

La mesure des inégalités de long terme avec des panels courts : 1990-2000

Stéphane Bonhomme and Jean-Marc Robin

Volume 84, Number 4, décembre 2008

URI: <https://id.erudit.org/iderudit/039323ar>

DOI: <https://doi.org/10.7202/039323ar>

[See table of contents](#)

Publisher(s)

HEC Montréal

ISSN

0001-771X (print)

1710-3991 (digital)

[Explore this journal](#)

Cite this article

Bonhomme, S. & Robin, J.-M. (2008). La mesure des inégalités de long terme avec des panels courts : 1990-2000. *L'Actualité économique*, 84(4), 325–363. <https://doi.org/10.7202/039323ar>

Article abstract

In this paper, we build a model of earnings dynamics and estimate it on a representative three-year panel, the French Labour Force Survey. We use the model to simulate individual earnings trajectories beyond the period of observation, and to compute permanent income values. We measure the ratio between inequality in present values of lifetime earnings and current inequality. This ratio is less than one, showing that earnings mobility is equalizing. Moreover, it is very sensitive to unemployment risk. Mobility reduces long-run inequality in a representative sample of employed and unemployed workers, in large part because unemployment is a transitory state. In contrast, unemployment risk generates a lot of inequality in the sample of employed workers, and reduces by one half the equalizing force of earnings mobility.

LA MESURE DES INÉGALITÉS DE LONG TERME AVEC DES PANELS COURTS : 1990-2000*

Stéphane BONHOMME

Centro de Estudios Monetarios y Financieros (CEMFI)

Jean-Marc ROBIN

Paris School of Economics,

Université Paris I Panthéon Sorbonne,

University College London

RÉSUMÉ – Dans cette étude, nous proposons un modèle de la dynamique salariale adapté à une estimation à partir de panels courts comme l'enquête Emploi. Nous utilisons le modèle pour simuler des trajectoires individuelles de salaires au-delà de la période d'enquête et calculer des revenus permanents. Nous mesurons le rapport entre l'inégalité de revenus permanents (inégalité de long terme) et l'inégalité salariale de coupe. Ce rapport est inférieur à un, preuve que la mobilité des revenus est égalisatrice. Cependant, nous constatons le rôle essentiel joué par le risque de chômage dans cette mesure. La mobilité réduit les inégalités sur le long terme dans un échantillon représentatif de travailleurs employés ou en chômage, en grande partie parce que le chômage ne dure pas éternellement. À l'inverse, le risque de chômage est fortement générateur d'inégalité dans l'échantillon des employés. Nous mesurons qu'il annule ainsi la moitié du bénéfice égalisateur de la mobilité salariale.

ABSTRACT – In this paper, we build a model of earnings dynamics and estimate it on a representative three-year panel, the French Labour Force Survey. We use the model to simulate individual earnings trajectories beyond the period of observation, and to compute permanent income values. We measure the ratio between inequality in present values of lifetime earnings and current inequality. This ratio is less than one, showing that earnings mobility is equalizing. Moreover, it is very sensitive to unemployment risk. Mobility reduces long-run inequality in a representative sample of employed and unemployed workers, in large part because unemployment is a transitory state. In contrast, unemployment risk generates a lot of inequality in the sample of employed workers, and reduces by one half the equalizing force of earnings mobility.

* Étude financée par la Direction de l'Animation de la Recherche, des Études et des Statistiques (DARES) du ministère de l'Emploi et de la Solidarité (convention no 4028). Cet article reprend pour une large part le contenu de l'article de S. Bonhomme et J.-M. Robin (2009) dont la *Review of Economic Studies* nous a autorisé la reproduction. Il a fourni l'essentiel du matériel de la conférence François-Albert-Angers donnée par J.-M. Robin au 48^e congrès de la Société canadienne de science économique (SCSE, Montebello, 14-16 mai, 2008).

