma \Omega_i^I \alpha^{I,Y})^2, \\
B_{i1} &= w_{i1} - XW_{i1} \delta^w - \rho_{v,\varepsilon} \sigma (y_{i1}^* - XY_{i1} \delta_0^Y), \\
\tilde{B}_{i1} &= y_{i1} B_{i1}, \\
\tilde{\Omega}_{i1}^I &= y_{i1} \Omega_i^I.
\end{aligned}$$

3. In the third term:

$$\begin{aligned}
(y_{it}^* - m_{y_{it}^*})^2 &= (C_{it} - \Omega_i^I \theta^{Y,I})^2, \\
C_{it} &= y_{it}^* - XY_{it} \delta^Y - \gamma^Y y_{it-1} - \gamma^M m_{it-1}.
\end{aligned}$$

4. In the fourth term:

$$\begin{aligned}
y_{it}(w_{it} - M_{w_{it}})^2 &= y_{it}(D_{it} - \Omega_i^I \theta^{W,I} + \rho_{v,\varepsilon} \sigma \Omega_i^I \theta^{Y,I})^2, \\
D_{it} &= w_{it} - XW_{it} \delta^w - \rho_{v,\varepsilon} \sigma C_{it}, \\
\tilde{D}_{it} &= y_{it} D_{it}, \\
\tilde{\Omega}_{it}^I &= y_{it} \Omega_i^I.
\end{aligned}$$

5. In the fifth term:

For $t > 1$:

$$\begin{aligned}
y_{it}(m_{it}^* - M_{m_{it}^*})^2 &= y_{it}(F_{it} + \Omega_i^I (-\theta_{M,I} + a\theta^{Y,I} + b\theta^{W,I}))^2, \\
F_{it} &= m_{it}^* - \gamma m_{it-1} - XM_{it} \delta^M - aC_{it} - b(w_{it} - XW_{it} \delta^w), \\
\tilde{F}_{it} &= y_{it} F_{it}.
\end{aligned}$$

For $t = 1$:

$$\begin{aligned}
y_{i1}(m_{i1}^* - M_{m_{i1}^*})^2 &= y_{i1}(G_{i1} + \Omega_i^I (-\alpha_{M,I} + a\alpha^{Y,I} + b\theta^{W,I}))^2, \\
G_{i1} &= m_{i1}^* - XM_{i1} \delta_0^M - aA_{i1} - b(w_{i1} - XW_{i1} \delta^w), \\
\tilde{G}_{i1} &= y_{i1} G_{i1}.
\end{aligned}$$

The posterior distribution satisfies then:

$$\begin{aligned}
l(\theta^E | \dots) &\propto \exp\left(-\frac{1}{2} \theta^{I'} D^{I,-1} \theta^I\right) \\
&\times \exp\left(-\frac{1}{2} \sum_{i=1}^n (A_{i1} - \Omega_i^I \alpha^{Y,I})^2 - \frac{1}{2V^w} \sum_i \left(\tilde{B}_{i1} - \tilde{\Omega}_{i1}^I (\theta^{W,I} - \rho_{v,\varepsilon} \sigma \alpha^{Y,I})\right)^2\right) \\
&\times \exp\left(-\frac{1}{2} \sum_i \sum_{t=2}^T (C_{it} - \Omega_i^I \theta^{Y,I})^2 - \frac{1}{2V^w} \sum_i \sum_{t=2}^T \left(\tilde{D}_{it} - \tilde{\Omega}_{it}^I (\theta^{W,I} - \rho_{v,\varepsilon} \sigma \theta^{Y,I})\right)^2\right) \\
&\times \exp\left(-\frac{1}{2V^m} \sum_i \left(\tilde{G}_{i1} + \tilde{\Omega}_{i1}^I (-\alpha^{M,I} + a\alpha^{Y,I} + b\theta^{W,I})\right)^2\right) \\
&\times \exp\left(-\frac{1}{2V^m} \sum_i \sum_{t=2}^{T-1} \left(\tilde{F}_{it} + \tilde{\Omega}_{it}^I (-\theta^{M,I} + a\theta^{Y,I} + b\theta^{W,I})\right)^2\right).
\end{aligned}$$

We define several projection operators. Let $P_1 = (I_J, 0_J, 0_J, 0_J, 0_J)$, so that $P_1 \theta^I = \alpha^{I,Y}$, and similarly, define P_2, \dots, P_5 , so that $P_2 \theta^I = \alpha^{I,M}$, $P_3 \theta^I = \theta^{I,Y}$, $P_4 \theta^I = \theta^{I,W}$, and $P_5 \theta^I = \theta^{I,M}$.

Appendix B

Table B.1: Descriptive Statistics

No.	Variable	HS Dropouts (CEP)		Vocational HS (CAP-BEP)		HS Grad. (Baccalauréat)		College Graduates	
		Mean	St.Dev	Mean	St.Dev	Mean	St.Dev	Mean	St.Dev
1.	Participation	0.5045	0.5000	0.5860	0.4926	0.4604	0.4984	0.4700	0.4991
2.	Log wage	4.0874	0.8093	4.1991	0.7463	4.1573	0.9801	4.7006	1.1599
3.	Mobility	0.0752	0.2637	0.0938	0.2916	0.1232	0.3287	0.0952	0.2935
4.	Tenure	7.6476	7.8048	5.9895	6.9210	4.1165	6.1718	6.2467	7.7694
5.	Experience	24.7733	11.4711	17.4464	10.7499	11.7819	9.9945	15.7837	9.6285
6.	Lives in couple	0.0648	0.2461	0.0613	0.2399	0.0661	0.2484	0.0551	0.2281
7.	Married	0.6347	0.4815	0.5650	0.4958	0.3540	0.4782	0.5706	0.4950
8.	Children between 0 and 3	0.0863	0.2809	0.1317	0.3381	0.1019	0.3025	0.1225	0.3279
9.	Children between 3 and 6	0.0898	0.2859	0.1142	0.3180	0.0749	0.2633	0.1096	0.3125
10.	Number of Children	1.3196	1.3771	1.0712	1.2158	0.5873	0.9889	0.9830	1.4218
11.	Live in Ile de France	0.1196	0.3245	0.1124	0.3160	0.1531	0.3600	0.2088	0.4064
12.	Live in Paris	0.1182	0.3228	0.1120	0.3153	0.1532	0.3602	0.2707	0.4443
13.	Live in a town	0.2033	0.4024	0.2167	0.4120	0.2417	0.4281	0.2251	0.4176
14.	Live in rural area	0.6785	0.4670	0.6714	0.4697	0.6051	0.4888	0.5043	0.5000
15.	Part-time	0.1846	0.3880	0.1528	0.3598	0.2394	0.4267	0.2137	0.4100
16.	Local unemp. rate	8.1940	3.6354	8.3736	3.4322	8.2858	3.5824	8.9162	2.7771

Table B.1: Descriptive Statistics (Continued)

