Assessing the equalizing force of mobility using short panels: France, 1990-2000

Stéphane Bonhomme
Jean-Marc Robin

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP2/08

Assessing the Equalizing Force of Mobility Using Short Panels: France, 1990-2000

Stéphane Bonhomme
CEMFI

Jean-Marc Robin
Paris School of Economics,
Université Paris I Panthéon Sorbonne,
University College London

REVISED VERSION – August 2007

Abstract

In this paper, we document whether and how much the equalizing force of earnings mobility has changed in France in the 1990s. For this purpose, we use a representative three-year panel, the French Labour Force Survey. We develop a model of earnings dynamics that combines a flexible specification of marginal earnings distributions (to fit the large cross-sectional dimension of the data) with a tight parametric representation of the dynamics (adapted to the short time-series dimension). Log earnings are modelled as the sum of a deterministic component, an individual fixed effect, and a transitory component which is assumed first-order Markov. The transition probability of the transitory component is modelled as a one-parameter Plackett copula. We estimate this model using a sequential EM algorithm.

We exploit the estimated model to study employment/earnings inequality in France over the 1990-2002 period. We show that, in phase with business cycle fluctuations (a recession in 1993 and two peaks in 1990 and 2000), earnings mobility decreases when cross-section inequality and unemployment risk increase. We simulate individual earnings trajectories and compute present values of lifetime earnings over various horizons. Inequality presents a hump-shaped evolution over the period, with a 9% increase between 1990 and 1995 and a decrease afterwards. Accounting for unemployment yields an increase of 11%. Moreover, this increase is persistent, as it translates into a 12% increase in the variance of log present values. The ratio of inequality in present values to inequality in one-year earnings, a natural measure of immobility or of the persistence of inequality, remains remarkably constant over the business cycle.

JEL classification numbers: D30, D63, J22, J64.

Keywords: Inequality, mobility, earnings dynamics, copulas.

Acknowledgements:

1 Introduction

It is generally agreed that measuring earnings dispersion in one single year does not provide an accurate picture of earnings inequality if individual earnings change over time. Forward-looking consumers can indeed self-insure against transitory earnings shocks to some extent. Given the scarcity of panel data on consumption, longer-run earnings inequality thus provides a useful characterization, albeit imperfect, of the inequality in the intertemporal consumption sets available to individuals (Haider, 2001).[1]

Transitory earnings shocks are usually smoothed by averaging at least five years of individual earnings observations.[2] However, this approach is not readily implementable because of business cycle fluctuations. For example, Gottschalk and Moffitt's (1994) well-cited paper compares earnings inequality in the 1970s and the 1980s. If business cycles are irregular or if their periodicity is different from 10 years, contrasting the distributions of 10-year average earnings in 1970-1979 and in 1980-1989 will not necessarily characterize long-run trends. This is why the literature has recently tried to develop models of earnings dynamics incorporating structural change.[3]

In this paper, we document changes in earnings inequality, and earnings and employment mobility in France in the 1990s. We show that the first half of the period is characterized by a simultaneous increase in wage inequality, wage immobility and unemployment risk, and the second half by a mirroring decrease. As 1988-1990 and 1998-2000 correspond to periods of fast growth and 1993 to a recession, business cycle conditions appear strongly associated with individual earnings dynamics.

Because of this structural instability, we first develop a model of individual earnings dynamics with time-varying parameters. Then, we use the estimated model to simulate individual

[1]Very few studies have tried to face the scarcity of panel data on consumption and document long-run consumption inequality. An essentially exhaustive list of references includes Cuttler and Katz (1992), Attanasio and Davis (1996), Blundell and Preston (1998), and Krueger and Perri (2001).

[2]Gottschalk and Moffitt (2002) claim indeed that "a shorter lag would [give] a more distorted picture of the growth of permanent variances because short lags would be contaminated by increasing transitory variances" (p. C72). For recent applications of this technique, see e.g. Burkhauser, Holtz-Eakin and Rhody (1997), Burkhauser and Poupore, (1997), Maasoumi and Trede (2001), Aaberge *et al.,* (2002) and Van Kerm (2004).

[3]Examples of models of earnings dynamics with time-varying parameters can be found in Moffitt and Gottschalk (2002), Heathcote, Storesletten and Violante (2004), Meghir and Pistaferri (2004), and Guvenen (2005).

trajectories for each year of the 1990-decade and compute measures of inequality of Present Values (PVs) of earnings sequences of various lengths. Finally, by computing the ratio of the inequality in PVs to earnings inequality in one cross-section, we document how the equalizing force of mobility changed in the 1990s.[4]

Our model is designed to fit the French Labour Force Survey (LFS), a sequence of 11 large three-year panels from 1990-1992 to 2000-2002. Contrary to other well-used panels like the PSID, LFSs typically follow large samples of individuals over short periods of time. The model therefore combines a flexible specification of marginal earnings distributions—to fit the large cross-sections—with a tight parametric specification of the dynamics—adapted to the small time-series dimension.

Log wages are modelled as a standard linear combination of two factors (an individual fixed effect and a first-order Markov transitory shock), with time-varying factor loadings. The distribution of the individual effect is assumed discrete with many support points. The marginal density of the transitory component is a normal mixture, and in order to map the marginals onto the joint distribution of couples of consecutive realizations we use a parametric copula. We find that Plackett's (1965) one-parameter copula captures the year-to-year dynamics quite well. Transitions between employment and unemployment are modelled as a Markov chain conditional on the fixed effect.

The model being non-identified when factor loadings are allowed to vary freely over time, we assume that they vary sufficiently slowly and smoothly for each three-year panel to face a specific set of time-invariant factor loadings. For each three-year panel, the model thus takes the form of a mixture of nonlinear first-order Markov models. The nonlinearity renders identification difficult to prove. We nevertheless discuss an identification algorithm and conjecture that the semi-parametric identification of the distribution of fixed effects and transitory shocks is likely to hold in general.

We estimate the model using a sequential EM algorithm (see Arcidiacono and Jones, 2003), because EM algorithm are particularly well adapted to the estimation of discrete mixtures.

[4]Theoretical justifications for measuring (im)mobility by this ratio can be found in Shorrocks (1978) and Fields (2005).

After estimating a specific set of parameters for each subpanel, we then simulate one earnings trajectory for every individual in each sub-panel, assuming that structural parameters keep their current values forever. This amounts to assuming that individuals have myopic expectations.[5]

We find that the permanent component (fixed effect) has a distribution that is close to Gaussian. The marginal distribution of the transitory component is leptokurtic, large shocks being much more probable than for a normal distribution. Unemployment is an important source of inequality. The variance of log wages (monthly salaries) increases by 9% between 1990 and 1995, and decreases afterwards. In contrast, imputing Unemployment Insurance (UI) benefits to unemployed workers using a replacement ratio of 60%, the variance of log earnings in the whole sample increases by 11% between 1990 and 1995. Moreover, this increase is persistent as the variance of the log PVs of lifetime earnings also increases by 12% in the same period, before decreasing at the end of that period. Lastly, the equalizing force of mobility, as measured by the ratio of longer-run inequality to cross-sectional inequality, changed remarkably little over the period.

Several other authors have recently focused on the inequality of PVs of lifetime earnings, as opposed to cross-sectional earnings inequality (*e.g.* Haider, 2001, Flinn, 2002, Bowlus and Robin, 2004, and Haider and Solon, 2006). We contribute to this literature by modelling unobserved heterogeneity,[6] by modelling unemployment risk, and by documenting business cycle effects.

One distinctive feature of the model is that it specifies the full distribution of earnings, not only its first moments. Modelling the full distribution is necessary in order to compute PVs. In a contribution related to ours, Geweke and Keane (2000) build an AR(1) linear model with an individual-specific drift, and with an AR(1) transitory component.[7] They adopt a flexible specification (normal mixtures) for the distributions of transitory shocks. By comparison, our model has a more constrained dynamic structure, that is adapted to very short panels. On the other hand, Geweke and Keane estimate a stationary model using twenty years of PSID data

[5] As we simulate one trajectory per individual, we thus focus on *realized* values of earnings. On the distinction between *ex-ante* and *ex-post* measures of inequality, see Bowlus and Robin (2004).

[6] Flinn (2002) tried without success to identify and estimate the distribution of unobserved heterogeneity.

[7] Geweke and Keane (2007) extends their 2000 paper by permitting state probabilities to depend on covariates.

(1968-1989), while our framework allows parameters to change with the business cycle.

The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 precisely describes the model and outlines the estimation method. Section 4 presents parameter estimates and Section 5 shows how the model fits the data. Section 6 shows the main results of the paper, comparing cross-sectional and longer-run earnings inequality in France in the 1990s. Section 7 provides a counterfactual analysis. The last Section concludes.

2 The data

The study is based on data drawn from the 1990-2002 French Labour Force Survey (LFS), conducted by INSEE, the French statistical institute. This survey is a rotating panel, a third of the sampling units (dwellings) being replaced, every year, by an equivalent number of new units. The 1990-2002 data contain 11 subpanels indexed by the date of entry of individuals, from 1990-1992 to 2000-2002. Each subpanel has about 150,000 individuals aged 15 or more, in 75,000 households, which are thus interviewed three times, in March of three subsequent years, about various aspects of their employment histories.

We use LFS data instead of administrative earnings data, as in Alvarez, Browning and Erjnaes (2002), because the French register data (DADS) do not follow individuals moving to public jobs, self-employment or non employment. Moreover, administrative data contain limited information on individual characteristics.

2.1 Sample construction

Self-reported usual weekly hours worked are likely to be mismeasured. For instance, we notice that many individuals alternatively report 39 hours worked in one week (the legal working time in the 1990s), and 40 hours in another. To limit the influence of measurement error in hours on our results, we use monthly salaries rather than hourly wages.[8] The employment status is measured at the time of the interview as well as earnings. Monthly earnings are deflated by the retail price index. We drop all earnings observations for students, retirees and self-employed workers. We only keep male employment/earnings trajectories to limit the role of labour supply

[8]In the sequel, we shall use "wage" for monthly salaries and "earnings" for a more general definition of labour income, that is either employment monthly wage or unemployment insurance benefits (calculated as a fraction of past wage).

as a determinant of earnings dispersion. Finally, we trim wages below the first and above the ninety ninth percentiles.

Geographical mobility, among other reasons, generates a significant amount of attrition. Only 53% of the individuals of the sample have complete trajectories, 25% drop out after the second year, and 22% after the first year. We assume attrition exogenous to the employment-earnings process.

2.2 Descriptive statistics

Table 1 shows a few descriptive statistics. All quantities reported are weighted using the LFS weights.

The upper part of the table displays these statistics by calendar year, from 1990 to 2000. We see that the French working population is slightly aging between 1990 and 2002, and is also becoming more educated. The log wage variance increases from .171 in 1990 to .186 in 1996, and returns to .178 in 2000. The amplitude of the increase is 8.8%. The unemployment rate sharply increases at the beginning of the period, from 9.5% in 1991 to 14.1% in 1994, then remains roughly constant between 1994 and 1999, and drops to 11.6% in 2000.

The lower part of the table completes the description by showing statistics by date of entry in the panel: from 1990(-1992) to 2000(-2002).[9] The panel data allow to document wage dynamics. We document average wage growth and relative wage immobility, as measured by Spearman's rho,[10] over one or two years. Wage mobility clearly decreases in the first half of the decade and increases in the second half.

Overall, the first half of the 1990s is thus characterized by a simultaneous increase in wage inequality, wage immobility and unemployment risk, and the second half by a mirroring decrease.

[9] All individuals in the 1990 cross-section are treated as individuals entering the panel in 1990 because, in the estimation, we are not estimating separate parameters for the 1988 and 1989 cohorts, earnings data being measured by interval before 1990. A proper treatment of the cohorts of individuals entering the LFS panel before 1990 would have thus required a more complicated model.

[10] Let X and Y be two random variables with c.d.f. F_X and F_Y, respectively. Spearman's rho, or rank's correlation, is given by $\rho = \mathrm{Corr}(F_X(X), F_Y(Y))$.

By calendar year											
t	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Number of individuals	32749	32402	32855	34629	35689	35674	35780	34927	34945	35309	35376
Age	37.9	38.0	38.1	38.2	38.4	38.6	38.8	39.0	39.1	39.2	39.3
Age at the end of school	17.5	17.5	17.6	17.8	17.9	18.0	18.1	18.2	18.3	18.5	18.6
Wages											
Mean log wage	8.90	8.92	8.93	8.94	8.93	8.92	8.91	8.92	8.92	8.93	8.93
Log wage variance	.171	.173	.172	.179	.183	.185	.186	.186	.179	.178	.178
Unemployment probability	.098	.095	.107	.123	.141	.130	.135	.141	.134	.137	.116

By year of entry in the panel											
t	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Nb of individuals entering at t	32749	16156	16022	16955	17151	16643	17030	16599	16947	17157	17869
Nb present at t+1	16246	10294	10652	11229	11236	10899	10973	10543	10897	10492	10995
Nb present at t+2	6539	7022	7409	7795	7851	7355	7455	7255	7015	6938	7204
Age	38.7	38.1	38.2	38.4	38.6	38.8	38.9	39.2	39.2	39.4	39.2
Age at the end of school	17.4	17.6	17.7	17.9	18.0	18.1	18.2	18.3	18.4	18.6	18.6
Wages											
Mean log wage	8.92	8.93	8.92	8.94	8.92	8.92	8.92	8.92	8.92	8.94	8.94
Log wage variance	.166	.173	.177	.181	.185	.188	.184	.182	.179	.179	.173
Wage growth over 1 year	.027	.013	.008	.009	.012	.010	.021	.027	.025	.026	.026
Wage growth over 2 years	.042	.023	.016	.018	.015	.017	.040	.047	.044	.046	.044
Spearman rho over 1 year	.866	.869	.887	.902	.901	.900	.905	.902	.896	.891	.892
Spearman rho over 2 years	.851	.859	.872	.883	.885	.892	.890	.882	.886	.871	.861
Unemployment probability											
Marginal	.091	.099	.119	.126	.140	.131	.134	.142	.131	.128	.112
Transition from employment	.034	.041	.050	.045	.042	.041	.040	.037	.038	.030	.033
Transition from unemployment	.63	.65	.66	.65	.64	.67	.68	.66	.64	.61	.64

Table 1: Data description, by calendar year and by year of entry in the panel

3 A model of earnings dynamics for short panels

We consider a series of short panel datasets $\mathcal{S}_j$, $j = 1, ..., J$, with many individual observations. Each subpanel $\mathcal{S}_j$ records individual log wages y_{it} and individual attributes x_{it} over a fixed time length T (3 years for the French LFS):[11]

$$\mathcal{S}_j = \{(y_{it}, x_{it}), i = 1, ..., N, t = t_j, t_j + 1, ..., t_j + T\},$$

where t_j denotes the first observation date for panel $\mathcal{S}_j$.

3.1 Parametric specification

The wage equation is an extension of Friedman's (1957) textbook model of individual income as the sum of a permanent component and a transitory component. Specifically, we assume that

$$y_{it} = \beta_j x_{it} + \sigma_j \eta_i + \exp(\alpha_j x_{it}) \varepsilon_{it}, \quad t = t_j, ..., t_j + T, \tag{1}$$

where y_{it} is log wage and x_{it} is a column-vector of individual covariates, with parameters (row-vectors) indexed by the subpanel index j. There are two error components: η_i is an individual-specific effect, and ε_{it} is a strongly stationary idiosyncratic component (both with zero mean and unitary variance). We complete the specification of model (1) as follows.