INTRODUCTION

L'objectif de cette recherche est de produire une méthodologie permettant d'étudier l'évolution des inégalités de salaires de long terme. Par inégalité de salaires de long terme, nous entendons une mesure de la dispersion de moyennes individuelles de salaires sur une période suffisamment longue pour lisser les chocs conjoncturels transitoires¹. La principale difficulté est la suivante : comment conjuguer une étude de l'évolution de ces inégalités avec une définition de la variable d'intérêt qui requiert elle-même des données individuelles temporelles? La dynamique des trajectoires individuelles se superposant à celle du changement macroéconomique, il convient de développer une méthodologie réduisant au maximum le nombre de période requises pour construire les salaires permanents.

Pour cela il nous faut abandonner l'idée initiale du calcul de moyennes individuelles. Ce n'est pas une procédure suffisamment performante pour éliminer les chocs transitoires². Nous modélisons la dynamique des salaires individuels comme une somme de deux composantes : une composante permanente (un effet fixe individuel) et une composante stochastique stationnaire. Ces deux composantes sont affectées de pondérations qui changent au cours du temps. Une fois le modèle estimé pour une date donnée, on simule des trajectoires individuelles sur une longue période et on calcule des salaires permanents simulés. De cette façon, une série de panels courts comme les différentes vagues de l'enquête Emploi de l'Institut national de la statistique et des études économiques (INSEE) suffit pour étudier, année après année, le changement des inégalités de salaires de long terme. Cette approche a déjà été employée par Flinn (2002) pour comparer les États-Unis et l'Italie, et par Haider (2001) et Bowlus et Robin (2004) pour étudier l'évolution des inégalités de long terme aux États-Unis. Ce qui fait l'originalité de la présente étude est le traitement de l'hétérogénéité inobservée, ici la composante permanente des salaires, qui est modélisée comme un effet fixe individuel, dans un modèle dynamique non linéaire.

Il est étonnant de constater que l'immense majorité des modèles de salaires sont linéaires. Ainsi, la structure la plus couramment utilisée représente les salaires comme la somme d'une composante déterministe (effet de l'éducation, de l'âge, etc.), un effet fixe individuel, une composante permanente stochastique modélisée comme une marche aléatoire ($y_{it}^p = y_{it-1}^p + \varepsilon_{it}$), une composante stochastique transitoire modélisée comme un processus moyenne mobile (un MA(1) : $y_{it}^T = \eta_{it} + \phi\eta_{it-1}$,

1. Shorrocks (1978) a axiomatisé la mesure de la mobilité individuelle de revenus et justifié la construction d'un indice de mobilité comme rapport des inégalités de long terme aux inégalités de coupe. On trouvera des applications récentes de cette approche dans Burkhauser, Holtz-Eakin and Rhody (1997), Burkhauser et Poupore, (1997), Maasoumi et Trede (2001), Aaberge *et al.*, (2002) et Van Kerm (2004).

2. Gottschalk and Moffitt (2002) soutiennent que cinq années au moins sont nécessaires pour le calcul des salaires ou revenus permanents. « *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).

par exemple), et enfin, parfois, une erreur de mesure³. A l'inverse, la plupart des études statistiques de la mobilité individuelle se fondent sur l'examen de matrices de probabilités de transition entre quartiles, quintiles ou déciles de salaires. La première approche cherche à filtrer différentes composantes additives, dont la dynamique est linéaire. La seconde met l'accent sur le caractère réducteur de cette simplicité, un petit nombre de paramètres d'autocorrélations ne pouvant décrire toute la complexité de la mobilité. En contrepartie, cette dernière filtre mal les erreurs de mesure et peine à reproduire correctement la dynamique de long terme.