No.	Variable	HS Dropouts (CEP)		Vocational HS (CAP-BEP)		HS Grad. (Baccalauréat)		College Graduates	
		Mean	St.Dev	Mean	St.Dev	Mean	St.Dev	Mean	St.Dev
17.	Agriculture	0.0424	0.2016	0.0379	0.1909	0.0216	0.1454	0.0130	0.1133
18.	Energy	0.0095	0.0969	0.0164	0.1272	0.0117	0.1076	0.0219	0.1463
19.	Intermediate goods	0.1007	0.3009	0.0960	0.2946	0.0489	0.2157	0.0634	0.2436
20.	Equipment goods	0.1122	0.3157	0.1305	0.3368	0.0691	0.2536	0.1277	0.3337
21.	Consumption goods	0.1127	0.3162	0.0741	0.2620	0.0551	0.2281	0.0555	0.2289
22.	Construction	0.0884	0.2840	0.1159	0.3201	0.0387	0.1930	0.0357	0.1856
23.	Retail and wholesale	0.1611	0.3676	0.1436	0.3507	0.1554	0.3623	0.0939	0.2917
24.	Transport	0.0606	0.2385	0.0629	0.2428	0.0699	0.2550	0.0531	0.2242
25.	Market services	0.2083	0.4061	0.2222	0.4158	0.3082	0.4617	0.3383	0.4731
26.	Insurance	0.0079	0.0886	0.0080	0.0893	0.0197	0.1390	0.0156	0.1238
27.	Banking and Finance	0.0155	0.1234	0.0213	0.1445	0.0558	0.2296	0.0452	0.2077
28.	Non-market services	0.0748	0.2630	0.0661	0.2484	0.1396	0.3466	0.1306	0.3370
29.	Born before 1929	0.2380	0.4258	0.0540	0.2261	0.0452	0.2078	0.0956	0.2941
30.	Born between 1930 and 1939	0.2132	0.4096	0.1215	0.3268	0.0480	0.2137	0.1314	0.3378
31.	Born between 1940 and 1949	0.2290	0.4202	0.2013	0.4010	0.1048	0.3063	0.2867	0.4522
32.	Born between 1950 and 1959	0.2329	0.4227	0.3084	0.4618	0.2123	0.4089	0.3266	0.4690
33.	Born between 1960 and 1969	0.0646	0.2459	0.2758	0.4469	0.4476	0.4973	0.1580	0.3648
34.	Born after 1970	0.0222	0.1475	0.0390	0.1935	0.1421	0.3492	0.0017	0.0417
	Number of Observations	32,596		12,405		34,071		7,579	

Notes: Source is the DADS-EDP from 1976 through 1996.

Table 1: Wage Equation

No.	Variable	High School Dropouts (CEP)			Vocational HS (CAP-BEP)			HS Grad. (Baccalauréat)			College Graduates						
		Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max				
Job characteristics:																	
1.	Constant	2.3537	0.057	2.150	2.634	2.5619	0.052	2.352	2.738	2.7754	0.050	2.594	2.949	2.5562	0.061	2.324	2.782
2.	Experience	0.0504	0.004	0.037	0.066	0.0590	0.003	0.046	0.070	0.0764	0.005	0.059	0.096	0.0537	0.004	0.041	0.067
3.	Exper. sq.	-0.0005	0.000	-0.001	-0.000	-0.0006	0.000	-0.001	-0.000	-0.0009	0.000	-0.001	-0.001	-0.0008	0.000	-0.001	-0.001
4.	Seniority	0.0027	0.003	-0.009	0.014	-0.0046	0.003	-0.015	0.006	-0.0062	0.005	-0.024	0.010	0.0264	0.003	0.016	0.038
5.	Sen. sq.	-0.0002	0.000	-0.001	0.000	-0.0002	0.000	-0.001	0.000	-0.0001	0.000	-0.001	0.000	-0.0004	0.000	-0.001	-0.000
6.	Part-time	-0.5199	0.012	-0.563	-0.472	-0.4432	0.011	-0.481	-0.404	-0.6559	0.013	-0.706	-0.600	-0.8207	0.015	-0.882	-0.768
Individual and family characteristics:																	
7.	Male	0.5242	0.035	0.382	0.643	0.4111	0.034	0.268	0.304	0.3811	0.038	0.258	0.518	0.5480	0.040	0.404	0.713
8.	Married	-0.0059	0.014	-0.057	0.046	0.0216	0.012	-0.023	0.000	-0.0070	0.018	-0.070	0.060	0.1960	0.022	0.112	0.288
9.	Couple	0.0101	0.019	-0.060	0.094	-0.0078	0.017	-0.073	0.071	-0.0106	0.025	-0.121	0.090	0.0178	0.023	-0.062	0.102
10.	Ile de France	0.0787	0.023	-0.005	0.169	0.0668	0.021	-0.018	0.066	0.1022	0.024	0.020	0.190	0.1358	0.020	0.067	0.207
11.	Unemp. rate	0.0055	0.095	-0.357	0.361	0.1361	0.094	-0.255	0.487	0.1902	0.096	-0.147	0.552	0.1523	0.095	-0.232	0.521
Non-French nationality:																	
12.	Own	0.0764	0.040	-0.057	0.207	0.0550	0.045	-0.088	0.214	0.0240	0.045	-0.140	0.187	-0.0752	0.034	-0.185	0.069
13.	Father	0.1064	0.083	-0.184	0.500	0.0460	0.071	-0.195	0.375	0.0856	0.069	-0.167	0.329	-0.0183	0.069	-0.344	0.237
14.	Mother	0.1250	0.079	-0.174	0.483	0.0176	0.072	-0.229	0.306	-0.0010	0.061	-0.240	0.230	0.0551	0.064	-0.217	0.282
Cohort effects (born between):																	
15.	<1929	0.0873	0.067	-0.173	0.405	0.0047	0.075	-0.256	0.281	-0.0284	0.084	-0.364	0.289	0.3720	0.068	0.118	0.635
16.	1930-1939	0.2579	0.057	0.054	0.484	0.2006	0.062	-0.051	0.432	-0.0411	0.075	-0.330	0.224	0.4487	0.060	0.205	0.686
17.	1940-1949	0.5079	0.054	0.2975	0.7440	0.3587	0.053	0.1463	0.5545	0.1396	0.059	-0.063	0.356	0.5208	0.052	0.326	0.725
18.	1950-1959	0.6393	0.053	0.4305	0.8285	0.5828	0.048	0.4076	0.7567	0.3839	0.052	0.216	0.591	0.6085	0.052	0.416	0.821
19.	1960-1969	0.4583	0.068	0.1732	0.7463	0.6027	0.051	0.4159	0.7751	0.4775	0.045	0.297	0.721	0.5722	0.061	0.332	0.819

Table 1: Wage Equation (Continued)