Observed heterogeneity. The vector x_{it} of observed individual characteristics comprises experience, squared experience, and five education dummy variables: primary education (I), some high school (II), high school (III), some college (IV), and college graduate (V).

Unobserved heterogeneity/Permanent component. We model unobserved heterogeneity η_i as a strictly exogenous random effect, with zero mean and unitary variance, independent of x_{it}, for all t, and that follows a discrete distribution with K support points, $\underline{\eta}_1, ..., \underline{\eta}_K$, with respective probability $p_1, ..., p_K$. The c.d.f. of η_i is thus

$$F_\eta(x) = \sum_{k=1}^{K} p_k \mathbf{1}\left\{x \geq \underline{\eta}_k\right\},$$

where $\mathbf{1}\{\cdot\}$ denotes the indicator function, with $\sum_{k=1}^{K} p_k \underline{\eta}_k = 0$ and $\sum_{k=1}^{K} p_k \underline{\eta}_k^2 = 1$.

[11] To simplify the exposition, we omit the dependence of T on i, due to attrition.

The support points are chosen a priori to form a grid on $\mathbb{R}$. In the application, we set $\underline{\eta}_k = 2\Phi^{-1}(k/(K+1))$, for $k = 1, ..., K$. We also set K large enough, and thus the grid fine enough, to approximate a continuous distribution well. The reason why we do not discretize the distribution of η_i in the frequency domain, like a normal mixture for instance, is that modelling η_i as a discrete variable will permit to exploit a simple EM algorithm for estimation.

We allow the standard deviation of the permanent component of log wages (σ_j) to vary with the panel index j. Thus, the dispersion of unobserved heterogeneity depends on calendar time. However, the distribution of η_i is assumed to be the same in all subpanels $\mathcal{S}_j$.

Transitory component. We model ε_{it} as a strongly stationary first-order Markov process with zero marginal mean and unitary variance. The only nonstationarity in the transitory component of the log wage process that is allowed for is the dependence to experience via the standard deviation $\exp(\alpha_j x_{it})$.

The number of observation periods ($T = 3$) does not allow for a rich structure of state dependence. However, the model can be quite flexible as far as the marginal distribution of ε_{it} is concerned because of the very large sample size (more than 30,000 individuals in each subpanel; see Table 1). We thus use a flexible specification for the marginal distribution of ε_{it}, namely a mixture of $M = 3$ normal distributions. The c.d.f. of ε_{it} is

$$F_\varepsilon(x) = \sum_{m=1}^{M} \pi_{mj} \Phi\left(\frac{x - \mu_{mj}}{\omega_{mj}}\right),$$

with $\sum_{m=1}^{M} \pi_{mj}\mu_{mj} = 0$ and $\sum_{m=1}^{M} \pi_{mj}\omega_{mj}^2 = 1$.

In order to map the couples of marginal distributions of ε_{it} and $\varepsilon_{i,t+1}$ onto the joint distributions of $(\varepsilon_{it}, \varepsilon_{i,t+1})$ we use a copula. Copulas are popular in financial econometrics for modelling dependence structures.[12] Here we use a parametric copula so as to tightly parameterize earnings mobility, while keeping marginal distributions flexible. Formally, we specify the joint c.d.f. of ε_{it} and $\varepsilon_{i,t+1}$ as

$$F_{(\varepsilon_{it}, \varepsilon_{i,t+1})}(x, y) = C\left(F_\varepsilon(x), F_\varepsilon(y)\right) \tag{2}$$

[12] Copulas were initially introduced by Sklar (1959). See Nelsen (1998) and Joe (1997) for references on copulas in statistics and finance.

where C, the copula, is the joint c.d.f. of the ranks of ε_{it} and $\varepsilon_{i,t+1}$ in their marginal distributions.

The copula representation (2) exists for any strongly stationary first-order Markov process, and the copula function C is unique if F_ε is continuous (Sklar's theorem). It is thus not less general than the standard decomposition of the joint p.d.f. of $(\varepsilon_{it}, \varepsilon_{i,t+1})$ as the product of the conditional p.d.f. of $\varepsilon_{i,t+1}$ given ε_{it} and the marginal p.d.f. of ε_{it}.

Moreover, copula representations may arise naturally in structural models. In Appendix A we derive the particular form of the copula for Burdett and Mortensen's (1998) steady-state equilibrium search model. The equilibrium earnings distribution F_ε depends on the distribution of firm productivity and search-matching parameters. Yet, the copula only depends on the search-matching parameters, because individual mobility is possible in a steady-state economy only insofar as individuals exchange positions. In this example, an error in the marginal distribution thus has no effect on the parameters identified by the copula.

After testing several popular copula specifications,[13] we selected the one-parameter Plackett copula:

$$C(u, v; \tau) = \frac{1}{2}\tau^{-1}\left\{1 + \tau(u + v) - [(1 + \tau(u + v))^2 - 4\tau(\tau + 1)uv]^{1/2}\right\}. \tag{3}$$

The copula parameter τ_j is modelled as a function of observed heterogeneity:

$$\tau_j(x_{it}) \equiv \exp\left(\gamma_j x_{it}\right) - 1.$$

It unambiguously indexes mobility from more mobile (low τ) to less mobile (high τ). In particular, Spearman's rho (rank's correlation) is an increasing function of τ (see Appendix B for details).

Comparing Plackett's copula to the Gaussian copula, which is implicitly assumed in linear autoregressive models with normal margins, we found that the latter tends to underestimate the dependence in the middle of the distribution. Table 2 compares the observed transition probabilities between two consecutive years to the predicted probabilities using Gaussian and Plackett copulas. The Gaussian copula clearly underestimates the probabilities of remaining in the second, third and fourth quintiles.[14]

[13]We tried Gaussian, Plackett, Frank, Gumbel, Joe, Clayton, FGM and Log copulas.

[14]In order to compute the transition probabilities in Table 2 we first regress log-wages on experience and

Observed					Gaussian					Plackett				
.68	.21	.08	.03	.00	.66	.25	.08	.01	.00	.73	.20	.04	.02	.01
.20	.50	.22	.06	.02	.25	.37	.26	.11	.01	.20	.52	.21	.05	.02
.07	.21	.47	.20	.05	.08	.26	.32	.26	.08	.04	.21	.50	.21	.04
.03	.06	.19	.53	.19	.01	.11	.26	.37	.25	.02	.05	.21	.52	.20
.02	.02	.04	.18	.74	.00	.01	.08	.25	.66	.01	.02	.04	.20	.73

Table 2: Fit of transition probabilities across earnings quintiles by Plackett and Gaussian copulas in 1990-1991

Unemployment. We also need to model transitions from and into unemployment. This is important for any reasonable description of labour market trajectories in France where unemployment rates are chronically above or around 10%. Let e_{it} be the binary variable that is equal to 1 if the individual is employed at time t and 0 otherwise. We assume that

$$e_{i1} = \mathbf{1}\{\delta_{0j}x_{i1} + \zeta_{0j}\eta_i + \xi_{i1} > 0\}, \tag{4}$$

$$e_{it} = \mathbf{1}\{\rho_j e_{i,t-1} + \delta_{1j}x_{it} + \zeta_{1j}\eta_i + \xi_{it} > 0\}, \quad t = 2, ..., T, \tag{5}$$

where ξ_{i1} and ξ_{it}, $t = 2, ..., T$, follow independent standard normal distributions.[15] All parameters are again indexed by the subpanel index j.

Calendar-time dependence. In order to reduce the number of parameters to estimate, we smooth the dynamics of certain parameters with respect to the subpanel index j. There are many ways to do that (orthogonal polynomials, splines, etc.). First attempts at estimating the model without smoothing constraints resulted in rather imprecise estimates, yet displaying rather monotonic variations. We were thus comforted by the idea that a simple cubic polynomial in j would provide ample enough flexibility. More precisely, we let

- parameters $\beta_j, \pi_{mj}, \mu_{mj}, \omega_{mj}$ vary with j without restriction,
- parameters $\sigma_j, \zeta_{0j}, \rho_j, \zeta_{1j}$ vary with j as a cubic polynomial,
- parameters $\alpha_j, \gamma_j, \delta_{0j}, \delta_{1j}$: the intercept and the parameters of education dummies vary with j as cubic polynomials; the parameters of experience and experience squared do not

experience squared. We then compute wage ranks using the empirical c.d.f. of log-wage residuals. Finally, we estimate the copula parameters by Maximum Likelihood, and compute transition probabilities.

[15] Note that an alternative would have been to nonparametrically model the unemployment probabilities by education and experience cells. However, this approach is difficult to implement when the number of support points for η_i (K) gets large, because of the usual curse of dimensionality.

depend on j.

3.2 Identification

Let us consider the model for one single cohort of workers, neglecting the employment/unemployment process. For each individual i, we observe log wages y_{it} at three dates $t = 1, 2, 3$, such that

$$y_{it} = \eta_i + \varepsilon_{it}.$$

If η_i, ε_{i1} and ε_{i2} are independently distributed, then the three marginal distributions of η_i, ε_{i1} and ε_{i2} are nonparametrically identified from the joint distribution of $(y_{i1,}y_{i2})$ (see Kotlarski, 1967). As we have one more year, we should, intuitively, be able to allow for some dependence between the shocks.

In Appendix C, we detail the following identification algorithm. First, η_i being independent of ε_{it}, given the marginal distribution of η_i one can identify the marginal distribution of ε_{it} from that of y_{it} by nonparametric deconvolution. Then, from the joint distribution of (y_{i1}, y_{i2}) or (y_{i2}, y_{i3}) one recovers the copula density, still for a given distribution of η_i. Lastly, using the first-order Markov assumption for (ε_{it}), the information in the joint distribution of (y_{i1}, y_{i3}) or (y_{i1}, y_{i2}, y_{i3}) reveals the marginal distribution of η_i.

Unfortunately, deriving precise conditions guaranteeing the validity of this simple constructive identification procedure is difficult and clearly exceeds the scope of this paper. The identification procedure is trivially correct for the linear case of a Gaussian copula with Gaussian margins, but, otherwise, involves complicated inverse problems. This is definitely not a straightforward generalization of Kotlarski's result. Yet, we speculate that it works under fairly general conditions. In the empirical analysis we will show that, when we increase the number of support points for the distribution of η_i, the estimated discrete distribution appears to converge to a continuous limit (see Figure 1 below).

Nonstationary case. Now, consider the model:

$$y_{it} = \sigma_t \eta_i + \omega_t \varepsilon_{it}, \quad t = 1, ..., T,$$

$$\mathrm{Var}\, \eta_i = \mathrm{Var}\, \varepsilon_{it} = 1, \quad \mathrm{Cov}\,(\varepsilon_{it}, \varepsilon_{i,t-1}) = \rho_t.$$

Parameters σ_t, ω_t and ρ_t are not identifiable from second-order moments with only $T = 3$ observations per individual (6 moments for 8 parameters).[16] It is only for $T \geq 5$ that the number of moment conditions $(T(T+1)/2)$ is greater than the number of parameters $(3T-1)$. Our identifying assumption is that σ_t, ω_t and ρ_t change slowly over time. We make this statement precise by assuming that σ_t, ω_t and ρ_t can be assumed constant in each three-year panel.

4 Estimation results

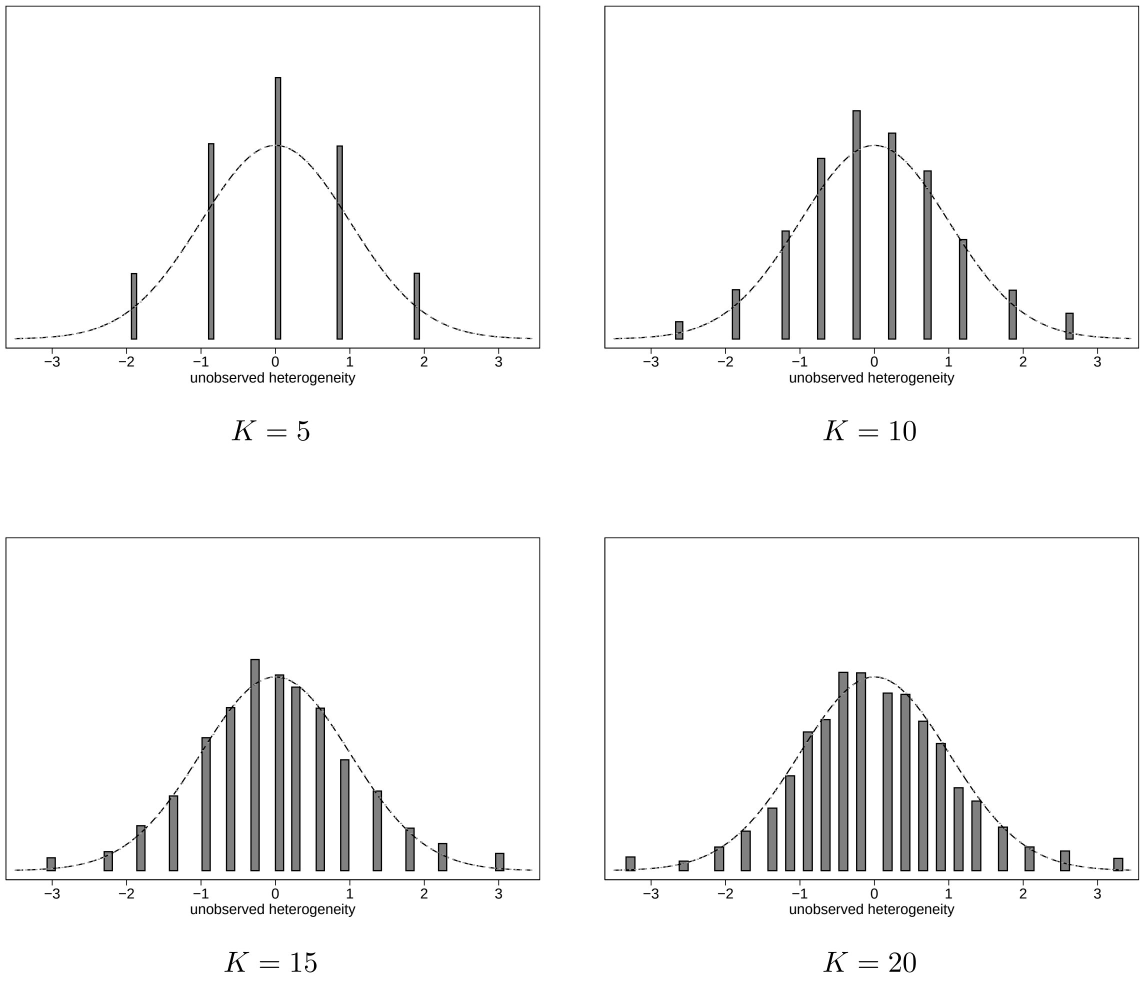
The true value of the unobserved heterogeneity component η_i being unknown but equal to $\underline{\eta}_k$ with probability p_k, the likelihood for one individual observation is a discrete mixture. We take advantage of the discrete mixture in the estimation by using a sequential variant of the EM algorithm that we detail in Appendix D. Using arguments similar to Arcidiacono and Jones (2003), this estimation procedure can be shown to deliver a consistent estimator of the model's parameters, asymptotically equivalent to a GMM estimator. This algorithm is very simple to implement but is slow to converge, as usually are EM-type algorithms. It delivers, as a by-product of the estimation, the posterior probabilities of unknown types $\underline{\eta}_k$, which can be used to simulate individual trajectories.