Nous combinons ces deux approches. Concrètement, nous retenons l'idée que le processus observé agrège des chocs de nature différente. Cependant, il est pertinent de chercher à décrire différemment la mobilité selon l'endroit de la distribution des salaires auquel on se trouve à un moment donné. Une matrice de probabilités de transitions entre quantiles est une approximation de la dynamique des *rangs* des salaires dans les distributions de coupe. La littérature statistique donne un nom à la loi jointe de deux rangs consécutifs : une *copule*. Nous verrons que l'on peut décrire assez précisément les transitions entre quantiles de salaires à l'aide de copules paramétriques simples. Par ailleurs, en modélisant les processus individuels comme agrégation de deux composantes, une constante individuelle spécifique et une composante markovienne, nous espérons améliorer la capacité des modèles de mobilité à décrire la dynamique de long terme.

Nous utilisons l'enquête Emploi de 1990 à 2002. Avant 1990 les salaires étaient mesurés en tranches. Après 2002 le protocole de l'enquête a changé. Malgré ces limitations, la période 1990-2002 présente l'avantage de décrire un cycle d'activité de l'économie française complet, 1990 et 2000 étant deux points hauts du cycle. Pour chaque vague différente de l'enquête Emploi (de 1990 à 2000), nous utilisons le modèle pour simuler des trajectoires individuelles de salaires au-delà de la période d'enquête et calculer des revenus permanents.

Nous mesurons le rapport entre l'inégalité de revenus permanents (inégalité de long terme) et l'inégalité salariale de coupe. Ce rapport est inférieur à un, preuve que la mobilité des revenus est égalisatrice. Cependant, nous constatons le rôle essentiel joué par le risque de chômage dans cette mesure. La mobilité réduit les inégalités sur le long terme dans un échantillon représentatif de travailleurs employés ou au chômage, en grande partie parce que le chômage ne dure pas éternellement. À l'inverse, le risque de chômage est fortement générateur d'inégalité dans l'échantillon des employés. Nous mesurons qu'il annule ainsi la moitié du bénéfice égalisateur de la mobilité salariale.

Le plan de cet article est le suivant. Dans un premier temps, nous développons un modèle de dynamique des revenus de l'emploi et du chômage. Puis nous exposons les données et décrivons la méthode d'estimation. Enfin, nous appliquons le

3. Voir par exemple Hall et Mishkin (1982), et Abowd et Card (1989).

modèle à l'étude de l'évolution des inégalités de salaires de long terme au cours des années 1990.

1. LE MODÈLE

Les données sont constituées d'une collection de J panels courts :

$$\{(y_{it}, x_{it}), i \in S_j, t = t_j, t_j + 1, \dots, t_j + T\}$$

où S_j est un ensemble d'indices individuels, t_j est la date de la première observation du j^e panel, y_{it} est le log du salaire de l'individu i à la date t , x_{it} est un vecteur de variables telles que l'éducation et l'âge et T est la longueur de chaque panel. Dans l'enquête Emploi de l'INSEE d'avant 2003, un tiers environ des ménages étaient renouvelés chaque année. Si bien que chaque individu restait au plus $T = 3$ ans dans le panel. Le vecteur x_{it} comprend les variables suivantes : l'expérience (âge moins âge de fin d'études), son carré ainsi que cinq indicatrices de niveau de diplôme : primaire (I), collège (II), bac (III), bac + 2 (IV) et bac + 5 (V).

L'avantage de ces panels rotatifs, par rapport à des panels longs comme le *Panel Survey of Income Dynamics* (PSID) américain, est que chaque échantillon contient un grand nombre d'observations, ce qui permet une excellente représentativité. L'avantage sur des panels administratifs, comme les DADS (Déclarations annuelles des salaires), est qu'on sait beaucoup de choses sur les individus et que ceux-ci sont suivis y compris lorsqu'ils passent par le chômage ou prennent un emploi public. Toutefois, à la différence du PSID ou du SIPP (*Survey of Income and Program Participation*) américains, l'enquête ne permet pas de suivre les individus lorsqu'ils changent de logement. C'est une source d'attrition qu'on supposera exogène⁴.