No.	Variable	High School Dropouts (CEP)			Vocational HS (CAP-BEP)			HS Grad. (Baccalauréat)			College Graduates						
		Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max				
Industry:																	
20.	Energy	0.1153	0.065	-0.152	0.356	0.0099	0.058	-0.187	0.225	0.1294	0.068	-0.110	0.374	0.2386	0.054	0.029	0.453
21.	Intermediate	0.1317	0.031	0.020	0.256	0.1856	0.027	0.092	0.279	0.2055	0.040	0.026	0.366	0.1364	0.035	-0.028	0.282
22.	Equipment	0.1827	0.031	0.072	0.323	0.1455	0.027	0.049	0.255	0.2490	0.037	0.113	0.394	0.2274	0.033	0.110	0.363
23.	Consumption	0.1411	0.030	0.028	0.251	0.1452	0.028	0.027	0.264	0.1865	0.038	0.015	0.335	0.1349	0.035	0.004	0.253
24.	Construction	0.0647	0.032	-0.057	0.198	0.1537	0.027	0.051	0.240	0.2773	0.044	0.124	0.434	0.1341	0.042	-0.033	0.277
25	Retail/Whole	0.0954	0.026	-0.020	0.182	0.0608	0.024	-0.039	0.151	0.1758	0.031	0.065	0.308	0.1664	0.031	0.047	0.283
26.	Transport	0.1018	0.035	-0.037	0.239	0.0697	0.032	-0.054	0.193	0.1812	0.040	0.013	0.333	0.1363	0.038	-0.017	0.273
27.	Services	0.0171	0.025	-0.086	0.103	0.0114	0.023	-0.084	0.107	0.1637	0.028	0.052	0.272	0.0777	0.027	-0.028	0.191
28.	Insurance	0.1034	0.074	-0.182	0.374	-0.0232	0.062	-0.242	0.196	0.2501	0.051	0.056	0.439	0.1614	0.056	-0.043	0.354
29.	Financial	0.2026	0.061	-0.074	0.425	0.1510	0.054	-0.045	0.364	0.2712	0.043	0.112	0.441	0.2214	0.043	0.061	0.392
30.	Oth. serv.	-0.0573	0.030	-0.184	0.061	-0.0902	0.028	-0.191	0.025	-0.1539	0.033	-0.281	-0.022	-0.3711	0.032	-0.489	-0.221
Job switch in first year:																	
31.	1st year	0.1077	0.065	-0.119	0.364	0.0154	0.065	-0.238	0.231	0.0401	0.069	-0.198	0.301	0.1441	0.074	-0.201	0.386
32.	Exp at $t - 1$	-0.0049	0.003	-0.018	0.007	-0.0083	0.005	-0.026	0.008	-0.0059	0.007	-0.031	0.019	-0.0030	0.008	-0.035	0.027
Constant of job switch that lasted:																	
33.	Up to 1 year	0.0529	0.015	-0.005	0.107	0.0301	0.010	-0.007	0.065	0.0451	0.013	-0.011	0.095	0.1829	0.015	0.120	0.238
34.	2 to 5 years	0.0985	0.030	-0.025	0.201	0.1967	0.024	0.107	0.294	0.1283	0.033	0.002	0.248	0.0529	0.030	-0.070	0.179
35.	6 to 10 years	-0.0164	0.067	-0.255	0.228	0.1635	0.059	-0.060	0.399	-0.0217	0.081	-0.304	0.270	0.1731	0.071	-0.109	0.424
36.	10+ years	0.0585	0.051	-0.141	0.249	-0.0619	0.048	-0.242	0.135	0.0851	0.072	-0.183	0.399	0.0582	0.053	-0.164	0.280
Coefficient on lagged experience for job that lasted:																	
37.	2 to 5 years	-0.0242	0.009	-0.060	0.015	-0.0316	0.008	-0.063	-0.003	-0.0344	0.012	-0.075	0.012	0.0308	0.010	-0.008	0.068
38.	6 to 10 years	-0.0025	0.009	-0.036	0.033	-0.0199	0.008	-0.054	0.014	-0.0156	0.013	-0.064	0.034	0.0079	0.010	-0.029	0.046
39.	10+ years	-0.0167	0.004	-0.034	-0.002	-0.0180	0.005	-0.037	-0.001	-0.0145	0.008	-0.051	0.014	0.0189	0.004	0.000	0.036
Coefficient on lagged seniority for job that lasted:																	
40.	Up to 1 year	-0.0055	0.001	-0.008	-0.002	-0.0029	0.001	-0.006	-0.001	-0.0059	0.001	-0.011	-0.001	-0.0095	0.001	-0.014	-0.005
41.	2 to 5 years	-0.0031	0.001	-0.007	0.001	-0.0070	0.001	-0.011	-0.003	-0.0036	0.002	-0.012	0.004	-0.0039	0.001	-0.009	0.002
42.	6 to 10 years	0.0016	0.002	-0.005	0.009	-0.0106	0.002	-0.018	-0.003	0.0003	0.004	-0.014	0.014	-0.0096	0.002	-0.018	-0.001
43.	10+ years	0.0060	0.002	-0.002	0.014	0.0094	0.002	0.001	0.019	-0.0011	0.004	-0.020	0.015	-0.0082	0.003	-0.020	0.002

Notes: Data sources are DADS-EDP from 1976 to 1996. The number of observations for the four education groups are 32,596, 12,405, 34,071, and 7,579, respectively. Estimation by Gibbs Sampling. 80,000 iterations for the first group with a burn-in equal to 65,000; 80,000 iterations and 70,000 for the second group; 60,000 iterations and 50,000 for the two last groups. The equation also includes (unreported) year indicators.