In the remainder of this section, we present the parameter estimates that the EM algorithm delivers for all subpanels 1990-1992 to 2000-2002.

4.1 Distribution of unobserved heterogeneity

Figure 1 shows the distribution of the individual effect η_i for different choices of support. We draw the density of the standard normal distribution for comparison. There is clear evidence of a "convergence" of the distribution of η_i to a continuous density when K increases. Moreover, this density resembles a standard normal density. This finding is consistent with what Horowitz and Markatou (1996) found for the U.S., using a model with an individual-specific effect and i.i.d. transitory shocks. It is also consistent with the idea that the unobserved heterogeneity component aggregates many missing individual characteristics.

[16] Yet they are if σ_t is constant, as the Appendix C shows. Appendix C develops an identification algorithm for the general nonstationary case where the marginal distribution of ε_{it} and the copula of $(\varepsilon_{it}, \varepsilon_{i,t+1})$ are indexed by t.



$K = 5$ $K = 10$

$K = 15$ $K = 20$

(solid line = data; dashed line = standard normal density)

Figure 1: Distribution of unobserved heterogeneity, η_i, for different values of K, the number of support points

4.2 Effects of covariates

Table 3 displays parameter estimates for subpanel 1990-1992, all other subpanels yielding similar estimates. Asymptotic standard errors are given in parentheses.[17]

The mean log wage is concave in experience, with a maximum at 31 years. The copula parameter is monotonically increasing in experience, and so relative mobility decreases with experience. The relation between unemployment risk and experience is U-shaped, with a minimum at 25 years of experience. This is true for the initial probability of being unemployed as well as for the transition probabilities into unemployment. Lastly, the log wage variance is also initially decreasing with experience, but starts to increase after only 8 years.

Education has an increasing effect on the mean log wage and the transitory log wage variance, and a decreasing effect on mobility and unemployment risk. The unobserved heterogeneity component η_i has the same effects as education. Lastly, the strong positive effect of e_{it-1} on e_{it} shows that the employment status is highly persistent, even after accounting for unobserved heterogeneity.

On Figure 2 we draw the density of ε_{it}. We observe that the distribution is highly non-normal, displaying a very peaked mode and some skewness to the left. This visual impression is confirmed by the skewness and kurtosis values, respectively equal to .26 and 7.5. The non-normality of transitory shocks ε_{it} contrasts with the near-normality of the permanent component η_i.

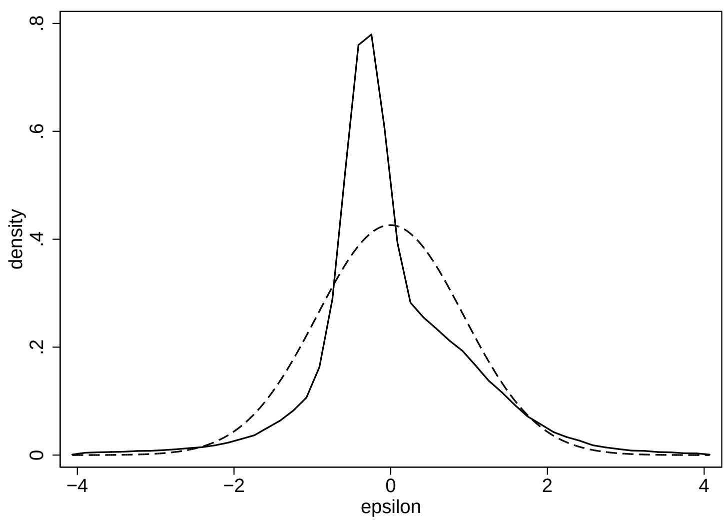
4.3 Changes across subpanels

The left panel in Figure 3 shows the evolution of the variance of the deterministic component of log wages, *i.e.* $\mathrm{Var}(\beta_j x_{it})$. It increases from .070 in 1990-1992 to .081 in 1995-1997 (a 16% increase), and returns to its initial level in 2000-2002. In Table 4 we report the results of an Oaxaca decomposition, taking 1990-1992 as reference. The contribution of age and education to the log wage variance (row labelled "composition effects" in the Table) increases in France in the 1990s because the average age and education of the labour force increase in this period (see Table 1). Composition effects account for 6 percentage points (out of 16) of the rise of

[17] We report the asymptotic standard errors of the last EM step before numerical convergence, *i.e.* taking the prior probability of each heterogeneity type as given. These are thus likely to be understated.

	Log-wages			Unemployment	
	Mean	Variance	Mobility	Initial	Transition
	(β,σ)	(α)	(γ)	(δ_0,ζ_0)	(δ,ρ,ζ_1)
expe	.03776 (.00112)	−.01172 (.0191)	.06154 (.00713)	−.1104 (.00351)	−.05234 (.00538)
(100×) expe^2	−.06113 (.00237)	.07065 (.0380)	−.04738 (.0144)	.2154 (.00751)	.1049 (.0110)
educ II	.1850 (.00808)	.1206 (.187)	.1026 (.0462)	−.5928 (.0259)	−.2576 (.0347)
educ III	.4269 (.0104)	.7416 (.194)	.7007 (.0586)	−.8137 (.0374)	−.4092 (.0518)
educ IV	.5869 (.0137)	.5230 (.264)	.6368 (.0766)	−1.282 (.0636)	−.4968 (.0770)
educ V	.8952 (.0134)	1.086 (.206)	.9898 (.0747)	−1.315 (.0633)	−.6853 (.0863)
η_i	.2304 (.00321)	-	-	−.7775 (.0137)	−.3334 (.0167)
$e_{i,t-1}$	-	-	-	-	1.744 (.0384)
constant	8.173 (.0130)	−3.486 (.266)	1.107 (.0869)	−.1165 (.0354)	−1.115 (.0612)

Table 3: Parameter estimates for subpanel 1990-1992



(solid line = data; dashed = standard normal density)

Figure 2: Density of ε_{it} for subpanel 1990-1992

	1990	1993	1995	1997	2000
Composition effects	.0700	.0704	.0741	.0734	.0742
+ returns to education	.0700	.0700	.0733	.0696	.0666
+ returns to experience	.0700	.0699	.0807	.0755	.0698

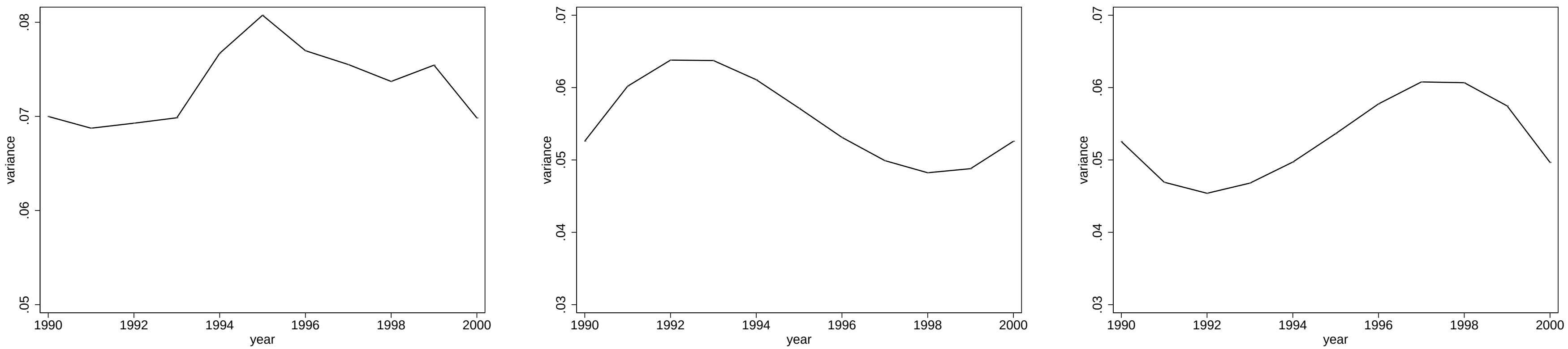
Table 4: Oaxaca decomposition of the deterministic component of the variance of log wages (reference = 1990-1992)

the log wage variance between 1990-1992 and 1995-1997. Allowing the return to education to change (in the second row), the variance decreases. This is because returns to education decrease in France in the 1990s (see Figure 4). An active educational policy is responsible of the rising high school graduation rate, from 35% in 1985 to 65% in 1995. The diminishing returns to education could thus indicate that the educational improvement does not truly reflect augmented productivity but rather lower certification levels. At the same time, returns to experience increase very steeply during the first half of the 1990s (30%), and compensate for some of the variance reduction due to lower returns to education.[18]

The central and right panels in Figure 3 display the evolution of two key parameters, that we shall refer to as the *permanent* and *transitory* components of the log wage variance. The first one is σ_j^2, and is approximately equal to the variance of the effect of unobserved heterogeneity η_i on log wages.[19] The second term is the average variance of stationary shocks, namely the mean of $\exp(2\alpha_j x_{it})$. Figure 3 shows that the deterministic, permanent and transitory components respectively account for about 40%, 30% and 30% of the log wage variance, respectively. Moreover, permanent and transitory variances follow opposite dynamics. While the permanent component increased by 20% at the beginning of the observation period, to reach a peak in 1992-1994, meanwhile, the transitory component decreased by 15%.

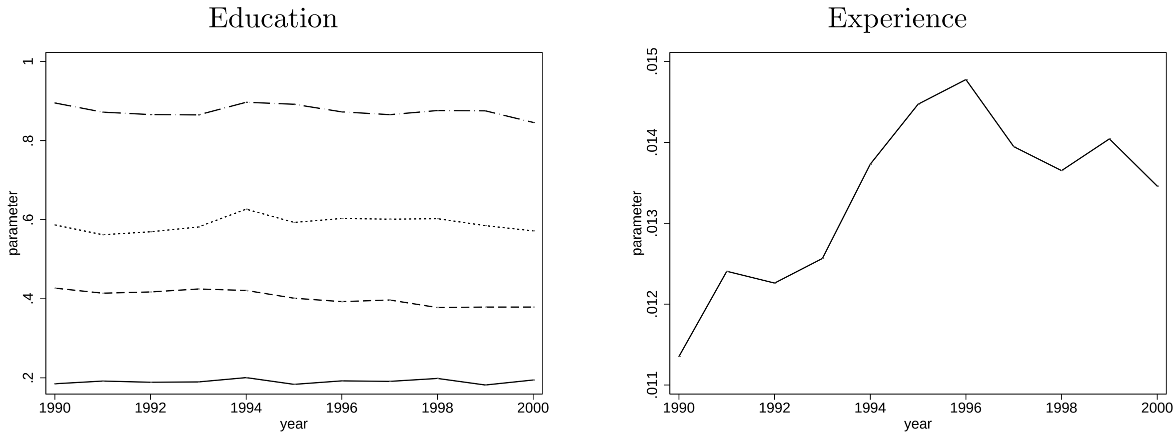
[18]The effect of experience on the mean log wage is quadratic in the model, so the return to experience is linear in experience. The mean return in the sample is reported on the right panel of Figure 4.

[19]The variance of η_i in the sample of employed workers is not exactly equal to 1 because unemployment risk also depends on x_{it} and η_i, and, therefore, the sample of employed workers, for whom we observe wages, is endogenously selected.



Deterministic $\left(\mathrm{Var}(\beta_j x_{it})\right)$ | Permanent $\left(\sigma_j^2\right)$ | Transitory $\left(\mathbb{E}\left[\exp(2\alpha_j x_{it})\right]\right)$

Figure 3: Decomposition of the cross-sectional variance of log wages, subpanels 1990-1992 to 2000-2002



(solid line = level II; dashed = III; dotted = IV;
dotted = V; Reference level is I)

Figure 4: Average returns to education and experience

5 Model fit

In this section, we evaluate the capacity of the model to reproduce several features of the data.

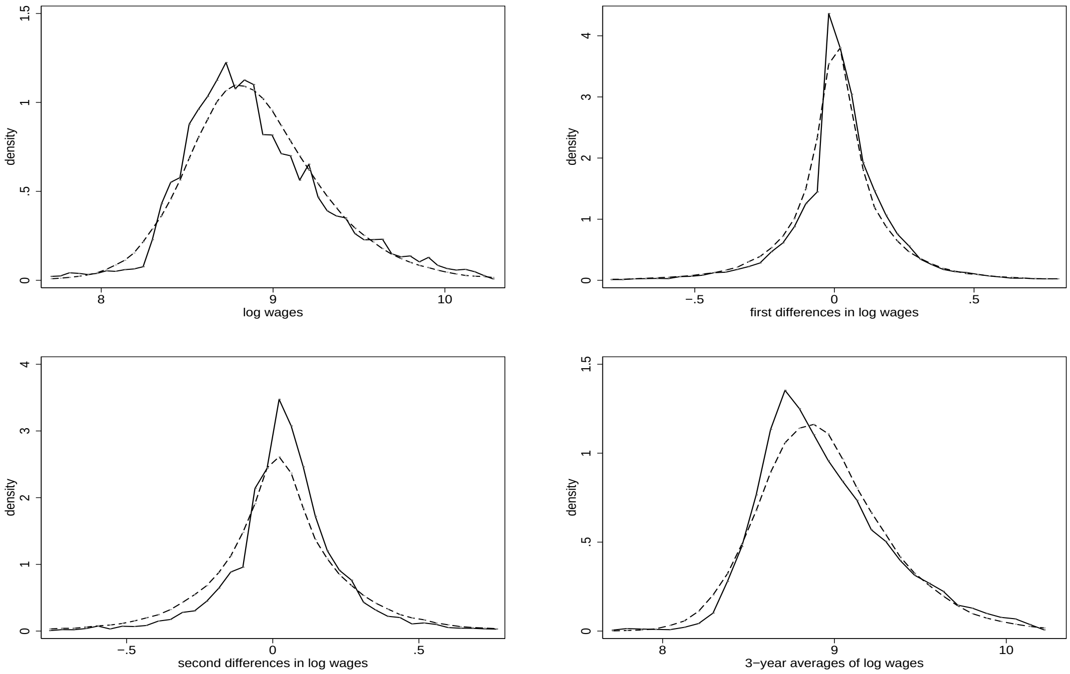
5.1 Marginal earnings distribution and earnings dynamics in 1990-1992

Figures 5 and 6 compare actual and predicted distributions of log wage levels y_{it} and log wage residuals $u_{it} = y_{it} - \beta_j x_{it}$ for 1990-1992.[20] In Figure 5, we separately consider the distributions of levels, y_{it}, first and second differences, $y_{it} - y_{it-1}$ and $y_{it} - y_{it-2}$, and 3-year averages, $(y_{it} + y_{it+1} + y_{it+2})/3$. Figure 6 repeats this comparison study for residuals. Overall, the fit is good although the model does not fully succeed in capturing the peak of the distribution of second differences. Moreover, comparing the upper left and lower right panels on the two figures shows that the additive structure of the model fails to fit precisely the skewness of log wages, though the full distribution of log wage residuals is very well fitted.