Nous postulons la décomposition additive suivante entre composantes déterministes, permanentes et transitoires :

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

Notez que les paramètres β_j , σ_j et α_j sont indexés par l'indice du panel. On voudrait qu'ils soient fonction du temps calendaire t . Mais cela compliquerait considérablement l'estimation. En les faisant dépendre de l'indice du panel, on voit qu'on peut en principe estimer des modèles séparés pour chaque vague. Une indexation stricte sur t obligerait à estimer tous les paramètres en même temps et à empiler tous les panels, soit plusieurs centaines de milliers d'observations. D'autre part, une indexation sur le numéro de panel permet de lisser les variations calendaires des paramètres.

4. Une expérience que nous avons réalisée avec les données SIPP (en comparant les résultats d'estimation sur l'échantillon où l'on suit les individus qui déménagent et un autre où on ne les suit pas) semblent donner raison à cette hypothèse. Pour cette raison, nous supposons, pour simplifier les notations, que chaque sous-panel est cylindré (d'une longueur fixe T).

1.1 Composante permanente

La variable d'hétérogénéité inobservée η_i est modélisé comme un effet aléatoire indépendant de x_{it} et du processus ε_{it} . La distribution de η_i dans la population est supposée stationnaire (ne changeant pas au cours du temps), de moyenne 0 et de variance unitaire. Notez néanmoins que la composante permanente $\sigma_j \eta_i$ voit sa variance σ_j^2 changer avec la vague d'enquête j .

De façon à conserver le maximum d'information, nous ne postulons pas une spécification paramétrique pour la loi de η_i . Nous privilégions plutôt une approche non paramétrique. Concrètement, nous supposons simplement que la loi de η_i est une loi discrète de support $\{\underline{\eta}_1, \dots, \underline{\eta}_K\}$. Soit p_k la probabilité de chaque point de masse $\underline{\eta}_k$. La standardisation impose les contraintes :

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

$G(\eta) = \sum_{k=1}^K p_k 1\{\underline{\eta}_k \leq \eta\}$ est donc la fonction de répartition de η_i .

Les points de support sont arbitrairement choisis selon une grille gaussienne sur \mathbb{R} : $\underline{\eta}_k = 2\Phi^{-1}(k / (K + 1))$, $k = 1, \dots, K$. Plus K sera grand, plus la grille sera dense et plus précise sera l'approximation discrète. Nous pourrions lisser cette approximation, mais le caractère discret trouve une utilité dans la construction de l'algorithme de maximisation de la vraisemblance (Cf. *infra*).

1.2 Composante transitoire

Nous modélisons le choc transitoire ε_{it} comme un processus stationnaire markovien d'ordre un; c'est-à-dire que la dernière observation ε_{it-1} suffit à résumer tout le passé de ε_{it} . La moyenne et la variance marginales de ε_{it} sont aussi normalisées à 0 et 1 dans la mesure où le terme $\exp(\alpha_j x_{it})$ devant ε_{it} nous permet de modéliser une dépendance très générale de la volatilité des salaires à x_{it} et au temps calendaire (via l'indexation du paramètre α_j sur le numéro du panel).

Comme pour η_i nous voulons être aussi non paramétrique que possible dans la dimension de coupe. La taille de l'échantillon nous y autorise. Cette fois, il n'y a aucune utilité à employer une approximation discrète. Nous préférons donc utiliser une approximation continue. Nous choisissons comme modèle un mélange de normales. Pratiquement, nous approximations la fonction de répartition de ε_{it} , $i \in S_j$ de la façon suivante :

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

où les pondérations π_{mj} , les moyennes μ_{mj} et écarts-types ω_{mj} de chaque composante du mélange (Φ dénote la fonction de répartition de la loi normale centrée réduite) sont fonction du temps par le biais de l'indice du panel j .