Table 2: Participation Equation

No.	Variable	High School Dropouts (CEP)			Vocational HS (CAP-BEP)			HS Grad. (Baccalauréat)			College Graduates						
		Mean	St.Dev	Max	Mean	St.Dev	Max	Mean	St.Dev	Max	Mean	St.Dev	Max				
Job characteristics:																	
1.	Constant	-1.2059	0.226	-2.046	-0.308	-1.1824	0.227	-2.150	-0.376	-0.8397	0.224	-1.685	0.024	-0.8050	0.230	-1.685	0.101
2.	Experience	0.2211	0.005	0.203	0.239	0.3273	0.005	0.309	0.348	0.4304	0.006	0.409	0.452	0.1777	0.005	0.158	0.196
3.	Exp. squared	-0.0040	0.000	-0.004	-0.004	-0.0061	0.000	-0.007	-0.006	-0.0088	0.000	-0.009	-0.008	-0.0040	0.000	-0.004	-0.004
4.	Unemp. rate	7.4407	0.286	6.337	8.561	8.6314	0.340	7.359	9.950	6.4389	0.280	5.466	7.410	6.4504	0.511	4.397	8.307
Lagged dependent variables:																	
5.	Mobility	0.2041	0.025	0.101	0.295	0.2081	0.024	0.125	0.293	0.2450	0.023	0.132	0.334	0.2665	0.026	0.176	0.350
6.	Participation	0.5192	0.009	0.485	0.556	0.3961	0.009	0.360	0.430	0.2600	0.009	0.224	0.292	0.3414	0.010	0.307	0.375
Individual and family characteristics																	
7.	Male	0.3000	0.040	0.168	0.441	0.2096	0.041	0.061	-0.073	-0.0796	0.000	-0.222	0.043	0.3530	0.045	0.199	0.518
8.	Children 0-3	-0.3319	0.031	-0.443	-0.222	-0.1685	0.027	-0.277	-0.185	-0.2517	0.031	-0.374	-0.125	0.1060	0.032	-0.031	0.217
9.	Children 4-6	-0.3387	0.029	-0.439	-0.234	-0.2828	0.027	-0.403	0.134	-0.4096	0.035	-0.534	-0.266	-0.0304	0.031	-0.153	0.084
10.	Couple	-0.0546	0.036	-0.187	0.085	-0.0060	0.041	-0.172	-0.127	-0.2256	0.042	-0.432	-0.031	0.0499	0.046	-0.138	0.206
11.	Married	-0.1070	0.025	-0.195	-0.013	-0.2155	0.027	-0.314	-0.336	-0.2101	0.030	-0.320	-0.110	0.1099	0.037	-0.015	0.228
12.	Ile de France	0.0659	0.033	-0.054	0.188	0.1177	0.037	0.000	0.251	0.0109	0.025	-0.085	0.102	6.0465	0.186	5.409	6.547
Non-French nationality:																	
13.	Own	-0.3225	0.045	-0.502	-0.157	-0.5103	0.055	-0.721	-0.336	-0.2459	0.052	-0.421	-0.040	-0.2613	0.043	-0.407	-0.121
14.	Father	0.2813	0.173	-0.276	0.915	0.2403	0.115	-0.132	0.689	0.0532	0.098	-0.323	0.381	0.1492	0.130	-0.273	0.572
15.	Mother	-0.0161	0.144	-0.588	0.472	0.0175	0.123	-0.381	0.463	-0.1245	0.081	-0.392	0.204	-0.1337	0.120	-0.560	0.249
Cohort effects:																	
16.	<1929	-2.4810	0.059	-2.750	-2.229	-3.3449	0.113	-3.758	-2.962	-3.0419	0.117	-3.464	-2.637	-2.5581	0.093	-2.918	-2.215
17.	1930-1939	-2.2575	0.068	-2.508	-2.023	-3.1446	0.084	-3.455	-2.866	-3.9734	0.111	-4.394	-3.621	-2.5230	0.086	-2.883	-2.167
18.	1940-1949	-1.9738	0.072	-2.228	-1.723	-2.9346	0.078	-3.226	-2.681	-3.7966	0.083	-4.104	-3.513	-2.3254	0.070	-2.608	-2.082
19.	1950-1959	-1.3127	0.065	-1.568	-1.041	-1.7896	0.059	-1.996	-1.579	-1.9638	0.055	-2.190	-1.785	-1.9746	0.062	-2.205	-1.762
20.	1960-1969	-0.4090	0.090	-0.730	-0.048	-0.6373	0.050	-0.812	-0.466	-0.8580	0.037	-0.995	-0.730	-1.4472	0.071	-1.719	-1.184

Notes: See Table 1 for details.

Table 3: Mobility Equation

No.	Variable	High School Dropouts (CEP)			Vocational HS (CAP-BEP)			HS Grad. (Baccalauréat)			College Graduates						
		Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max				
Job characteristics:																	
1.	Constant	-0.9936	0.319	-2.061	0.144	-1.0834	0.277	-2.107	-0.113	-0.5756	0.271	-1.557	0.449	-1.0982	0.425	-2.596	0.495
2.	Experience	-0.0119	0.012	-0.058	0.034	-0.0344	0.011	-0.083	0.003	-0.0557	0.014	-0.098	-0.010	-0.0072	0.009	-0.037	0.024
3.	Exp. squared	0.0001	0.000	-0.001	0.001	0.0006	0.000	0.000	0.001	0.0013	0.000	0.000	0.002	0.0003	0.000	-0.001	0.001
4.	Seniority	-0.0315	0.009	-0.060	0.002	-0.0315	0.008	-0.060	0.001	-0.0335	0.012	-0.078	0.021	-0.0197	0.008	-0.046	0.005
5.	Sen. squared	0.0018	0.000	0.000	0.003	0.0017	0.000	0.001	0.003	0.0017	0.001	-0.001	0.004	0.0010	0.000	0.000	0.002
6.	Unemp. rate	0.4218	0.741	-2.102	3.242	0.3087	0.684	-2.264	2.868	-0.1282	0.684	-2.794	2.409	0.0516	0.697	-2.874	2.900
7.	Part-time	0.4738	0.045	0.291	0.648	0.6054	0.039	0.443	0.755	0.4737	0.039	0.318	0.645	0.4497	0.041	0.308	0.600
Lagged dependent variables:																	
8.	Mobility	0.0150	0.051	-0.175	0.223	0.0477	0.040	-0.102	0.190	-0.0638	0.046	-0.241	0.098	-0.0161	0.043	-0.194	0.141
Individual and family characteristics																	
9.	Male	0.2164	0.058	0.010	0.448	0.1942	0.057	-0.005	0.392	0.1455	0.056	-0.033	0.309	0.0932	0.061	-0.162	0.310
10.	Children 0-3	-0.3319	0.031	-0.443	0.207	-0.0116	0.044	-0.158	0.143	-0.0512	0.059	-0.249	0.222	0.0042	0.049	-0.166	0.177
11.	Children 4-6	-0.3387	0.029	-0.439	0.155	-0.0724	0.045	-0.267	0.089	-0.1150	0.069	-0.383	0.114	-0.0832	0.047	-0.274	0.097
12.	Married	-0.1400	0.051	-0.313	0.088	-0.0989	0.070	-0.348	0.128	-0.2122	0.077	-0.492	0.053	-0.1520	0.056	-0.343	0.062
13.	Couple	-0.0055	0.076	-0.315	0.283	-0.1666	0.045	-0.321	0.006	-0.2443	0.055	-0.437	-0.057	0.0116	0.078	-0.286	0.315
14.	Ile de France	0.1744	0.069	-0.074	0.407	0.0931	0.062	-0.107	0.325	0.1122	0.061	-0.134	0.362	0.2268	0.050	0.023	0.426
Non-French Nationality:																	
15.	Own	0.1540	0.067	-0.066	0.391	0.0313	0.080	-0.212	0.311	0.0035	0.077	-0.254	0.251	-0.0358	0.058	-0.235	0.180
16.	Father	0.0417	0.220	-0.824	0.887	0.0291	0.132	-0.438	0.509	-0.0410	0.132	-0.507	0.420	0.0162	0.157	-0.554	0.581
17.	Mother	0.2582	0.208	-0.432	1.070	-0.1501	0.145	-0.683	0.335	0.1056	0.105	-0.256	0.495	0.0848	0.147	-0.421	0.706
Cohort effects:																	
18.	<1929	-1.2194	0.327	-2.518	-0.026	-0.9357	0.302	-1.997	0.099	-1.9495	0.372	-3.292	-0.631	-0.9637	0.387	-2.364	0.503
19.	1930-1939	-1.1365	0.286	-2.168	0.122	-0.7139	0.233	-1.608	0.097	-1.1395	0.264	-2.004	-0.230	-0.6508	0.373	-1.896	0.810
20.	1940-1949	-0.7240	0.247	-1.570	0.359	-0.4582	0.187	-1.263	0.208	-0.9366	0.182	-1.530	-0.349	-0.3260	0.365	-1.613	0.916
21.	1950-1959	-0.6070	0.217	-1.411	0.252	-0.2721	0.150	-1.004	0.239	-0.5250	0.123	-0.959	-0.083	0.0686	0.363	-1.149	1.304
22.	1960-1969	-0.4095	0.218	-1.199	0.393	-0.2520	0.132	-0.734	0.253	-0.2829	0.086	-0.617	0.032	0.4750	0.368	-0.688	1.730