5.2 Transition probabilities in 1990-1992

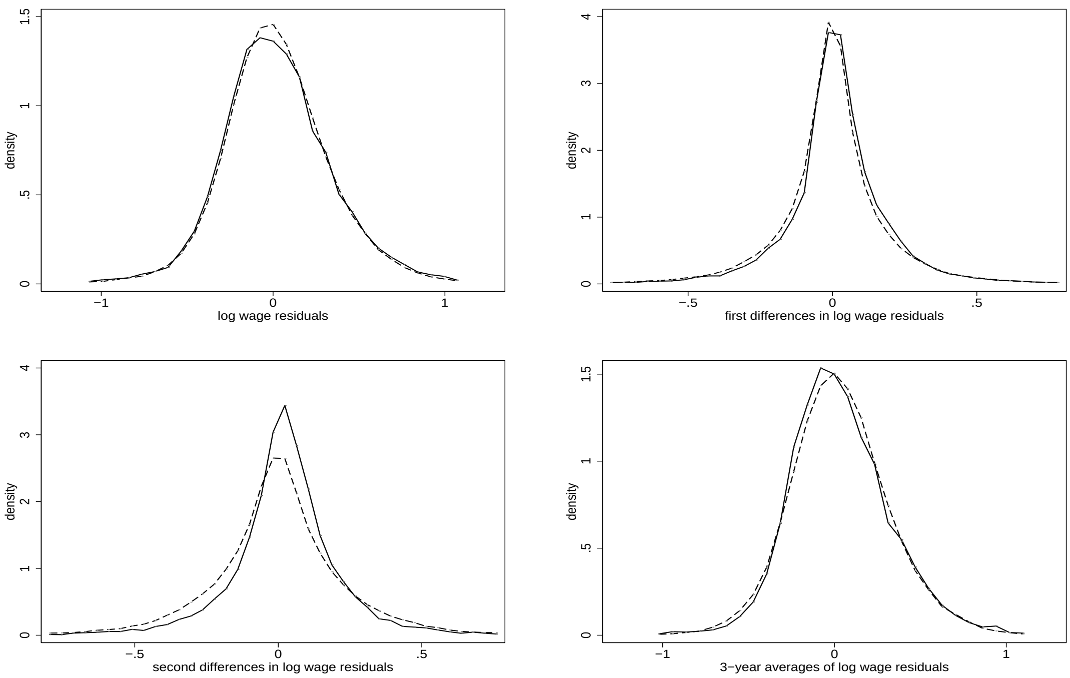
Next, we consider the transition probabilities across earnings quintiles between t and $t+1$, and between t and $t+2$. Actual and predicted probability matrices are displayed in Table 5. The fit is there also quite good, despite a certain tendency to overestimate mobility in the upper part of the distribution.

[20] We predicted these distributions by simulation, using parameters estimated on the subpanel 1990-1992.



(solid line = data; dashed = model's prediction)

Figure 5: Fit of the marginal distribution of log wage levels, subpanel 1990-1992



(solid line = data; dashed = model's prediction)

Figure 6: Fit of the marginal distribution of log wage residuals, subpanel 1990-1992

	Actual					Predicted				
	.70	.20	.07	.02	.01	.76	.18	.04	.01	.00
	.21	.54	.19	.04	.01	.19	.56	.19	.05	.01
1990/1991	.06	.22	.53	.17	.02	.04	.20	.53	.19	.03
	.02	.03	.18	.63	.14	.01	.05	.20	.57	.16
	.01	.01	.02	.14	.82	.00	.01	.03	.16	.79
	.67	.23	.08	.02	.01	.68	.22	.07	.03	.01
	.23	.49	.22	.04	.01	.22	.46	.22	.08	.02
1990/1992	.07	.23	.50	.18	.03	.07	.22	.43	.22	.06
	.02	.04	.19	.59	.16	.02	.08	.23	.47	.20
	.01	.01	.02	.16	.80	.01	.02	.05	.20	.72

Table 5: Fit of transition probabilities across earnings quintiles in 1990-1992

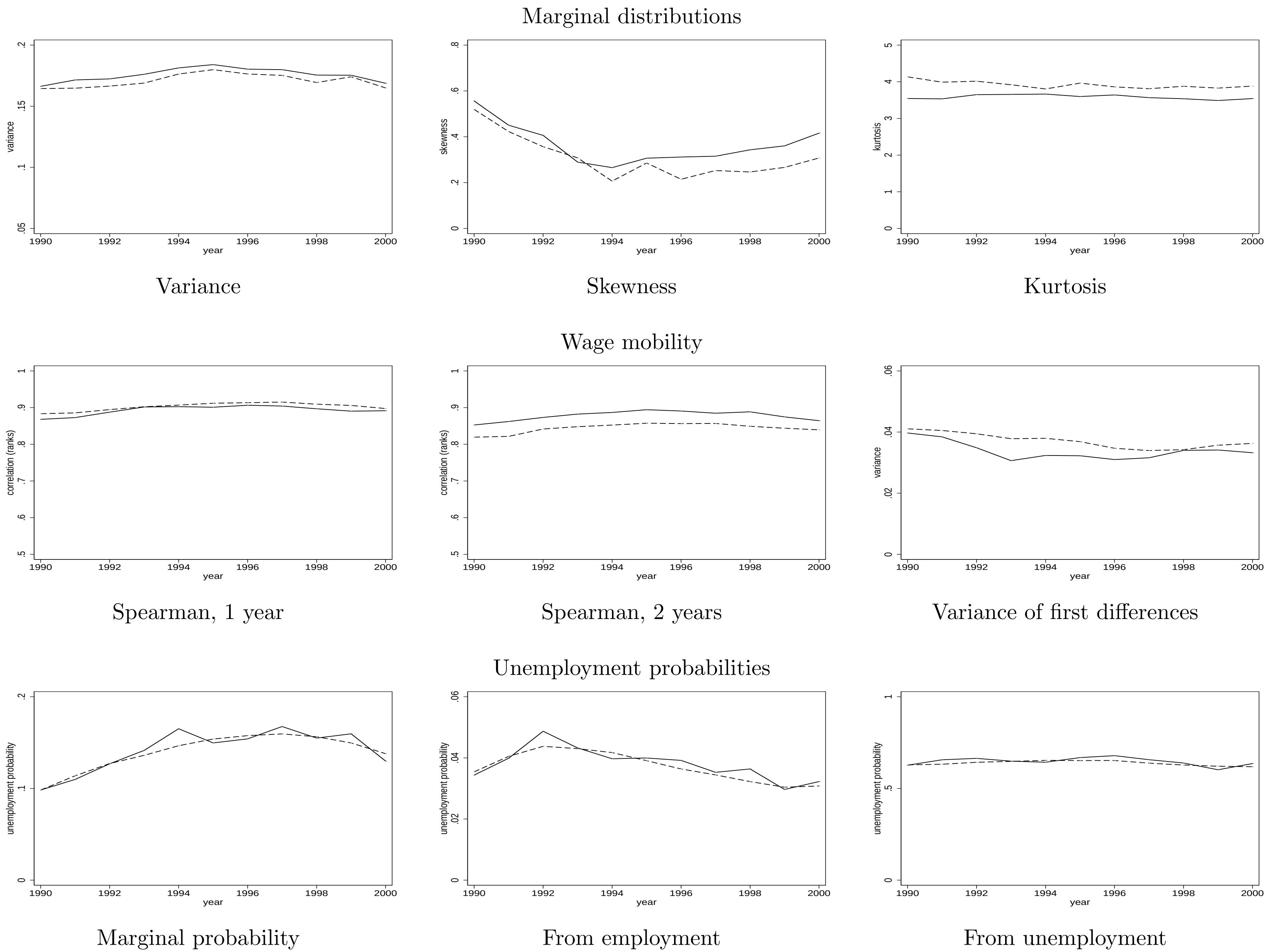
5.3 Evolution 1990-2000

Finally, we look at the capacity of the model to reproduce the changes in the log wage distribution, wage dynamics and unemployment risk across subpanels, from 1990-1992 to 2000-2002. The top panels of Figure 7 show how the model fits the variance, the skewness and the kurtosis of log wages. The middle panels compare Spearman's rho for actual and simulated log wage dynamics over one and two years, and the variance of first-differences of log earnings levels. The lower panels plot unemployment probabilities in the first year, and transition probabilities from employment and from unemployment. Again, the overall fit is quite satisfactory, although the central graph shows that the model underpredicts Spearman's rho two years apart. It is likely that the model predicts too much mobility because of the first-order-Markov assumption.

6 Longer-run inequality

With a fully parametric model in hand, we can now simulate individual earnings trajectories. Specifically, we assign to every individual in the survey the parameters of the subpanel s/he belongs to.[21] We then draw a value for the unobserved heterogeneity component η_i conditional on observations, using the posterior distribution of latent variables that the EM algorithm delivers. Next, given η_i and given the first three employment/earnings observations $Y_{it}, Y_{i,t+1}, Y_{i,t+2}$, we simulate employment/earnings trajectories after $t+2$, $Y_{i,t+3}, Y_{i,t+4}, ...$, until retirement at age

[21] So, in a given calendar year two individuals who entered the data at different dates are assigned different parameters, and thus draw their earnings from different distributions. This introduces some degree of smoothing in the dynamics of cross-sectional distributions of present values.



(solid line = data; dashed = model's prediction)

Figure 7: Model fit, subpanels 1990-1992 to 2000-2002

65. Lastly, we impute a replacement income to unemployed individuals equal to 60% of the last wage.[22] For any individual who is unemployed in the first year, we draw a "last wage" from the equilibrium (marginal) wage distribution.

For a given H, we then compute annuitized present values[23] of finite sequences of H future earnings as

$$Y_{itH}^{P}=\frac{Y_{it}+\beta Y_{i,t+1}+...+\beta^{H_{it}-1}Y_{i,t+H_{it}-1}}{1+\beta+...+\beta^{H_{it}-1}},$$

where $H_{it}=\min\{H,65-Age_{it}\}$ is the minimum of H and the remaining time before retirement, and $\beta=.95$. When the horizon H is less than the subpanel length, Y_{itH}^{P} is a sample average. Thus, our approach extends the classical way of measuring permanent income to short panels, by simulating earnings trajectories beyond the observation period. And for this purpose, we need a model that specifies the full distribution of the data.

Table 6 displays the variances of log PVs, $y_{itH}^{P}=\ln Y_{itH}^{P}$, together with other inequality indices (Gini index and 90/10 percentile ratio), for horizons $H=1,5,10,\infty$ years. In the remainder of this Section, we comment on these statistics, displaying them in graphical form to make interpretation easier.

6.1 The contribution of unemployment to earnings inequality

We start by studying the impact of unemployment risk on earnings inequality. In Figure 8, we compare the variance of y_{itH}^{P} computed in the sample of employees to the corresponding variance computed in the whole sample of both employed and unemployed workers.

In a cross-section ($H=1$) the log wage variance is one third less than the variance of log earnings. When the horizon increases, the inequality differential between both samples is reduced but never completely disappears (about 15% for lifetime earnings). The memory of the initial inequality does not vanish as the horizon increases, despite the stationarity assumed for simulating trajectories, because of time discounting and a finite retirement age.

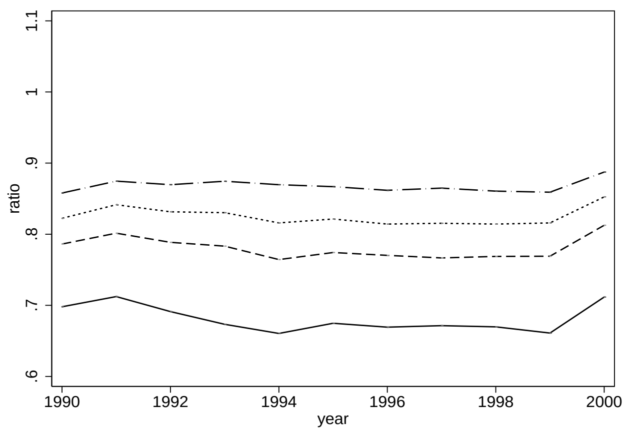
The Figure also shows that the ratio of the variance of log PVs of earnings of employees

[22] 60 % is the average replacement ratio in France over the period, see Martin (1996). Note that we neglect several important features, such as heterogeneity in the replacement ratio, nonstationary UI schemes and specific disutility of unemployment.

[23] In the rest of the paper, present values are always annuitized and we use PV as a shorthand for *Annuitized Present Value*. Computing annuities is important in order to compare the present values of earnings sequences of different lengths.

	Whole sample				Employees only			
Year	1-year	5-year	10-year	Lifetime	1-year	5-year	10-year	Lifetime
				Variance of log				
1990	.2449	.2012	.1858	.1753	.1708	.1594	.1539	.1512
1991	.2435	.2036	.1869	.1765	.1728	.1626	.1569	.1544
1992	.2494	.2064	.1902	.1791	.1718	.1622	.1576	.1552
1993	.2621	.2170	.1991	.1864	.1792	.1717	.1670	.1647
1994	.2758	.2259	.2072	.1943	.1830	.1745	.1705	.1694
1995	.2720	.2276	.2085	.1958	.1849	.1778	.1717	.1694
1996	.2729	.2258	.2054	.1913	.1856	.1758	.1703	.1677
1997	.2748	.2249	.2048	.1907	.1858	.1750	.1690	.1661
1998	.2647	.2180	.1995	.1877	.1785	.1702	.1642	.1622
1999	.2659	.2151	.1954	.1830	.1778	.1663	.1600	.1579
2000	.2511	.2060	.1886	.1784	.1778	.1655	.1596	.1575
				Gini				
1990	.2689	.2490	.2416	.2367	.2409	.2317	.2275	.2255
1991	.2675	.2499	.2421	.2374	.2410	.2332	.2290	.2273
1992	.2686	.2501	.2429	.2380	.2390	.2317	.2285	.2270
1993	.2740	.2558	.2485	.2433	.2419	.2369	.2345	.2333
1994	.2792	.2598	.2526	.2476	.2435	.2387	.2368	.2365
1995	.2774	.2603	.2522	.2470	.2440	.2400	.2365	.2352
1996	.2796	.2599	.2514	.2452	.2453	.2389	.2358	.2341
1997	.2803	.2594	.2506	.2444	.2453	.2388	.2350	.2332
1998	.2755	.2557	.2472	.2421	.2415	.2356	.2317	.2306
1999	.2757	.2538	.2448	.2389	.2410	.2335	.2292	.2274
2000	.2708	.2506	.2425	.2373	.2424	.2339	.2298	.2277
				Percentile ratio, P90/P10				
1990	3.323	2.954	2.836	2.782	2.826	2.714	2.638	2.633
1991	3.253	2.963	2.857	2.789	2.802	2.714	2.664	2.652
1992	3.459	2.994	2.893	2.843	2.799	2.708	2.684	2.687
1993	3.694	3.104	2.950	2.879	2.831	2.784	2.738	2.728
1994	3.905	3.173	3.012	2.950	2.916	2.797	2.782	2.779
1995	3.820	3.183	3.041	2.966	2.884	2.828	2.784	2.792
1996	3.833	3.203	3.022	2.940	2.834	2.787	2.749	2.770
1997	3.894	3.180	3.019	2.955	2.788	2.804	2.767	2.773
1998	3.845	3.116	2.954	2.902	2.729	2.766	2.722	2.736
1999	3.839	3.084	2.920	2.872	2.729	2.766	2.722	2.736
2000	3.482	3.022	2.880	2.858	2.855	2.745	2.689	2.710

Table 6: Longer-run inequality for various horizons, 1990-2000 (replacement ratio = 60%, discount rate = 5%)



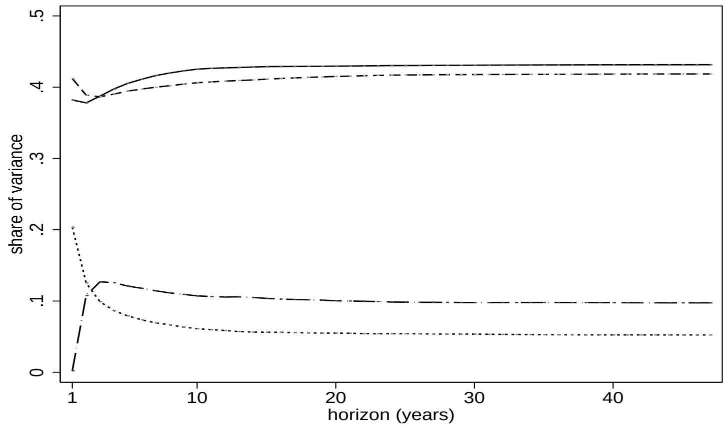
(solid line = 1 year; dashed = 5 years;
dotted = 10 years; dashed-dotted = lifetime)

Figure 8: Variance of log PVs of earnings of employees as a proportion of the variance of log PVs in the whole sample (replacement ratio = 60% and discount rate = 5%)

to that in the whole sample decreases at the beginning of the period, from .70 in 1990 to .66 in 1994. Thus, earnings inequality increases more than wage inequality following the 1993 recession, mirroring the increase in the unemployment rate. This business cycle effect vanishes in the long run.