La normalisation de la moyenne et de la variance de ϵ_{it} impose :

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

1.3 Dynamique transitoire

Il reste à modéliser la dynamique de ϵ_{it} . Comme nous avons déjà défini la loi marginale de ϵ_{it} , il suffit de décrire la façon dont les individus se déplacent dans cette distribution⁵. Pratiquement, il s'agit de décrire la dynamique de $F_j(\epsilon_{it})$. Nous supposons que le processus $F_j(\epsilon_{it})$ est markovien du premier ordre. La fonction de répartition de la loi de $F_j(\epsilon_{it})$ sachant $F_j(\epsilon_{it-1})$ est une *copule*⁶. On peut penser à la copule comme la version continue des matrices de probabilités de transition entre quintiles ou déciles. Différentes copules paramétriques ont été proposées dans la littérature statistique. La plus connue est sans doute la copule gaussienne qui revient à supposer que $\Phi^{-1}(F_j(\epsilon_{it}))$ est un processus AR(1).

Appelons $C_j(u, v)$ la copule de $(\epsilon_{it-1}, \epsilon_{it})$ en un point (u, v) , c'est-à-dire la fonction de répartition de $(F_j(\epsilon_{it-1}), F_j(\epsilon_{it}))$ en (u, v) ⁷. La factorisation habituelle de la densité de la loi jointe de $(\epsilon_{it-1}, \epsilon_{it})$ en le produit de la loi conditionnelle de ϵ_{it} sachant ϵ_{it-1} et de la marginale de ϵ_{it-1} peut être avantageusement remplacée par le produit des deux marginales et de la densité des rangs :

$$\begin{aligned} \ell[\epsilon_{it-1}, \epsilon_{it}] &= \ell[\epsilon_{it} | \epsilon_{it-1}] \ell[\epsilon_{it-1}] \\ &= \ell[\epsilon_{it}] \ell[\epsilon_{it-1}] \ell[F_j(\epsilon_{it-1}), F_j(\epsilon_{it})], \quad j \in S_j, \end{aligned}$$

en notant $\ell[X]$ la fonction de densité d'une variable X . De plus, pour tout $(u, v) \in [0, 1]^2$, la densité de $(F_j(\epsilon_{it-1}), F_j(\epsilon_{it}))$ en (u, v) est

$$\begin{aligned} \ell[F_j(\epsilon_{it-1}), F_j(\epsilon_{it})](u, v) &= c_j(u, v) \text{ (disons)} \\ &= \frac{\partial^2 C_j(u, v)}{\partial u \partial v}. \end{aligned}$$

C'est une factorisation très utile car elle permet de séparer complètement la paramétrisation des lois marginales et celle de la dynamique du processus, à la différence de la factorisation habituelle où la marginale de ϵ_{it} se déduit par intégration de la loi jointe. Une fois qu'on a spécifié $\ell[\epsilon_{it} | \epsilon_{it-1}]$ et $\ell[\epsilon_{it-1}]$, il ne reste plus de marge de liberté pour spécifier $\ell[\epsilon_{it}]$.

5. Nous avons imposé la stationnarité des distributions marginales. Rien n'empêche cependant d'autoriser des variations conjoncturelles et tendanciennes.

6. On doit la notion de copule à Sklar (1959). Cf. Nelsen (1998) et Joe (1997) pour une utilisation des copules en statistique et en finance.

7. Notez que pour deux variables aléatoires uniformes U et V , la loi jointe de (U, V) et la loi conditionnelle de U sachant V ou de V sachant U coïncident.

Enfin, du point de vue de l'interprétation économique, la factorisation copule permet de donner une représentation statistique à deux dynamiques distinctes : celle des inégalités de revenus d'une part (c'est-à-dire la façon dont $\ell[\varepsilon_{it}]$ change avec t) et celle des trajectoires individuelles relatives, d'autre part (décrite par C_j). Les inégalités de revenu peuvent s'accroître sans que changent les positions relatives des individus dans la distribution des revenus. Les inégalités de revenu peuvent rester inchangées alors même que les individus sont mobiles et échangent leurs positions.