Table 3: (Continued)

No.	Variable	High School Dropouts (CEP)			Vocational HS (CAP-BEP)			HS Grad. (Baccalauréat)			College Graduates						
		Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max				
Industry:																	
23.	Energy	-0.0565	0.737	-3.066	2.255	-1.2534	0.247	-2.250	-0.447	-1.1784	0.297	-2.186	-0.098	-0.7395	0.192	-1.453	-0.074
24.	Intermediate	-0.3374	0.319	-1.640	0.955	-0.1361	0.099	-0.491	0.201	-0.0747	0.143	-0.589	0.515	-0.3868	0.126	-0.854	0.072
25.	Equipment	0.3442	0.313	-0.764	1.538	-0.2215	0.098	-0.573	0.125	-0.0281	0.135	-0.536	0.480	-0.4903	0.123	-1.000	-0.067
26.	Consumption	0.5253	0.295	-0.517	1.738	0.0385	0.102	-0.327	0.386	-0.0535	0.136	-0.590	0.520	-0.4103	0.133	-0.989	0.100
27.	Construction	0.0887	0.297	-1.230	1.227	0.2799	0.092	-0.051	0.648	0.2324	0.147	-0.436	0.725	-0.1671	0.144	-0.664	0.374
28.	Retail/Whole	0.3397	0.274	-0.727	1.302	0.0454	0.089	-0.286	0.372	0.1300	0.121	-0.272	0.576	-0.4348	0.123	-1.071	0.011
29.	Transport	0.4488	0.380	-1.129	1.949	-0.2613	0.107	-0.655	0.112	-0.3924	0.143	-0.924	0.171	-0.7423	0.143	-1.307	-0.220
30.	Services	0.5872	0.275	-0.525	1.617	0.1569	0.086	-0.123	0.501	0.1053	0.114	-0.348	0.529	-0.4290	0.113	-0.989	-0.006
31.	Insurance	0.6985	0.576	-1.345	3.295	-0.0065	0.230	-0.825	0.830	-0.0588	0.174	-0.892	0.583	-0.3698	0.200	-1.073	0.500
32.	Financial	-0.7652	0.821	-4.565	2.069	-0.5870	0.186	-1.271	0.087	-0.2091	0.141	-0.744	0.349	-0.6057	0.147	-1.136	-0.108
33.	Oth. serv.	0.0160	0.470	-2.190	1.745	-0.3484	0.109	-0.723	0.112	-0.2473	0.125	-0.719	0.182	-0.5744	0.124	-1.170	-0.098

Notes: See Table 1 for details.

Table 4: Elements of Variance Covariance Matrices

No.	Variable	High School Dropouts (CEP)			Vocational HS (CAP-BEP)			HS Grad. (Baccalauréat)			College Graduates						
		Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max				
Individual specific effects:																	
1.	$\rho_{y_0m_0}$	0.0254	0.025	-0.047	0.077	-0.0143	0.023	-0.068	0.040	0.0397	0.024	-0.008	0.098	0.0214	0.014	-0.017	0.067
2.	ρ_{y_0y}	0.1458	0.029	0.074	0.200	0.1238	0.023	0.073	0.182	0.1199	0.033	0.050	0.195	0.1364	0.019	0.085	0.183
3.	ρ_{y_0w}	0.0596	0.036	-0.007	0.149	-0.0120	0.021	-0.057	0.036	0.0576	0.033	0.002	0.128	0.0525	0.019	0.009	0.103
4.	ρ_{y_0m}	0.0206	0.019	-0.041	0.063	-0.0261	0.016	-0.059	0.026	0.0422	0.037	-0.021	0.129	-0.0311	0.018	-0.066	0.010
5.	ρ_{m_0y}	0.0087	0.023	-0.065	0.055	-0.0025	0.034	-0.069	0.054	0.0245	0.023	-0.030	0.072	-0.0033	0.035	-0.078	0.060
6.	ρ_{m_0w}	0.0407	0.030	-0.008	0.138	0.0202	0.029	-0.043	0.088	-0.0168	0.036	-0.120	0.041	0.0045	0.011	-0.028	0.035
7.	ρ_{m_0m}	-0.0312	0.024	-0.087	0.028	0.0155	0.020	-0.035	0.060	-0.0154	0.016	-0.054	0.023	-0.0091	0.017	-0.049	0.029
8.	ρ_{yw}	0.2538	0.037	0.174	0.327	0.1260	0.021	0.077	0.174	0.2068	0.022	0.149	0.250	0.1773	0.020	0.124	0.219
9.	ρ_{ym}	-0.0677	0.019	-0.129	-0.024	-0.1013	0.017	-0.160	-0.060	-0.0643	0.020	-0.111	-0.009	-0.0287	0.020	-0.071	0.023
10.	ρ_{wm}	-0.1578	0.028	-0.219	-0.103	-0.1897	0.037	-0.244	-0.088	-0.0275	0.018	-0.081	0.012	-0.0629	0.021	-0.110	-0.009
Idiosyncratic terms:																	
11.	σ^2	0.2669	0.002	0.259	0.277	0.2653	0.002	0.257	0.274	0.3964	0.004	0.383	0.413	0.3728	0.004	0.364	0.387
12.	ρ_{wm}	-0.0595	0.013	-0.109	0.000	-0.0574	0.011	-0.103	0.000	-0.0553	0.014	-0.117	0.000	-0.0835	0.013	-0.135	0.000
13.	ρ_{ym}	0.1455	0.057	-0.030	0.359	0.1097	0.054	-0.092	0.254	0.0383	0.050	-0.119	0.201	0.2026	0.050	-0.043	0.343
14.	ρ_{yw}	0.0171	0.016	-0.044	0.068	-0.0099	0.016	-0.071	0.054	0.0161	0.020	-0.051	0.079	0.0373	0.022	-0.057	0.114

Notes: See Table 1 for details.