6.2 The structure of longer-run earnings inequality

In this section we decompose the variance of log PVs of earnings (y^P_{itH}) into four components. The first one is the *deterministic* contribution of differences in education and experience across individuals. It is calculated as the between-group variance of y^P_{itH} across education and experience cells. The second one, called *permanent*, captures unobserved heterogeneity differences (η_i). We compute it as the within-group variance—by education and experience—of y^P_{itH} simulated with transitory shocks set to zero in order to cancel out any other source of persistence but η_i. The third and fourth components measure the contribution of transitory shocks. Because transitory shocks are persistent, we decompose their contribution as follows. A first transitory component (simply called *transitory*) is calculated as the within-group variance—by education, experience and unobserved heterogeneity—of y^P_{itH} simulated with i.i.d. transitory shocks, *i.e.* setting the copula parameter to zero. A *persistent* component is lastly computed as the residual variance.



(solid line = deterministic component; dashed = permanent;
dotted = transitory; dashed-dotted = persistent)

Figure 9: Decomposition of lifetime earnings inequality (variance of log PVs of lifetime earnings, with a replacement ratio of 60% and a discount rate of 5%)

Figure 9 shows the results for the year 1990. The deterministic and permanent components both explain approximately 40% of the total variance of y^P_{itH}, for all H. Transitory shocks contribute as follows. Initially, in cross-section, the transitory component accounts for 20% of the variance. When the horizon increases its share decreases rapidly to 8% after five years and 5% after 20 years. In contrast, the share of the persistent component rises to 13% after 3 years, and then decreases slightly to 10%. The total contribution of transitory shocks to lifetime earnings inequality decreases slightly with the horizon, and accounts for 15-20% of the total variance of the log PVs of lifetime earnings.

6.3 Evolution, 1990-2000

In this section we document the evolution of longer-run inequality over the 1990s. Panel a) of Figure 10 shows the evolution of the variance of y^P_{itH} in the sample of employees and the whole sample, separately. Longer-run inequality follows a similar hump-shaped pattern as cross-sectional wage inequality. However, the amplitude is significantly larger. For example, the variance of log PV of lifetime earnings ($H = \infty$) in the whole population of both employed and unemployed workers grows from .175 in 1990 to .196 in 1995—a 12% increase—and then returns to .178 in 2000. For one-year earnings, the increase between 1990 and 1995 is 11%.

Panel b) of Figure 10 displays the evolution of the variance of y^P_{itH} in proportion to the variance of log earnings y_{it1}^{P}. As more mobility reflects that individuals exchange position more frequently, this ratio is an index of immobility, and its evolution measures the changes in the equalizing force of mobility over the decade (see Fields, 2005). Two remarks are in order. First, according to this index, there is much less mobility in the sample of employees than in the whole sample. Considering employed and unemployed individuals in the initial period, the ratio of 5-year, 10-year and lifetime inequality to the cross-sectional variance of log-earnings ranges around 82%, 76% and 72%, respectively. In the sample of employees, the immobility index is always close to, or higher than, 90%. This difference between the two samples is due to unemployment being a transitory state.

Second, and importantly, the equalizing force of mobility remains remarkably constant over the decade. Indeed, in the whole sample the ratio of lifetime to current earnings inequality varies between 72.5% (in 1991) and 68.8% (in 1999). This near constancy is all the more noticeable that all inequality indices follow a marked hump-shaped pattern.[24]

7 Counterfactual analyses

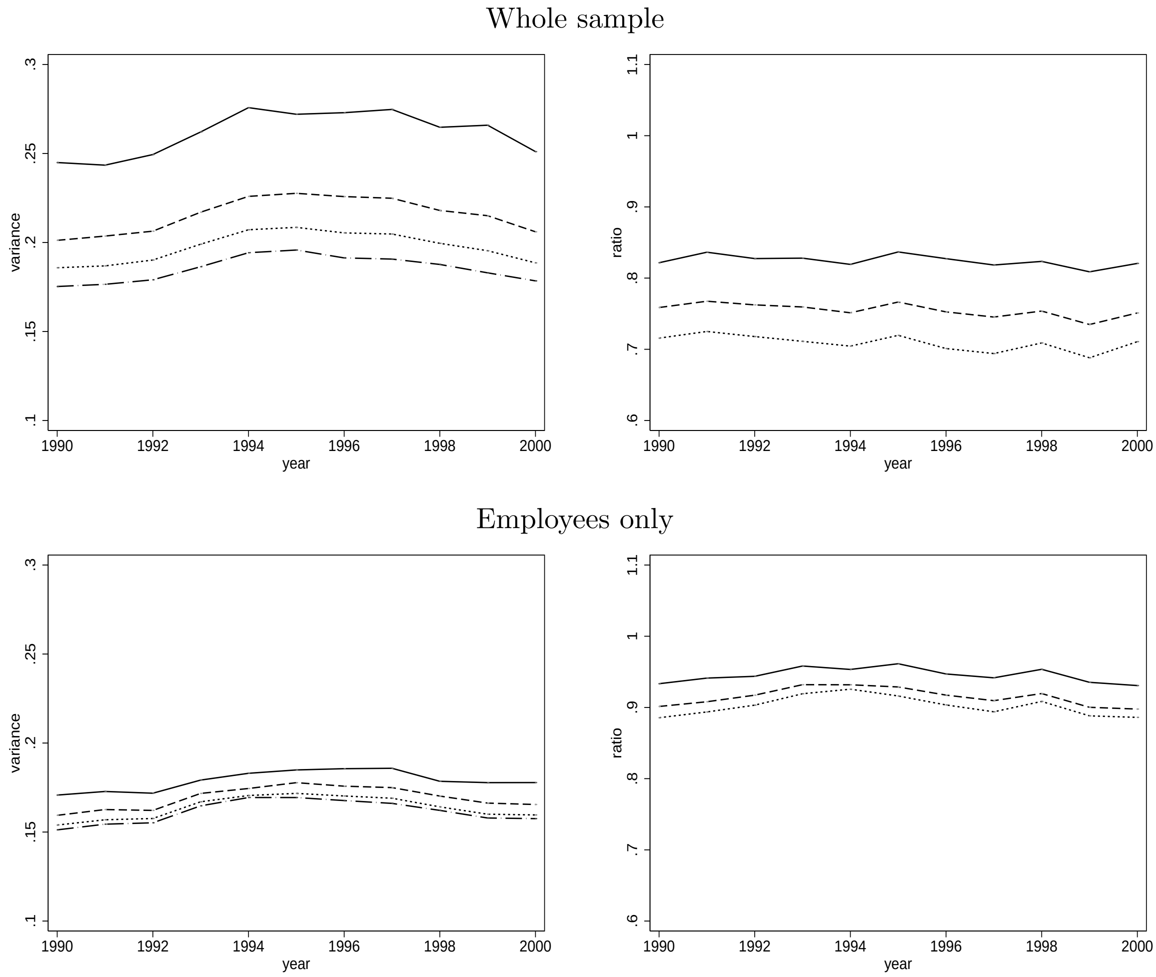
In this section, we test the sensitivity of the previous results to unemployment risk, we evaluate the effect of neglecting unobserved heterogeneity on longer-run inequality, and we measure the consequences of non Gaussianity and nonlinearity of the process of transitory shocks.

7.1 Varying unemployment risk

First, we study the sensitivity to unemployment insurance by varying the replacement ratio of Unemployment Insurance (UI) benefits between 40% and 80%.[25] Given the UI legislation in France, these percentages can be seen as conservative lower and upper bounds for unemployment compensation. Figure 11 shows that lowering the replacement ratio from 80% to 40% has very significant effects on longer-run earnings inequality. Inequality rises by about 25% (lifetime)

[24] The Gini index gives a similar picture, the ratio of the Gini of the PVs of lifetime earnings to the Gini of one-year earnings remaining constant over the decade, at 88%. In contrast, using the 90/10 percentile ratio yields a decrease in the ratio, from 86% in 1991 to 76% in 1994, followed by an increase to 82% in 2000 (see Table 6). This difference is due to the sensitivity of the 90/10 ratio to the lower tail of the earnings distribution, hence to the unemployment rate.

[25] In addition to the replacement ratio, we checked the robustness of the results to variation in the discount rate (from 1% to 10%). We also experimented with the way older individuals are accounted for in the calculation of inequality. Both modifications had little effect on the results.



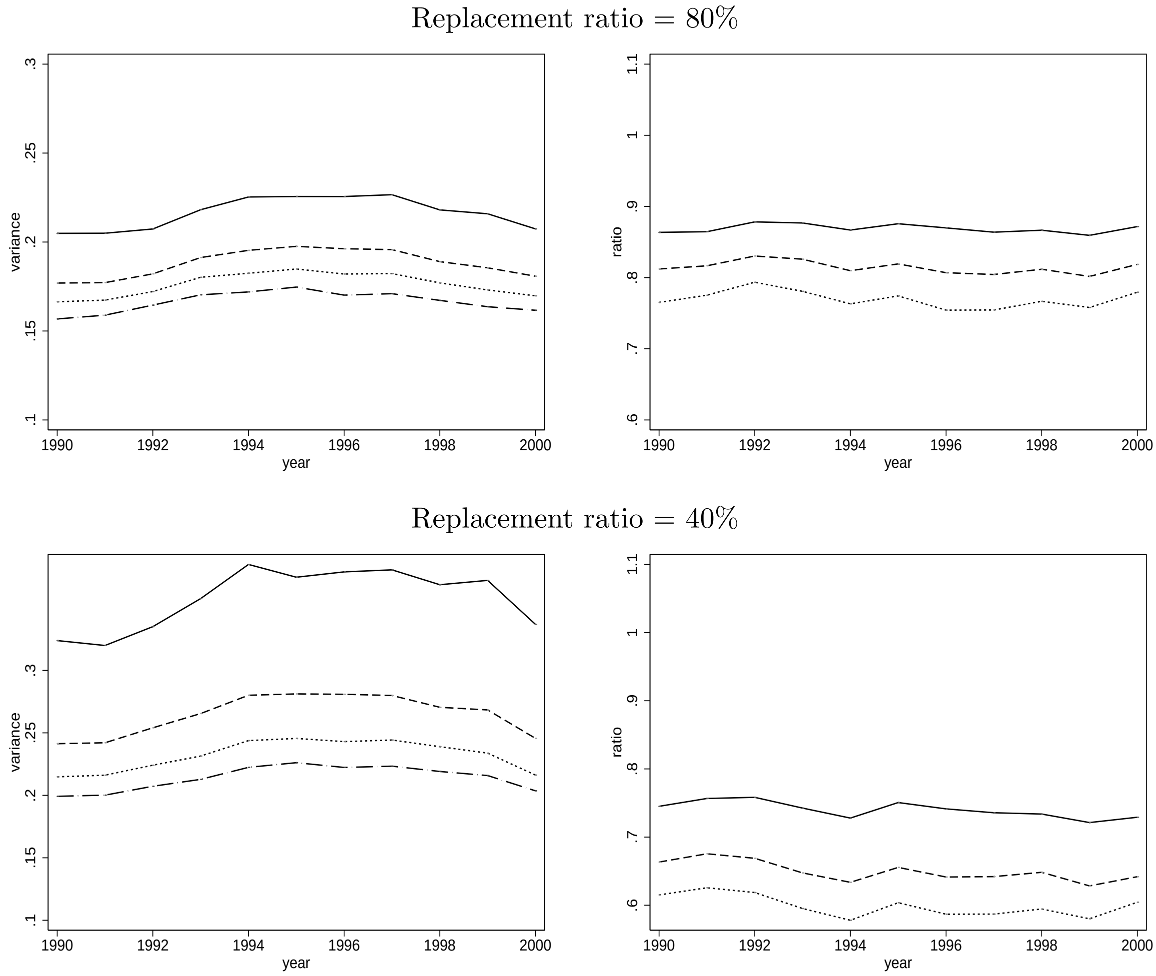
(solid line = 1 year; dashed = 5 years;
dotted = 10 years; dashed-dotted = lifetime)

a) Variance of log PVs

(solid line = 5 years; dashed = 10 years;
dotted = lifetime)

b) Immobility index

Figure 10: Longer-run earnings inequality (replacement ratio = 60% and discount rate = 5%)



(solid line = 1 year; dashed = 5 years; dotted = 10 years; dashed-dotted = lifetime)

(solid line = 5 years; dashed = 10 years; dotted = lifetime)

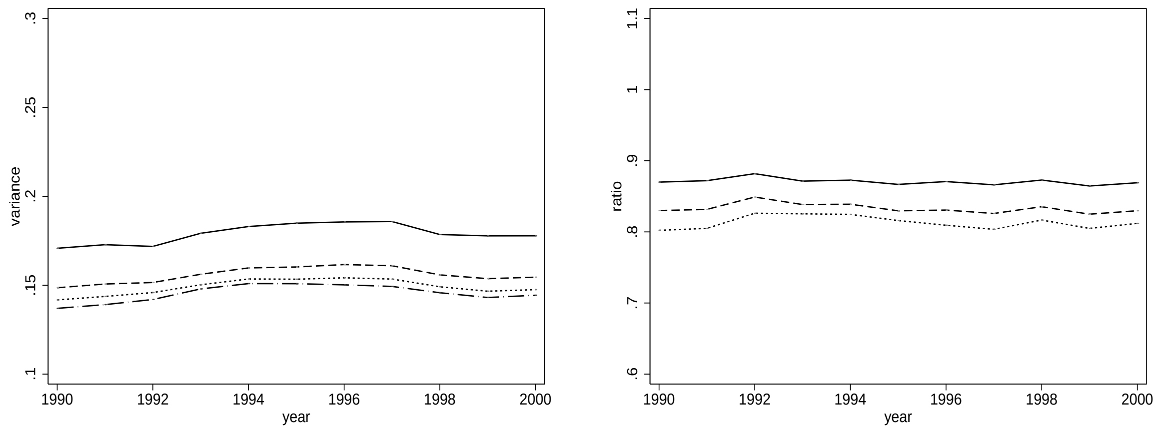
a) Variance of log PVs

b) Immobility index

Figure 11: Counterfactual simulation of longer-run earnings inequality for replacement ratios of 40% and 80% (whole sample)

or 60% (one year). The depressing effect of the 1993-1995 recession is also more pronounced when the replacement ratio is lower, as unemployment risk increases in 1993. Finally, the equalizing force of mobility is stronger, the immobility index decreasing from 75% to 60% when the replacement ratio decreases from 80% to 40%. This is because employment spells last on average much longer than unemployment spells, making unemployment shocks to income intense but transitory.