La littérature statistique offre un grand choix de copules paramétriques. Nous en avons testé un certain nombre. Pour cela, nous avons pris les observations de 1990 et 1991 (voir la prochaine section pour une description des données). Nous avons régressé les logs des salaires sur les variables habituelles d'éducation et d'expérience. Puis nous avons estimé par la méthode du maximum de vraisemblance les paramètres de quelques copules monoparamétriques sur les rangs, en 1990 et 1991, des résidus des régressions⁸. Enfin, nous avons calculé les probabilités des transitions interquintiles induites par chaque modèle. Le tableau 1 donne une illustration des résultats obtenus pour la copule gaussienne et celle qui a donné les meilleurs résultats, la copule de Plackett :

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

Le paramètre τ caractérise de façon non ambiguë la mobilité. Plus τ est grand et moins les individus sont mobiles (lorsque $\tau \rightarrow \infty$, la matrice de mobilité tend vers l'identité : $C(u, v) = \min(u, v)$). Plus τ est faible et plus les individus sont mobiles en échangeant leurs positions (lorsque $\tau = -1$, la matrice de mobilité tend vers une matrice miroir de la matrice identité, avec des uns sur la diagonale opposée : $C(u, v) = \max(u + v - 1, 0)$). Lorsque $\tau \rightarrow 0$ ($C(u, v) = uv$), on a l'indépendance complète (la matrice de mobilité est telle que chaque transition est équiprobable).

L'ajustement des probabilités de transitions interquintiles obtenu avec la copule de Plackett dépasse largement celui des autres copules et celui de la copule gaussienne en particulier. Nous retiendrons donc la copule de Plackett en spécifiant le paramètre τ comme une fonction de l'indice du panel j et de x_{it} :

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

1.4 Transitions emploi/non-emploi

Nous décrivons aussi les transitions emploi/non-emploi. Soit ε_{it} une variable indicatrice qui vaut 1 si l'individu $i \in S_j$ est employé en t et 0 sinon. On supposera que

8. Nous avons testé séparément la copule gaussienne, de Plackett, Frank, Gumbel, Joe, Clayton, FGM et log. Un ajustement vraisemblablement supérieur serait obtenu avec des copules plus complexes comme la copule de Student.

TABLEAU 1
AJUSTEMENT DES TRANSITIONS INTERQUINTILES POUR LES COPULES
DE GAUSS ET DE PLACKETT
(1990-1991)

Observé	Gauss	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

$$e_{it} = 1\{\delta_{0j}x_{it} + \zeta_{0j}\eta_i + \xi_{it} > 0\}, \quad (3)$$

$$e_{it} = 1\{\rho_j e_{i,t-1} + \delta_{1j}x_{it} + \zeta_{1j}\eta_i + \xi_{it} > 0\}, \quad t = 2, \dots, T \quad (4)$$

où ξ_{it} et ξ_{it} , $t = 2, \dots, T$, suivent des lois normales centrées-réduites. Tous les paramètres sont indexés par l'indice du panel j .

2. LES DONNÉES

Nous utilisons les données de l'enquête Emploi de l'INSEE collectées sur la période 1990-2002. Il s'agit d'un échantillon de logements dont on mesure la relation à l'emploi des habitants de plus de 14 ans. Chaque année, environ un tiers des logements est reconstitué. Nous sommes ainsi en mesure de construire une série de 11 panels de trois ans : 1990-1992 à 2000-2002. Chaque panel rassemble environ 150 000 individus âgés de 15 ans ou plus, dans 75 000 ménages. Chaque ménage reçoit un enquêteur trois fois, au mois de mars de trois années consécutives.

La variable de salaire est le salaire mensuel déflaté par l'indice des prix. On élimine de l'échantillon les étudiants, les retraités et les indépendants. On ne conserve de plus que le sous-échantillon des hommes pour réduire l'incidence du temps partiel. Enfin, pour éliminer les valeurs abérantes, on coupe l'échantillon en-dessous du premier pourcentile et au-dessus du dernier.

L'attrition est importante : sur les 47 % de trajectoires incomplètes, 25 % des individus sortent de l'échantillon après la deuxième année et 22 % après la première année.