Table 5: Comparison of the United States and France, College Graduates

No.	Variable	College Graduates, U.S.				College Graduates, France			
		Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max
Wage equation									
Main characteristics:									
1.	Experience	0.0580	0.0032	0.0518	0.0643	0.0537	0.0035	0.0410	0.0667
2.	Exp. squared	-0.0013	0.0001	-0.0015	-0.0012	-0.0008	0.0001	-0.0010	-0.0005
3.	Seniority	0.0518	0.0029	0.0460	0.0576	0.0264	0.0027	0.0157	0.0375
4.	Sen. squared	-0.0005	0.0001	-0.0007	-0.0004	-0.0004	0.0001	-0.0007	-0.0001
Constant of job switch that lasted:									
5.	Up to 1 year	0.2240	0.0172	0.1905	0.2572	0.1829	0.0150	0.1198	0.2384
6.	2 to 5 years	0.1648	0.0189	0.1274	0.2018	0.0529	0.0300	-0.0704	0.1786
7.	6 to 10 years	0.3231	0.0683	0.1861	0.4572	0.1731	0.0708	-0.1087	0.4239
8.	10+ years	0.4717	0.0869	0.3031	0.6425	0.0582	0.0530	-0.1644	0.2796
Coefficient on lagged seniority for job that lasted:									
9.	2 to 5 years	0.0567	0.0070	0.0432	0.0709	0.0308	0.0102	-0.0080	0.0678
10.	6 to 10 years	0.0111	0.0097	-0.0079	0.0303	0.0079	0.0101	-0.0290	0.0457
11.	10+ years	0.0062	0.0055	-0.0050	0.0166	0.0189	0.0042	0.0004	0.0357
Coefficient on lagged experience for job that lasted:									
12.	Up to 1 year	-0.0071	0.0016	-0.0102	-0.0040	-0.0095	0.0012	-0.0135	-0.0045
13.	2 to 5 years	-0.0058	0.0016	-0.0090	-0.0027	-0.0039	0.0014	-0.0094	0.0015
14.	6 to 10 years	-0.0025	0.0025	-0.0073	0.0024	-0.0096	0.0024	-0.0181	-0.0013
15.	10+ years	-0.0026	0.0033	-0.0090	0.0036	-0.0082	0.0027	-0.0196	0.0018
Participation equation									
16.	Lagged mobility	0.3336	0.1646	0.0111	0.6274	0.2665	0.0255	0.1764	0.3501
17.	Lagged participation	2.0046	0.0944	1.8178	2.1978	0.3414	0.0095	0.3067	0.3748
Mobility equation									
18.	Seniority	-0.0878	0.0074	-0.1024	-0.0734	-0.0197	0.0080	-0.0462	0.0049
19.	Sen. squared	0.0020	0.0003	0.0015	0.0026	0.0010	0.0003	0.0000	0.0022
20.	Lagged mobility	-0.9019	0.0552	-1.0133	-0.7953	-0.0161	0.0432	-0.1940	0.1412
Variance-covariance elements									
Individual effects:									
21.	$\rho_{y_0m_0}$	0.8040	0.0556	0.7024	0.9005	0.0214	0.0136	-0.0165	0.0674
22.	ρ_{y_0y}	0.5716	0.0286	0.5190	0.6224	0.1364	0.0187	0.0852	0.1834
23.	ρ_{y_0w}	0.1335	0.0757	0.0169	0.2714	0.0525	0.0194	0.0087	0.1025
24.	ρ_{y_0m}	-0.6044	0.0773	-0.7595	-0.4892	-0.0311	0.0179	-0.0659	0.0101
25.	ρ_{m_0y}	0.2896	0.0429	0.2268	0.3845	-0.0033	0.0353	-0.0775	0.0595
26.	ρ_{m_0w}	-0.1450	0.0884	-0.2586	0.0403	0.0045	0.0110	-0.0278	0.0346
27.	ρ_{m_0m}	-0.4234	0.0789	-0.5691	-0.2668	-0.0091	0.0173	-0.0491	0.0292
28.	ρ_{yw}	0.2174	0.0553	0.1066	0.3017	0.1773	0.0200	0.1243	0.2185
29.	ρ_{ym}	-0.5061	0.0656	-0.6172	-0.3874	-0.0287	0.0200	-0.0707	0.0231
30.	ρ_{wm}	-0.5352	0.0590	-0.6371	-0.4131	-0.0629	0.0214	-0.1099	-0.0089
Idiosyncratic terms:									
31.	σ^2	0.2062	0.0023	0.2016	0.2104	0.3728	0.0035	0.3640	0.3869
32.	ρ_{yw}	0.0013	-0.0111	0.0075	0.0161	-0.0835	0.0134	-0.1350	0.0000
33.	ρ_{ym}	-0.0005	0.0113	-0.0217	0.0188	0.2026	0.0499	-0.0427	0.3431
34.	ρ_{wm}	-0.0496	0.0124	-0.0672	-0.0205	0.0373	0.0222	-0.0571	0.1140

Table 6: Comparison of the United States and France, High School Dropouts

No.	Variable	College Graduates, U.S.				College Graduates, France			
		Mean	St.Dev	Min	Max	Mean	St.Dev	Min	Max
Wage equation									
Main characteristics:									
1.	Experience	0.0283	0.0027	0.0229	0.0334	0.0504	0.0035	0.0372	0.0658
2.	Exp. squared	-0.0007	0.0000	-0.0007	-0.0006	-0.0005	0.0001	-0.0007	-0.0003
3.	Seniority	0.0517	0.0034	0.0455	0.0580	0.0027	0.0030	-0.0091	0.0143
4.	Sen. squared	-0.0005	0.0001	-0.0008	-0.0003	-0.0002	0.0001	-0.0006	0.0000
Constant of job switch that lasted:									
5.	Up to 1 year	0.0923	0.0144	0.0635	0.1203	0.0529	0.0146	-0.0051	0.1067
6.	2 to 5 years	0.0958	0.0219	0.0526	0.1386	0.0985	0.0298	-0.0245	0.2012
7.	6 to 10 years	0.1229	0.1027	-0.0569	0.3076	-0.0164	0.0667	-0.2545	0.2278
8.	10+ years	0.2457	0.1078	0.0474	0.4606	0.0585	0.0509	-0.1409	0.2493
Coefficient on lagged seniority for job that lasted:									
9.	2 to 5 years	0.0293	0.0084	0.0127	0.0456	-0.0242	0.0093	-0.0603	0.0147
10.	6 to 10 years	0.0213	0.0109	0.0003	0.0422	-0.0025	0.0090	-0.0357	0.0327
11.	10+ years	0.0350	0.0053	0.0238	0.0444	-0.0167	0.0043	-0.0341	-0.0015
Coefficient on lagged experience for job that lasted:									
12.	Up to 1 year	0.0009	0.0012	-0.0015	0.0033	-0.0055	0.0008	-0.0082	-0.0023
13.	2 to 5 years	-0.0007	0.0016	-0.0038	0.0024	-0.0031	0.0010	-0.0071	0.0005
14.	6 to 10 years	0.0007	0.0030	-0.0049	0.0060	0.0016	0.0018	-0.0053	0.0085
15.	10+ years	-0.0090	0.0029	-0.0150	-0.0035	0.0060	0.0021	-0.0017	0.0139
Participation equation									
16.	Lagged mobility	0.5295	0.1258	0.3043	0.7836	0.2041	0.0249	0.1008	0.2945
17.	Lagged participation	1.7349	0.0660	1.5999	1.8530	0.5192	0.0090	0.4845	0.5557
Mobility equation									
18.	Seniority	-0.0812	0.0115	-0.1007	-0.0605	-0.0315	0.0089	-0.0603	0.0018
19.	Sen. squared	0.0018	0.0003	0.0011	0.0024	0.0018	0.0004	0.0003	0.0032
20.	Lagged mobility	-0.7190	0.0738	-0.8544	-0.5807	0.0150	0.0509	-0.1753	0.2228
Variance-covariance elements									
Individual effects:									
21.	$\rho_{y_0m_0}$	-0.1020	0.1146	-0.2589	0.1067	0.0254	0.0245	-0.0472	0.0771
22.	ρ_{y_0y}	0.7548	0.0566	0.6525	0.8747	0.1458	0.0285	0.0742	0.2000
23.	ρ_{y_0w}	0.3447	0.0351	0.2732	0.4142	0.0596	0.0357	-0.0066	0.1485
24.	ρ_{y_0m}	0.0278	0.2007	-0.2908	0.2281	0.0206	0.0193	-0.0414	0.0632
25.	ρ_{m_0y}	0.1972	0.0746	0.0260	0.2971	0.0087	0.0229	-0.0653	0.0545
26.	ρ_{m_0w}	0.0646	0.0505	-0.0061	0.1794	0.0407	0.0301	-0.0079	0.1382
27.	ρ_{m_0m}	-0.0573	0.1666	-0.2619	0.2194	-0.0312	0.0236	-0.0865	0.0281
28.	ρ_{yw}	0.2958	0.0282	0.2292	0.3560	0.2538	0.0369	0.1737	0.3272
29.	ρ_{ym}	-0.2100	0.1053	-0.3832	-0.0429	-0.0677	0.0185	-0.1291	-0.0241
30.	ρ_{wm}	-0.2744	0.0799	-0.4083	-0.1348	-0.1578	0.0282	-0.2188	-0.1027
Idiosyncratic terms:									
31.	σ^2	0.2448	0.0064	0.2331	0.2539	0.2669	0.0024	0.2591	0.2765
32.	ρ_{yw}	-0.0055	0.0074	-0.0185	0.0072	-0.0595	0.0133	-0.1085	0.0000
33.	ρ_{ym}	0.0029	0.0077	-0.0117	0.0160	0.1455	0.0573	-0.0299	0.3588
34.	ρ_{wm}	-0.0346	0.0072	-0.0497	-0.0183	0.0171	0.0160	-0.0436	0.0677