Second, we study the consequence of suppressing unemployment risk altogether. The results of the counterfactual simulation are displayed in Figure 12. Lifetime earnings inequality



(solid line = 1 year; dashed = 5 years; dotted = 10 years; dashed-dotted = lifetime)

a) Variance of log PVs

(solid line = 5 years; dashed = 10 years; dotted = lifetime)

b) Immobility index

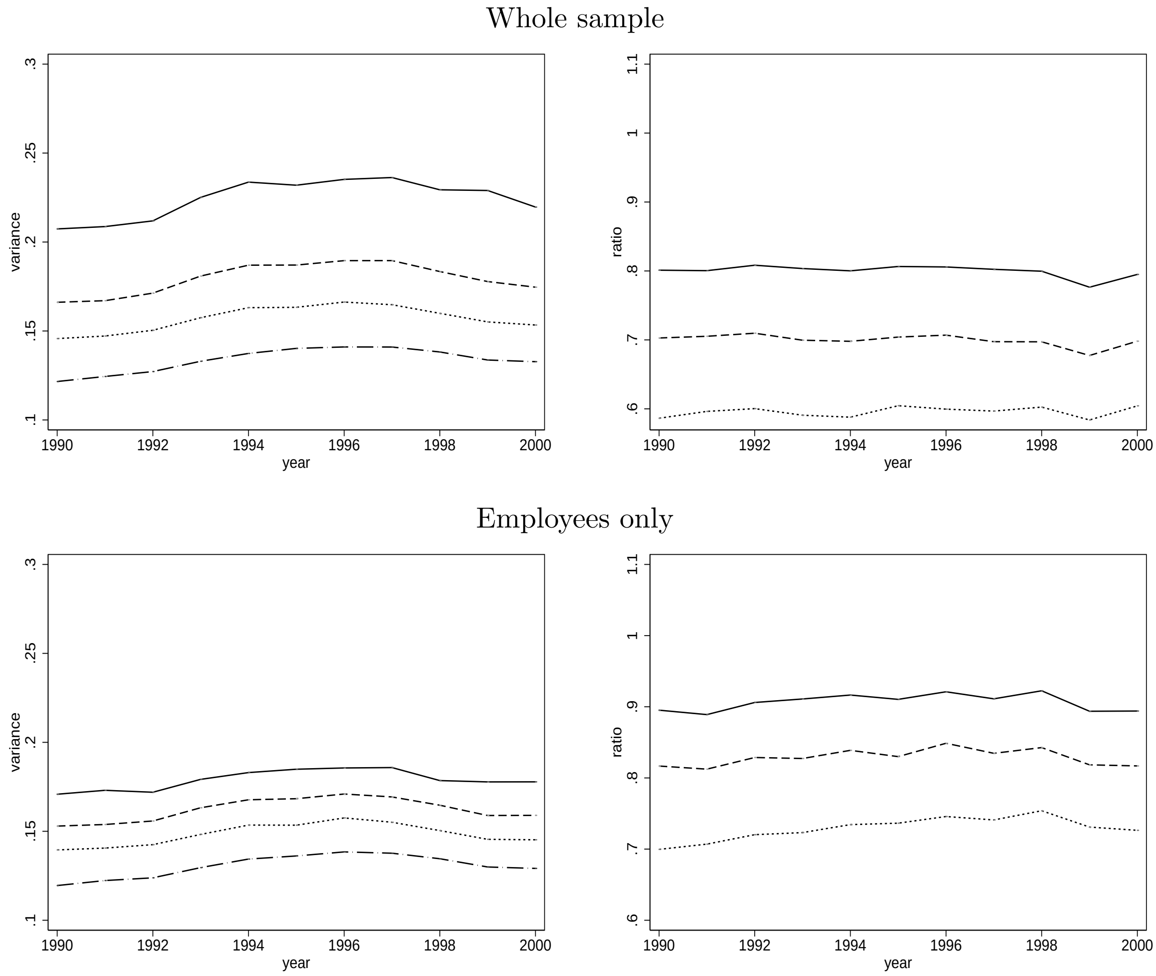
Figure 12: Counterfactual simulation of longer-run earnings inequality with no unemployment risk (whole sample, replacement ratio = 60% and discount rate = 5%)

increases by 11% from 1990 to 1995, a close match to the 12% increase generated by the actual parameters. So, cancelling unemployment risks has small effects on the evolution of the equalizing force of mobility. However, we see that the level of inequality is now noticeably smaller (*e.g.* .155 versus .175 in 1990), and that lifetime earnings inequality is more persistent (index of immobility of about 80% instead of 72%).

7.2 Neglecting unobserved heterogeneity

Most studies of lifetime earnings inequality are based on models of earnings dynamics that neglect unobserved heterogeneity (Bowlus and Robin, 2004, Flinn, 2002). To measure the consequences of this omission, we now estimate and simulate a model where the unobserved heterogeneity term η_i is set to zero. In the model with unobserved heterogeneity, unemployment is associated to lower values of η_i, and lower η_i to lower wages. So, omitting unobserved heterogeneity should yield higher UI benefits for unemployed workers.

Figure 13 shows the simulation of longer run inequality. The variance of log earnings in cross-section is 15% lower than the corresponding figures plotted in Figure 10, while the variance of log PVs of lifetime earnings is 30% lower. Moreover, the index of immobility is now lower than before and is strongly countercyclical. This result contrasts with the near-constancy of



(solid line = 1 year; dashed = 5 years; dotted = 10 years; dashed-dotted = lifetime)

(solid line = 5 years; dashed = 10 years; dotted = lifetime)

a) Variance of log PVs

b) Immobility index

Figure 13: Counterfactual longer-run earnings inequality according to the model with no unobserved heterogeneity (replacement ratio = 60% and discount rate = 5%)

the equalizing force of mobility that the model with unobserved heterogeneity predicts.

So, neglecting unobserved heterogeneity in earnings leads to strongly understate both the level and the persistence of inequality, and fails at capturing the evolution of mobility.

7.3 Gaussian transitory shocks and copula

Most earnings models assume linear dynamics and Gaussian innovations. In our model, the transitory process (ε_{it}) is AR(1) if margins and copula are all Gaussian. We already established that the distribution of the permanent component η_i was approximately normal, but that

	Log-wages			Unemployment	
	Mean	Variance	Mobility	Initial	Transition
	(β,σ)	(α)	(γ)	(δ_0,ζ_0)	(δ,ρ,ζ_1)
expe	.03369 (.00112)	.003746 (.0336)	.02598 (.00186)	−.07973 (.00302)	−.03773 (.00518)
(100×) expe^2	−.05071 (.00237)	.04423 (.0671)	−.03307 (.00351)	.1566 (.00649)	.07755 (.0106)
educ II	.1607 (.00807)	.2389 (.391)	.1633 (.0165)	−.4507 (.0226)	−.1889 (.0336)
educ III	.4378 (.0158)	1.182 (.375)	.4283 (.0163)	−.5630 (.0320)	−.3057 (.0497)
educ IV	.5624 (.0137)	.5444 (.559)	.2997 (.0221)	−.9377 (.0534)	−.4021 (.0748)
educ V	.8976 (.0134)	1.447 (.394)	.4545 (.0172)	−.9491 (.0521)	−.5345 (.0816)
η_i	.2965 (.00342)	-	-	−.1025 (.0106)	−.07642 (.0157)
$e_{i,t-1}$	-	-	-	-	2.033 (.0357)
constant	8.215 (.0130)	−4.453 (.507)	−.4010 (.0258)	−.1961 (.0315)	−1.283 (.0594)

Table 7: Parameter estimates for subpanel 1990-1992, model with Gaussian margins and Gaussian copula.

neither the margins of (ε_{it}), nor the copula, were Gaussian. We now test the effect on longer-run inequality of estimating and simulating a fully Gaussian model.[26]

Table 7 shows the parameter estimates. Comparing with Table 3, we see that the Gaussian model yields a very different decomposition of log wages into deterministic, permanent and transitory components. Indeed, the variance of the permanent component is $\sigma_j^2 = .2965^2 = .088$, compared to $.2304^2 = .053$, that is two thirds more. In the Gaussian model, the unobserved heterogeneity component thus accounts for 50% of the variance of log wages, compared to 30% in the non Gaussian model. As a consequence, transitory shocks have a smaller variance and are much less persistent. Moreover, the Gaussian model predicts *more* state dependence in employment (2.03 versus 1.74 for the autoregressive coefficient in unemployment), but a *lower* dependence to unobserved heterogeneity (compare the seventh row in the two Tables).

Hence, as expected, the choice of the parametric copula used to describe dynamic tail dependence does have important consequences on the estimation of the factor loadings.

Figure 14 shows the effect on longer-run inequality. On panel a), we see that the Gaussian AR(1) model predicts about 12% less earnings inequality in cross-section. This is due to the fact

[26]With Gaussian margins, the parameter of the Gaussian copula is also the autoregressive parameter of the AR(1) model. We model it as $\tau_j(x_{it}) = \exp\left(\gamma_j x_{it}\right) - 1$.

that unobserved heterogeneity is less strongly associated with unemployment than in the non Gaussian model. The difference in terms of lifetime inequality is smaller, as the Gaussian model predicts about 7% less lifetime inequality in the whole sample, and 2% more in the sample of employees. The Gaussian model predicts an immobility index that is about 5 percentage points higher in the whole sample. For employees, the ratio is close to 90% in both simulations. The evolutions of inequality and immobility indices are also nearly identical. It is quite remarkable that two models with so different decompositions of the log wage variance yield so similar lifetime inequality patterns. Hence, on the three-year data we use in this study, large differences in the classical distinction between unobserved heterogeneity (permanence) and state dependence (persistence) translate into rather similar pictures in terms of longer-run inequality.

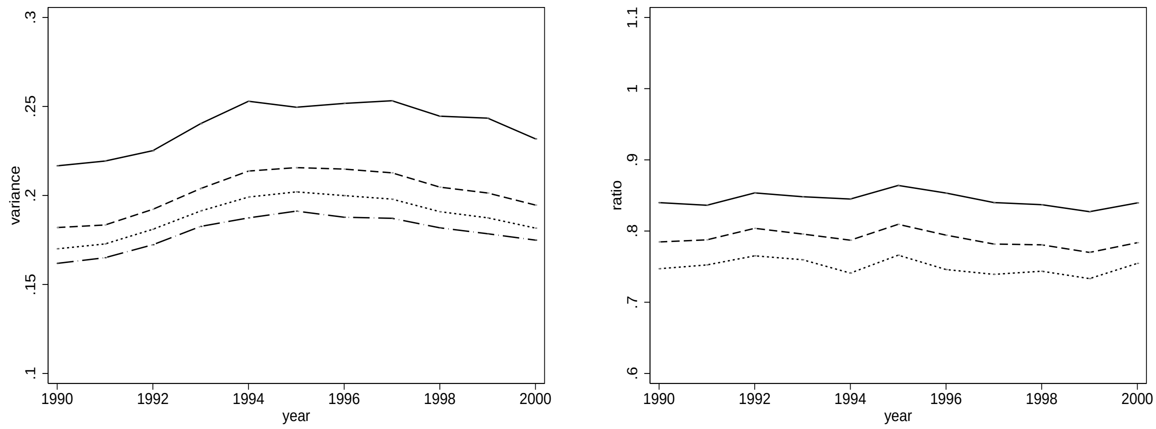
8 Conclusion

In this paper, we study the evolution of cross-sectional and longer-run measures of inequality in France in the 1990s, a period characterized by significant business cycle fluctuations. For this purpose, we use the French Labor Force Survey. In order to fit this type of data, we construct a model that combines a flexible specification of the marginal earnings distributions (adapted to the large cross-sectional dimension) with a tight parameterization of the dependence structure (that the small number of recording periods requires).

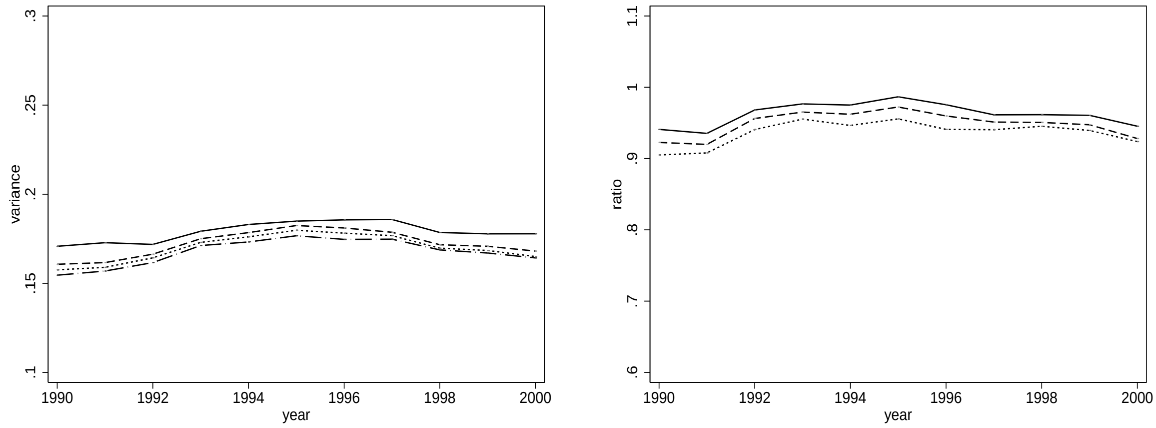
The model's structure is standard, with an individual effect and stationary transitory shocks, and we allow parameters to change with the date of entry of individuals into the panel. We simulate individual trajectories beyond the observation period by drawing from the conditional distribution given observations. The ratio of the inequality of long-run averages of earnings to the inequality in a cross-section is an aggregate measure of the persistence of inequality. Thus, our approach adapts an empirical methodology that has been previously developed for long panels to short panels like labour force survey data.