Table 7: Cumulative and Marginal Returns to Experience and Seniority for the United States and France

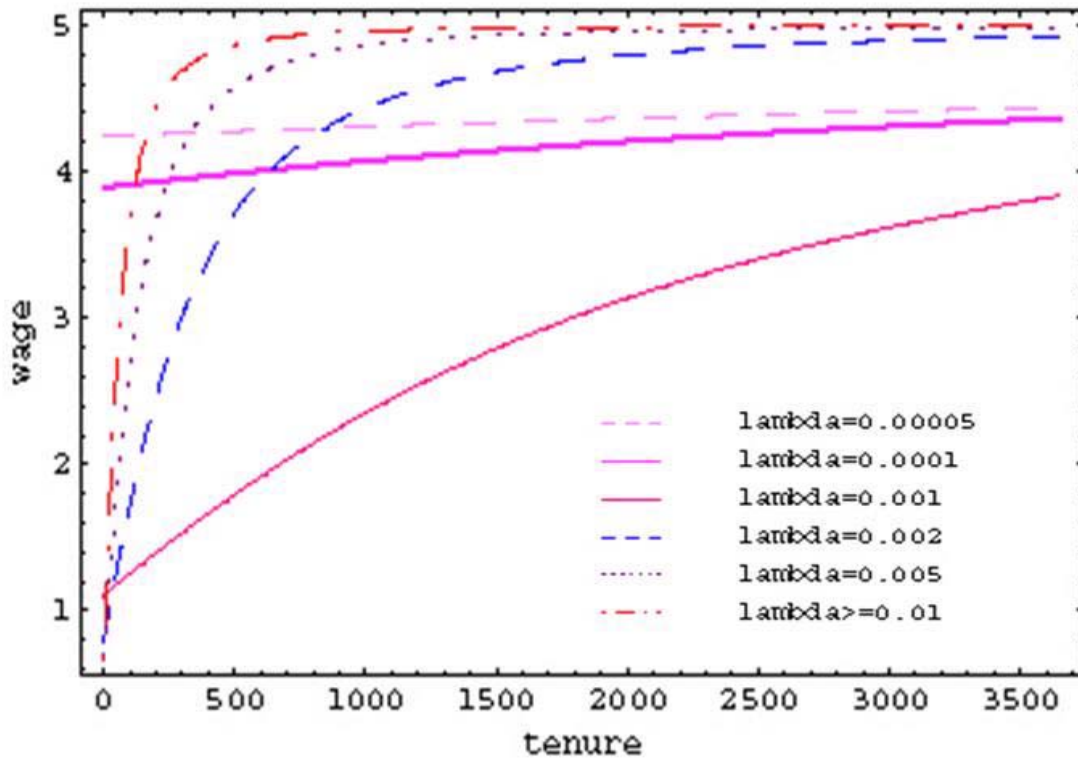
Country and Group	Cumulative			Marginal		
	Years			Years		
	5	10	15	5	10	15
Panel A: Return to Experience						
United States:						
1. High school dropouts	0.101 (0.005)	0.246 (0.012)	0.472 (0.022)	2.661 (0.286)	1.959 (0.248)	1.256 (0.215)
2. College graduates	0.256 (0.015)	0.446 (0.029)	0.567 (0.040)	4.455 (0.285)	3.109 (0.253)	1.763 (0.231)
France:						
3. High school dropouts	0.240 (0.017)	0.458 (0.031)	0.652 (0.045)	4.575 (0.314)	4.112 (0.285)	3.648 (0.265)
4. College graduates	0.249 (0.016)	0.459 (0.030)	0.630 (0.042)	4.587 (0.303)	3.808 (0.263)	3.030 (0.241)
Panel B: Return to Seniority						
United States:						
5. High school dropouts	0.266 (0.029)	0.347 (0.040)	0.392 (0.050)	4.721 (0.216)	4.314 (0.169)	3.906 (0.156)
6. College graduates	0.236 (0.014)	0.449 (0.025)	0.637 (0.034)	4.485 (0.247)	4.001 (0.209)	3.517 (0.205)
France:						
7. High school dropouts	0.006 (0.014)	-0.002 (0.027)	-0.025 (0.039)	-0.021 (0.269)	-0.307 (0.257)	-0.594 (0.267)
8. College graduates	0.122 (0.012)	0.225 (0.023)	0.307 (0.031)	2.245 (0.225)	1.849 (0.199)	1.453 (0.207)

Table 8: Comparison of Alternative Estimates of the Return to Seniority

	United States		France	
	High school dropouts	College graduates	High school dropouts	College graduates
Instrumental variable (Altonji-Williams):				
Linear tenure	0.046 (0.006)	0.037 (0.006)	-0.004 (0.003)	-0.001 (0.003)
Linear experience	0.055 (0.011)	0.058 (0.014)	0.036 (0.004)	0.051 (0.004)
Cumulative returns to tenure:				
2 years	0.062 (0.010)	0.043 (0.008)	-0.008 (0.005)	-0.002 (0.007)
5 years	0.112 (0.017)	0.078 (0.014)	-0.020 (0.014)	-0.005 (0.017)
10 years	0.131 (0.019)	0.092 (0.015)	-0.042 (0.029)	-0.013 (0.035)
15 years	0.131 (0.018)	0.090 (0.016)	-0.064 (0.046)	-0.022 (0.056)
20 years	0.146 (0.019)	0.099 (0.020)	-0.088 (0.066)	-0.034 (0.080)
Ordinary least-squares:				
Linear tenure	0.058 (0.006)	0.040 (0.005)	0.001 (0.002)	0.021 (0.002)
Linear experience	0.059 (0.010)	0.059 (0.008)	0.037 (0.003)	0.039 (0.004)
Cumulative returns to tenure:				
2 years	0.099 (0.009)	0.068 (0.007)	0.003 (0.004)	0.041 (0.005)
5 years	0.197 (0.015)	0.136 (0.012)	0.013 (0.009)	0.096 (0.012)
10 years	0.273 (0.016)	0.189 (0.013)	0.042 (0.020)	0.171 (0.026)
15 years	0.300 (0.017)	0.208 (0.015)	0.087 (0.032)	0.226 (0.041)
20 years	0.328 (0.017)	0.223 (0.018)	0.147 (0.046)	0.260 (0.059)
Our model:				
Linear tenure	0.052 (0.003)	0.052 (0.003)	0.003 (0.003)	0.026 (0.003)
Linear experience	0.028 (0.003)	0.058 (0.003)	0.050 (0.004)	0.054 (0.004)

Figure 1: Wage Profile as a Function of Tenure (in days) and σ

a. $\sigma = 0.2$



b. $\sigma = 0.4$

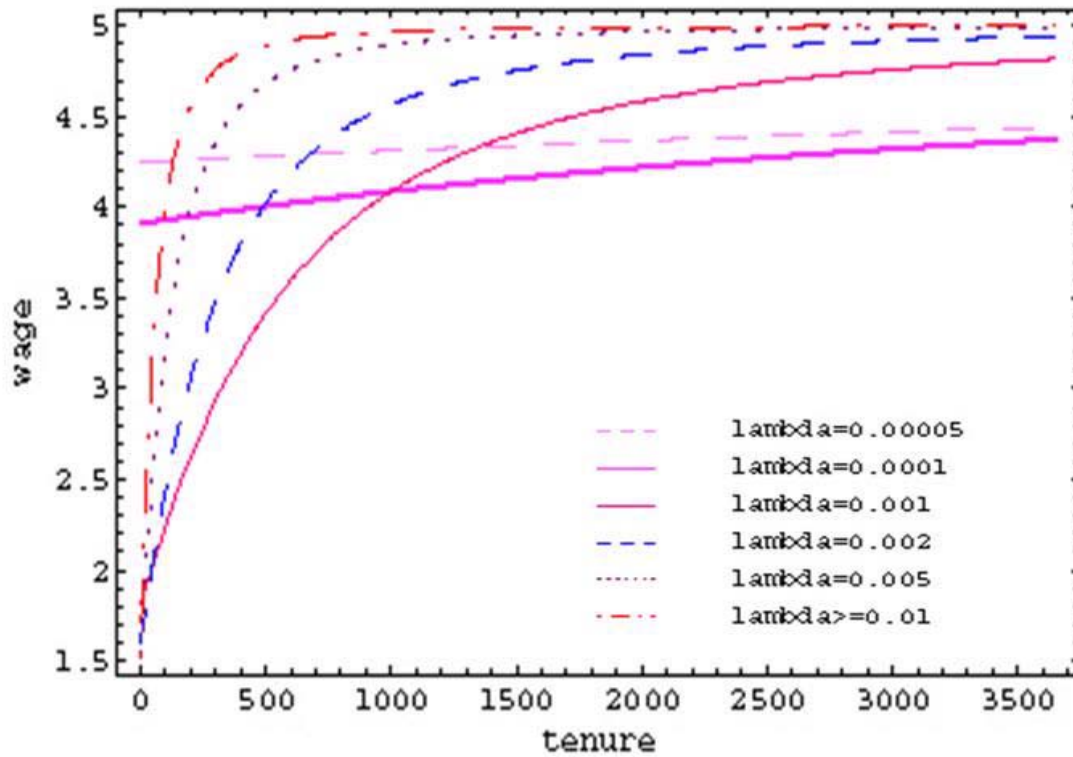
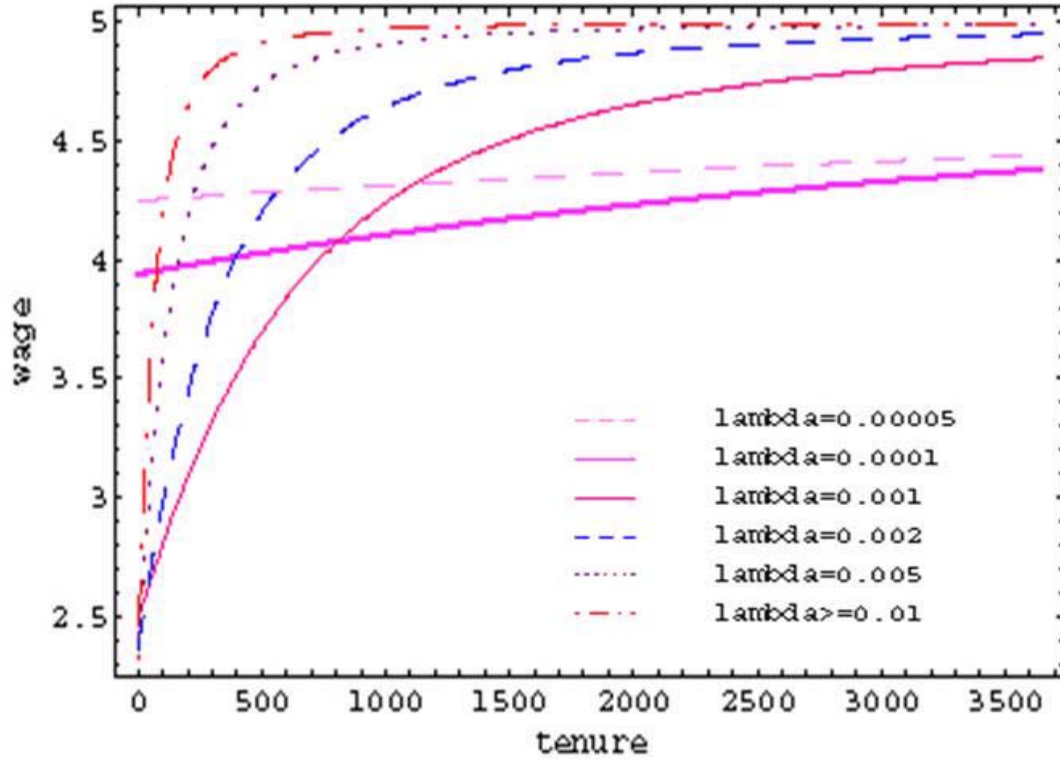


Figure 1: (Continued)

c. $\sigma = 0.8$



d. $\sigma = 1.4$