The representation of log earnings as the convolution of several factors is standard in the literature. The necessity of allowing for structural change (i.e. time-varying factor loadings) is now recognized of utter importance. A few recent papers also point at nonlinearities in the process of transitory shocks. Our model has all these characteristics while still being simple

Whole sample



Employees only



(solid line = 1 year; dashed = 5 years; dotted = 10 years; dashed-dotted = lifetime)

(solid line = 5 years; dashed = 10 years; dotted = lifetime)

a) Variance of log PVs

b) Immobility index

Figure 14: Counterfactual longer-run earnings inequality under full Gaussianity (replacement ratio = 60% and discount rate = 5%)

enough to authorize a rigorous discussion of identification. Our identification discussion is yet mostly at a conjectural stage. Further work is needed to formulate precise identification statements.

We find that the distribution of the unobserved heterogeneity component is approximately Gaussian and the distribution of transitory shocks is leptokurtic. Relative mobility, as measured by the joint distribution of the ranks of two consecutive transitory shocks (*i.e.* the copula), is also non Gaussian, and best approximated by Plackett's copula. Unemployment risk and unemployment insurance are found to be an important source of permanent inequality. Despite significant business cycle fluctuations in wage inequality, wage mobility and unemployment risk in the 1990s, the measure of the persistence of inequality (the immobility index) is remarkably stable. Counterfactual analyses also confirm the importance of allowing for unobserved heterogeneity.

Moreover, the choice of parametric specification for the copula has important consequences on the estimation of the factor loadings. A Gaussian copula attributes to unobserved heterogeneity a greater share of the total log wage variance. As, at the same time, the unemployment process depends less on unobserved heterogeneity, the Gaussian-copula model yields a very similar picture in terms of longer-run inequality. If our conjecture on identification is correct, the copula is identified. In the future, it will be interesting to push the analysis further and estimate a model with a flexible distribution for both the marginal distributions of fixed effects and transitory shocks, and the copula.

It will be also interesting to test the model on longer panels. The basic structure of the model could be extended for that purpose. With four or five observation periods per individual, it should be possible to identify and estimate a model with second-order Markov transitory dynamics, or with an additional measurement error. It would also be interesting to allow for an additional individual factor governing the unemployment process specifically. The estimation algorithm performs well and could easily be modified to incorporate these extensions.

References

[1] AABERGE, R., A. BJØRKLUND, M. JÄNTTI, M. PALME, P.J. PEDERSEN, N. SMITH and T. WENNEMO (2002): "Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States," *Review of Income and Wealth*, 48(4), 443-69.

[2] ARCIDIACONO, A., and J. JONES (2003): "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica*, 71, 933-946.

[3] ALVAREZ, J., BROWNING, M. and M. EJRNÆS (2002): "Modelling Income Processes with lots of heterogeneity," CAM Working Paper, 2002-01, University of Copenhagen.

[4] ATTANASIO, O., and S. DAVIS (1996): "Relative Wage Movements and the Distribution of Consumption," Journal of Political Economy, 104(6) , 1227-1262.

[5] BLUNDELL, R., and I. PRESTON (1998): "Consumption Inequality and Income Uncertainty," Quarterly Journal of Economics 113 (2), 603-640.

[6] BOWLUS, A., and J.M. ROBIN (2004): "Twenty years of Rising Inequality in US Lifetime labour Income Values," *Review of Economic Studies*, vol. 71, 3, pp. 709-742..

[7] BUCHINSKY, M., and J. HUNT (1999): "Wage Mobility in the United States," *Review of Economics and Statistics*, 351-68.

[8] BURDETT, K., and D. T. MORTENSEN (1998): "Wage Differentials, Employer Size and Unemployment", *International Economic Review*, 39, 257-73.

[9] BURKHAUSER, R. V., and J. G. POUPORE (1997): "A Cross-National Comparison of Permanent Inequality in the United States and Germany," *The Review of Economics and Statistics*, 79(1), 10-17.

[10] BURKHAUSER, R. V., D. HOLTZ-EAKIN and S. E. RHODY (1997): "Labor Earnings Mobility and Inequality in the United States and Germany during the Growth Years of the 1980s," *International Economic Review*, 38(4), 775-794.

[11] CUTLER, D. M., and L. F. KATZ (1992): "Rising Inequality? Changes in the Distribution of Income and Consumption in the 1980's," *American Economic Review*, 82(3), 546-551.

[12] FIELDS, G.S. (2005): "Does Income Mobility Equalize Longer-Term Incomes? New Measures of an Old concept," Cornell University.

[13] FLINN, C. (2002): "Labour Market Structure and Inequality: a Comparison of Italy and the US," *The Review of Economic Studies*, 69(3), 611-645.

[14] FRÉCHET, M. (1935): "Généralisations du Théorème des Probabilités Totales," *Fund. Math.*, 25, 379-387.

[15] FRIEDMAN, M. (1957): *A Theory of the Consumption Function*, Princeton: Princeton University Press.

[16] FRIEDMAN, M. (1962): *Capitalism and Freedom*, Chicago: University of Chicago Press.

[17] GEWEKE, J., and M. KEANE (2000): "An Empirical Analysis of Income Dynamics among Men in the PSID: 1968-1989," *Journal of Econometrics*, 96, 293-356.

[18] GEWEKE, J., and M. KEANE (2007): "Smoothly Mixing Regressions," *Journal of Econometrics*, 138, 252-290.

[19] GOTTSCHALK, P., and R. MOFFIT (1994): "The Growth of Earnings Instability in the US labour Market," *Brookings Papers on Economic Activity*, 2, 217-72.

[20] GUVENEN, F. (2005): "An Empirical Investigation of Labor Income Processes," mimeo.

[21] HAIDER, S. (2001): "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991," *Journal of Labor Economics*, vol. 19(4), 799-836.

[22] HAIDER, S. and G. SOLON (2006): "Life-Cycle Variation in the Association between Current and Lifetime Earnings," *American Economic Review*, vol. 96, 4, 1308-1320.

[23] HEATHCOTE, J., STORESLETTEN, K., and G. VIOLANTE (2004): "The Cross-Sectional Implications of Rising Wage Inequality in the United States," *mimeo.*

[24] HOROWITZ, J. L., and M. MARKATOU (1996): "Semiparametric Estimation of Regression Models for Panel Data," *Review of Economic Studies*, 63, 145-168.

[25] KRUEGER, D., and F. PERRI (2001): "Does Income Inequality Lead to Consumption Inequality? Empirical Findings and a Theoretical Explanation," mimeo.

[26] JOE, H. (1997): *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.

[27] JOLIVET, G., F. POSTEL-VINAY, and J.-M. ROBIN (2006): "The Empirical Content of the Job Search Model: Labor Mobility and Wage Dispersion in Europe and the U.S.," *European Economic Review*, 50(4), 877-907.

[28] KOTLARSKI, I. (1967): "On Characterizing the Gamma and Normal Distribution," *Pacific Journal of Mathematics*, 20, 69-76.

[29] MAASOUMI, E., and M. TREDE (2001): "Comparing Income Mobility in Germany and the United States using Generalized Entropy Mobility Measures," *The Review of Economics and Statistics*, 83(3), 551-559.

[30] MARTIN, J. P. (1996): "Measures of Replacement Rates for the Purpose of International Comparisons: a Note," *OECD Economic Survey,*, 26, 99-115.

[31] MEGHIR, C., and L. PISTAFERRI (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72, 1-32.

[32] MOFFIT, R., and P. GOTTSCHALK (2002): "Trends in the Transitory Variance of Earnings in the US," *The Economic Journal*, 112, C68-C73.

[33] NELSEN, R. B. (1998): *An Introduction to Copulas*. New-York: Springer Verlag.

[34] PLACKETT, R. L., (1965): "A Class of Bivariate Distributions," *J. Am. Statis. Assoc.*, 60, 516-522.

[35] POSTEL-VINAY, F., and J-M. ROBIN (2006): "Microeconometric Search-Matching Models and Matched Employer-Employee Data," in *The Proceedings of the 9th World Congress*

of the Econometric Society, edited by Richard Blundell, Whitney Newey, and Torsten Persson, Cambridge University Press.

[36] SHORROCKS, A. (1978): "Income Inequality and Income Mobility," *Journal of Economic Theory*, 19, 376-393.

[37] SKLAR, A. (1959): "Fonctions de Répartition à n Dimensions et leurs Marges," *Publ. Inst. Statist. Univ. Paris*, 8, 229-231.

[38] VAN DEN BERG G.J., and G. RIDDER (1998): "An empirical equilibrium search model of the labor market", *Econometrica*, 66, 1183-1221

[39] VAN KERM, P. (2004): "What Lies Behind Income Mobility? Reranking and Distributional Change in Belgium, Western Germany and the USA," *Economica*, 71(282), 223-239.

APPENDIX

A Copulas: an economic model

The main structure of our model is a single factor model with first-order Markov errors. We model individual transitory dynamics using copulas. In Section 3, we justified this choice by the ability of the copula representation to tightly parameterize the dependence structure while being flexible as far as marginal distributions are concerned. In this section of the Appendix, we argue that modelling the law of individual earnings processes using copulas also arises naturally in steady-state equilibrium models.

The prototypical steady-state search model describes an economy where a population of heterogeneous firms post different wages, and where a population of homogeneous workers sequentially search for a job both when they are unemployed and when they are employed. The wage offer distribution is some function $F(w)$ that depends on the distribution of productivity across different occupations and the search friction parameters. One can allow for heterogeneity in worker ability by assuming that the labour market is segmented (see for example Van den Berg and Ridder, 1998).

Search frictions are modelled as follows. Time is discrete. At the end of each period, an employee can be laid-off with probability δ or may draw an alternative wage offer with probability λ. This alternative offer is a draw w' from the equilibrium distribution F. Employees accept alternative offers w' strictly greater than the current wage w. In steady state, the cross-sectional earnings distribution is related to the wage offer distribution by the one-to-one relationship:

$$G(w) = \frac{F(w)}{1+\kappa\left[1-F(w)\right]} \Leftrightarrow F(w) = \frac{(1+\kappa)\,G(w)}{1+\kappa G(w)}, \tag{1}$$

where $\kappa = \lambda/\delta$.[27]

Conditional on being employed at time t and $t+1$ (alternative offers thus accrue with conditional probability $\frac{\lambda}{1-\delta}$), individual wage dynamics follows a first-order Markov process such that

$$\begin{aligned}\Pr\left\{w_{t+1} \leq w' | w_t = w\right\} &= \left(1 - \frac{\lambda}{1-\delta} + \frac{\lambda}{1-\delta}F(w')\right)\mathbf{1}\left\{w' \geq w\right\} \\ &= \frac{1}{1-\delta}\left[1 - \frac{\delta+\lambda}{1+\kappa G\left(w'\right)}\right]\mathbf{1}\left\{w' \geq w\right\},\end{aligned}$$

where $1\left\{w' \geq w\right\} = 1$ if $w' \geq w$ and 0 otherwise. Hence

$$\begin{aligned}\Pr\left\{w_{t+1} \leq w', w_t \leq w\right\} &= \frac{1}{1-\delta}\left[1 - \frac{\delta+\lambda}{1+\kappa G\left(w'\right)}\right]\left(G(w)\mathbf{1}\left\{w' \geq w\right\} + G(w')\mathbf{1}\left\{w' < w\right\}\right) \\ &= \frac{1}{1-\delta}\left[1 - \frac{\delta+\lambda}{1+\kappa G\left(w'\right)}\right]\min\left(G(w), G(w')\right).\end{aligned}$$

[27] See Burdett and Mortensen (1998). $\delta + \lambda\left[1 - F\left(w\right)\right]$ is the exit rate from the stock of employees paid less than w, which is proportional to $G\left(w\right)$. The inflow is itself proportional to $F\left(w\right)$.

The copula thus follows as:

$$C(u,v) = \frac{1}{1-\delta}\left[1 - \frac{\delta+\lambda}{1+\kappa v}\right]\min(u,v). \tag{2}$$

Hence, the steady-state search-wage-posting model naturally delivers a simple formula for the copula, which only depends on the friction parameters δ and λ and not on the equilibrium wage distribution F. We do not use this specification because it would not fit earnings mobility data well (as wage cuts are impossible). Yet, we want to put forward the interpretation that a steady-state labour market can be characterized, first, by the equilibrium wage distribution G (*i.e.* the marginal distribution of employees' wages), and, second, by a description of the way workers exchange positions (via mobility rate parameters embodied in the copula C).

Now, the economy is clearly not in steady state. Structural parameters like G, δ and λ do fluctuate over time. However, there is some evidence that steady-state flow conditions like (1) or the classical equality between the unemployment rate and $\delta/(\delta+\lambda_0)$, where λ_0 is the exit rate from unemployment, are empirically well verified (see Postel-Vinay and Robin, 2006, and Jolivet et al., 2006). This happens because net annual flows are small in relation to gross flows. Over a limited period of time, macro trends can thus be neglected and it seems that workers only exchange positions. This is how we justify our assumption that each subpanel S_j approximately reflects a steady-state economy with a specific set of parameters.

B The Plackett Copula

Plackett (1965) generalizes the independence condition for contingency tables. He shows that, if U and V are uniform random variables on $[0,1]$, then the following equation:

$$\frac{P(U \le u, V \le v)P(U > u, V > v)}{P(U \le u, V > v)P(U > u, V \le v)} = \tau + 1 \quad \forall (u,v), \tag{3}$$

where $\tau > -1$ is a given constant, has one single solution. This solution is a copula, and writes:

$$C(u,v;\tau) = \frac{1}{2}\tau^{-1}\left\{1 + \tau(u+v) - [(1+\tau(u+v))^2 - 4\tau(\tau+1)uv]^{1/2}\right\}.$$

Equation (3) shows that τ is a natural mobility index. More precisely, let us define the following ordering $\precsim_c$ on copulas, called the concordance ordering (*e.g.* Joe, 1997):

$$C_1 \precsim_c C_2 \quad \text{iff} \quad C_1(u,v) \le C_2(u,v), \quad \forall (u,v).$$

$\precsim_c$ is the first-order stochastic dominance ordering. It measures relative mobility: the ranks process governed by copula C_1 will be said more mobile than the ranks process governed by C_2 if $C_1 \precsim_c C_2$. Moreover, the concordance ordering possesses a lower and a upper bound (Fréchet, 1935). The lower bound C_L satisfies: $C_L(u,v) = \max(u+v-1,0)$. The upper bound C_U satisfies: $C_U(u,v) = \min(u,v)$.

The Plackett copula satisfies the following properties (Joe, 1997):

1. $C(.,.;\tau_1) \precsim_c C(.,.;\tau_2)$ for all $\tau_2 > \tau_1$.

2. $C(.,.;\tau) \to C_L$ when $\tau \to -1$.

3. $C(.,.;\tau) \to C_U$ when $\tau \to \infty$.

4. $C(.,.;\tau) \to C^{\perp}$ when $\tau \to 0$, where $C^{\perp}(u,v) = uv$ is the independent copula.

Therefore, the Plackett copula is mobility decreasing with respect to its parameter, and the Plackett family covers the whole range of bivariate dependence, from immobility (C_U) to independence ($C^{\perp}$) and perfect mobility (C_L).

Simulation. Let u be earnings ranks at time t. For given u, we here show how to simulate earnings ranks at $t+1$ (v) from the Plackett copula. Note that for $u \in [0,1]$ given, $\frac{\partial C}{\partial u}(u,v)$ is uniform(0,1). So in order to simulate a v draw a uniform(0,1) variable w and solve for v in $\frac{\partial C}{\partial u}(u,v) = w$. In the case of the Plackett copula, this implicit equation writes as a quadratic polynomial equation with two roots in $[0,1]$. Choose the highest or the lowest root with probability $\frac{1}{2}$.

C Remarks on identification

Let us consider the model

$$y_{it} = \eta_i + \varepsilon_{it}, \quad t = 1,2,3.$$

We assume that individual trajectories are i.i.d. and that η_i is independent of (ε_{it}). Let F_X denote the c.d.f. of each variable X_i from an i.i.d. sample $\{X_1, ..., X_N\}$, with $f = F'$ for the density. Assume that (ε_{it}) is first-order Markov with $C_t(\cdot,\cdot)$ being the copula of $(\varepsilon_{it}, \varepsilon_{it+1})$, for $t = 1,2$. We now provide an algorithm to reconstruct F_η, F_{ε_t}, and C_t from the distribution of (y_{i1}, y_{i2}, y_{i3}).

Start by assuming that the distribution of the individual-specific effect η_i is known. One can solve each convolution equation $y_{it} = \eta_i + \varepsilon_{it}$ for the distribution of ε_{it}. So, F_{ε_t} is known given F_η and F_{y_t}.

Then, consider the joint distribution of (y_{i1}, y_{i2}):

$$F_{y_1,y_2}(x_1,x_2) = \int f_\eta(u) C_1 \left[F_{\varepsilon_1}(x_1 - u), F_{\varepsilon_2}(x_2 - u)\right] du.$$

This equation is a functional equation that relates the bivariate functions F_{y_1,y_2} and C_1. Assuming that this functional equation has a unique solution in C_1, then C_1 is known given F_η. The same argument identifies C_2 from the distribution of (y_{i2}, y_{i3}).

One can finally consider the joint distribution of (y_{i1}, y_{i2}, y_{i3}):

$$F_{y_1,y_2,y_3}(x_1,x_2,x_3) = \int f_\eta(u) du \int_0^{F_{\varepsilon_2}(x_2-u)} \partial_2 C_1 \left(F_{\varepsilon_1}(x_1-u), v\right) \partial_1 C_2 \left(v, F_{\varepsilon_3}(x_3-u)\right) dv,$$

or the distribution of (y_{i1}, y_{i3}):

$$F_{y_1,y_3}(x_1,x_3) = \int f_\eta(u)du \int_0^1 \partial_2 C_1 \left(F_{\varepsilon_1}(x_1-u), v\right) \partial_1 C_2 \left(v, F_{\varepsilon_3}(x_3-u)\right) dv,$$

where $\partial_1 C_1$ and $\partial_2 C_1$ denote the partial derivatives of C_1 with respect to first and second arguments. These equations potentially (over)identify the distribution of η_i.

We terminate this section by considering, as an example, the case of a Gaussian copula with Gaussian margins (or a nonstationary Gaussian AR(1) model):

$$\frac{\varepsilon_{it}}{\sigma_t} = \tau_t \frac{\varepsilon_{it-1}}{\sigma_{t-1}} + \sqrt{1-\tau_t^2}\nu_{it}, \quad t = 2,3,$$

with $\varepsilon_{it} \sim N\left(0, \sigma_t^2\right)$ and $\nu_{it} \sim N(0,1)$.[28] Then,

$$\begin{aligned}
\text{Var}(y_{it}) &= \text{Var}(\eta_i) + \text{Var}(\varepsilon_{it}) = \text{Var}(\eta_i) + \sigma_t^2, \\
\text{Cov}(y_{i1}, y_{i2}) &= \text{Var}(\eta_i) + \text{Cov}(\varepsilon_{i1}, \varepsilon_{i2}) = \text{Var}(\eta_i) + \tau_2\sigma_1\sigma_2, \\
\text{Cov}(y_{i2}, y_{i3}) &= \text{Var}(\eta_i) + \text{Cov}(\varepsilon_{i2}, \varepsilon_{i3}) = \text{Var}(\eta_i) + \tau_3\sigma_2\sigma_3, \\
\text{Cov}(y_{i1}, y_{i3}) &= \text{Var}(\eta_i) + \text{Cov}(\varepsilon_{i1}, \varepsilon_{i3}) = \text{Var}(\eta_i) + \tau_2\tau_3\sigma_1\sigma_3.
\end{aligned}$$

The first equation identifies σ_t^2 given $\text{Var}(\eta_i)$. The second and third equations identify τ_2 and τ_3 given $\text{Var}(\eta_i)$. Finally, the last equation identifies $\text{Var}(\eta_i)$ as, for example:

$$\text{Var}(\eta_i) = \frac{\text{Cov}(y_{i1}, y_{i2})\,\text{Cov}(y_{i2}, y_{i3}) - \text{Var}(y_{i2})\,\text{Cov}(y_{i1}, y_{i3})}{\text{Cov}(y_{i1}, y_{i2}) + \text{Cov}(y_{i2}, y_{i3}) - \text{Var}(y_{i2}) - \text{Cov}(y_{i1}, y_{i3})}.$$

D Estimation algorithm

Fix K, and a grid $\underline{\eta}_1, ..., \underline{\eta}_K$ for the distribution of the permanent component η_i, and fix M the number of Gaussian components for the marginal distribution of transitory shocks ε_{it}.

Let $\mathbf{x}_i = (x_{i1}, ..., x_{iT})$, $\mathbf{y}_i = (y_{i1}, ..., y_{iT})$ and $\mathbf{e}_i = (e_{i1}, ..., e_{iT})$ be the observations for individual i. Let $j(i) \in \{1, ..., J\}$ be the index of the panel sample these observations belongs to (*i.e.* $j(i) = j$ such that $i \in \mathcal{S}_j$).

The likelihood for one individual observation is the discrete mixture:

$$\ell_j(\mathbf{y}_i, \mathbf{e}_i | \mathbf{x}_i) = \sum_{k=1}^K p_k \ell_j(\mathbf{y}_i, \mathbf{e}_i | \mathbf{x}_i, k), \quad i \in \mathcal{S}_j,$$

[28] With non Gaussian margins, the Gaussian copula corresponds to the model:

$$\Phi^{-1}\left[F_{\varepsilon_t}(\varepsilon_{it})\right] = \tau_t \Phi^{-1}\left[F_{\varepsilon_t}(\varepsilon_{it-1})\right] + \sqrt{1-\tau_t^2}\nu_{it}.$$

where

$$\ell_j(\mathbf{y}_i, \mathbf{e}_i|\mathbf{x}_i, k) = P_j^0(e_{i1}|x_{i1}, k) \times \prod_{t=1}^{T-1} P_j\left(e_{i,t+1}|e_{it}, x_{i,t+1}, k\right)$$
$$\times \prod_{t=1}^{T} \left[\frac{1}{\exp\left(\alpha_j x_{it}\right)} f_j\left(\epsilon\left(y_{it}|x_{it}, k\right)|x_{it}\right)\right]^{e_{it}}$$
$$\times \prod_{t=1}^{T-1} c\left[F_j\left(\epsilon\left(y_{it}|x_{it}, k\right)|x_{it}\right), F_j\left(\epsilon\left(y_{it+1}|x_{it+1}, k\right)\right); \exp\left(\gamma_j x_{it}\right) - 1\right]^{e_{it}e_{it+1}},$$

where c is Plakett's copula density, with parameter $\exp\left(\gamma_j x_{it}\right) - 1$, and where

$$\begin{aligned}
\epsilon\left(y_{it}|x_{it}, k\right) &= \frac{y_{it} - \beta_j x_{it} - \sigma_j \underline{\eta}_k}{\exp\left(\alpha_j x_{it}\right)}, \\
P_j^0(e_{i1}|x_{i1}, k) &= \Phi\left(\delta_{0j} x_{i1} + \zeta_{0j}\underline{\eta}_k\right)^{e_{i1}} \Phi\left(-\delta_{0j} x_{i1} - \zeta_{0j}\underline{\eta}_k\right)^{1-e_{i1}}, \\
P_j(e_{i,t+1}|e_{it}, x_{i,t+1}, k) &= \Phi\left(\rho_j e_{it} + \delta_{1j} x_{i,t+1} + \zeta_{1j}\underline{\eta}_k\right)^{e_{i,t+1}} \Phi\left(-\rho_j e_{it} - \delta_{1j} x_{i,t+1} - \zeta_{1j}\underline{\eta}_k\right)^{1-e_{i,t+1}}, \\
f_j(\varepsilon_{it}|x_{it}) &= \sum_{m=1}^{M} \pi_{mj} f_j(\varepsilon_{it}|x_{it}, m) = F_j'(\varepsilon_{it}|x_{it}), \\
f_j(\varepsilon_{it}|x_{it}, m) &= \frac{1}{\omega_{mj}} \varphi\left(\frac{\varepsilon_{it} - \mu_{mj}}{\omega_{mj}}\right) = F_j'(\varepsilon_{it}|x_{it}, m).
\end{aligned}$$

Note that we allow for a specific marginal distribution of transitory shocks ε_{it} to each subpanel j.

E-step. For initial values of the parameters p_k, $\left(\mu_{mj}, \pi_{mj}, \omega_{mj}\right)$, $\left(\delta_{0j}, \mu_{0j}\right)$, $\left(\rho_j, \delta_{1j}, \mu_{1j}\right)$, $\left(\beta_j, \sigma_j, \alpha_j\right)$, and γ_j, for each individual in the samples $\mathcal{S}_j$, $j = 1, ..., J$, compute the posterior probabilities of the latent types of η_i and ε_{it}:

$$\begin{aligned}
p_j(k|\mathbf{y}_i, \mathbf{e}_i, \mathbf{x}_i) &= \frac{p_k \ell_j(\mathbf{y}_i, \mathbf{e}_i|\mathbf{x}_i, k)}{\ell_j(\mathbf{y}_i, \mathbf{e}_i|\mathbf{x}_i)}, \quad k = 1, ..., K, \\
\pi_j(m|y_{it}, x_{it}, k) &= \frac{\pi_{mj} f_j(\epsilon\left(y_{it}|x_{it}, k\right)|x_{it}, m)}{f_j(\epsilon\left(y_{it}|x_{it}, k\right)|x_{it})}, \quad m = 1, ..., M.
\end{aligned}$$

M-step. With $p_j(k|\mathbf{y}_i, \mathbf{e}_i, \mathbf{x}_i)$ and $\pi_j(m|\varepsilon_{it}, x_{it}, k)$ set to these values,

- Update p_k by maximizing

$$\sum_{j=1}^{J} \sum_{i \in \mathcal{S}_j} \sum_{k=1}^{K} p_j(k|\mathbf{y}_i, \mathbf{e}_i, \mathbf{x}_i) \ln p_k,$$

  with respect to p_k and subject to:

$$\sum_{k=1}^{K} p_k = 1, \quad \sum_{k=1}^{K} p_k \underline{\eta}_k = 0, \quad \sum_{k=1}^{K} p_k \underline{\eta}_k^2 = 1.$$

  Then for each subpanel j perform the following steps:

- Maximize
$$\sum_{i\in\mathcal{S}_j}\sum_{m=1}^{M}\sum_{k=1}^{K}p_j(k|\mathbf{y}_i,\mathbf{e}_i,\mathbf{x}_i)\sum_{t/e_{it}=1}\pi_j(m|y_{it},x_{it},k)\ln\pi_{mj},$$
with respect to π_{mj} and s.t.
$$\sum_{m=1}^{M}\pi_{mj}=1,\quad \sum_{m=1}^{M}\pi_{mj}\mu_{mj}=0,\quad \sum_{m=1}^{M}\pi_{mj}\omega_{mj}^2=1.$$

- Maximize
$$\sum_{i\in\mathcal{S}_j}\sum_{k=1}^{K}p_j(k|\mathbf{y}_i,\mathbf{e}_i,\mathbf{x}_i)\ln P_j^0(e_{i1}|x_{i1},k;\delta_{0j},\mu_{0j}),$$
with respect to $\left(\delta_{0j},\mu_{0j}\right)$, and
$$\sum_{i\in\mathcal{S}_j}\sum_{k=1}^{K}\sum_{t=1}^{T-1}p_j(k|\mathbf{y}_i,\mathbf{e}_i,\mathbf{x}_i)\ln P_j(e_{i,t+1}|e_{it},x_{i,t+1},k;\rho_j,\delta_{1j},\mu_{1j}),$$
with respect to $\left(\rho_j,\delta_{1j},\mu_{1j}\right)$.

- Minimize
$$\sum_{i\in\mathcal{S}_j}\sum_{k=1}^{K}p_j\left(k|\mathbf{y}_i,\mathbf{e}_i,\mathbf{x}_i\right)\left(y_{it}-\beta_j x_{it}-\sigma_j\underline{\eta}_k\right)^2,$$
with respect to β_j and σ_j. Then, compute residuals $u\left(y_{it}|x_{it},k\right)=y_{it}-\beta_j x_{it}-\sigma_j\underline{\eta}_k$ and minimize
$$\sum_{i\in\mathcal{S}_j}\sum_{k=1}^{K}p_j\left(k|\mathbf{y}_i,\mathbf{e}_i,\mathbf{x}_i\right)\left[u\left(y_{it}|x_{it},k\right)^2-\exp\left(2\alpha_j x_{it}\right)\right]^2,$$
with respect to α_j.

- Compute standardized residuals $\epsilon\left(y_{it}|x_{it},k\right)$ and maximize
$$\sum_{i\in\mathcal{S}_j}\sum_{k=1}^{K}p_j\left(k|\mathbf{y}_i,\mathbf{e}_i,\mathbf{x}_i\right)\sum_{m=1}^{M}\sum_{t/e_{it}=1}\pi_j(m|y_{it},x_{it},k)\ln\left[\frac{1}{\omega_{mj}}\varphi\left(\frac{\epsilon\left(y_{it}|x_{it},k\right)-\mu_{mj}}{\omega_{mj}}\right)\right],$$
with respect to μ_{mj} and ω_{mj}. The solution can be written as weighted means and variances of the standardized residuals.

- Compute ranks:
$$r\left(y_{it}|x_{it},k\right)=F_j\left(\epsilon\left(y_{it}|x_{it},k\right)|x_{it}\right).$$
and maximize
$$\sum_{i\in\mathcal{S}_j}\sum_{k=1}^{K}\sum_{m=1}^{M}p_j\left(k|\mathbf{y}_i,\mathbf{e}_i,\mathbf{x}_i\right)\sum_{t/e_{it}=e_{i,t+1}=1}\ln c\left[r\left(y_{it}|x_{it},k\right),r\left(y_{i,t+1}|x_{i,t+1},k\right);\exp\left(\gamma_j x_{it}\right)-1\right],$$
with respect to γ_j.

- Lastly, pooling all subpanels together regress the estimates of j-specific parameters on a constant, j, j^2 and j^3. Compute the restricted parameter estimates, and start a new E-step. Iterate until convergence.

A STATA programme is available from the authors.