Le tableau 2 fournit quelques statistiques descriptives. L'échantillon est redressé pour être représentatif. On voit que l'échantillon vieillit sensiblement au

cours de la période. Le niveau d'éducation augmente aussi un peu. L'effet de la conjoncture se traduit par une augmentation de la variance des salaires en période de creux, soit 8,8 % d'augmentation en 1996 relativement à 1990 et 2000. Le taux de non-emploi s'accroît parallèlement. La mobilité salariale évolue de façon synchronisée, se réduisant lorsque l'activité se réduit⁹. Au total, en période d'embellie (deuxième moitié des années quatre-vingt-dix), l'inégalité salariale et le risque de chômage diminuent et la mobilité s'accroît; lorsque l'activité se détériore, c'est l'inverse.

3. MÉTHODE D'ESTIMATION

Un principe général de l'identification en statistique est qu'il faut une observation par aléa. Ici, nous avons η_i , l'hétérogénéité inobservée, et les chocs transitoires qu'il faut ainsi « mesurer ». Trois observations de salaires sont donc nécessaires pour espérer identifier solidement le modèle. La loi de y_{it} sachant y_{it-1} identifiera la dynamique transitoire sachant la distribution de η_i . La loi de y_{it} sachant y_{it-2} sélectionnera une distribution pour η_i parmi toutes les distributions possibles. L'enquête Emploi nous offre ces trois observations. On peut donc espérer que nos résultats d'estimation seront robustes aux choix de spécification paramétrique.

Le modèle étant paramétrique (bien que flexible), la méthode du maximum de vraisemblance s'impose pour des raisons d'efficacité. La variable manquante η_i étant discrète, le modèle est un mélange discret de lois continues. Nous choisissons l'algorithme EM (Espérance-Maximisation), pour sa simplicité. Nous en décrivons brièvement le principe.

3.1 L'algorithme EM

Soit $\{y_i, x_i, i = 1, \dots, N\}$ un échantillon i.i.d. Soit $f(y_i | x_i; \theta)$ la vraisemblance de y_i sachant x_i , fonction d'un paramètre θ . Supposons que x_i n'est pas observé (en général, une partie seulement de x_i n'est pas observée). C'est une variable discrète qui prend les valeurs $k = 1, \dots, K$, avec probabilité p_k ($p_1 + \dots + p_K = 1$). La vraisemblance marginale de y_i est

$$\ell(y_i; \theta, p_1, \dots, p_K) = \sum_{k=1}^K p_k f(y_i | x_i = k; \theta).$$

L'estimateur du maximum de vraisemblance de $(\theta, p_1, \dots, p_K)$ s'obtient par maximisation de la log-vraisemblance de l'échantillon :

$$L(\theta, p_1, \dots, p_K) = \sum_{i=1}^N \ln \ell(y_i; \theta, p_1, \dots, p_K). \quad (5)$$

9. Le tableau fournit des informations sur le taux de croissance des salaires et sur la corrélation des rangs (rho de Spearman).

TABLEAU 2

DESCRIPTION DES DONNÉES, PAR DATE ET ANNÉE D'ENTRÉE DANS LE PANEL

<i>t</i>	Par date										
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Nombre d'individus	32 749	32 402	32 855	34 629	35 689	35 674	35 780	34 927	34 945	35 309	35 376
Âge	37,9	38,0	38,1	38,2	38,4	38,6	38,8	39,0	39,1	39,2	39,3
Âge de fin d'études	17,5	17,5	17,6	17,8	17,9	18,0	18,1	18,2	18,3	18,5	18,6
Log de salaires											
• Moyenne	8,90	8,92	8,93	8,94	8,93	8,92	8,91	8,92	8,92	8,93	8,93
• Variance	0,171	0,173	0,172	0,179	0,183	0,185	0,186	0,186	0,179	0,178	0,178
Taux de non-emploi	0,098	0,095	0,107	0,123	0,141	0,130	0,135	0,141	0,134	0,137	0,116
<i>t</i>	Par date d'entrée dans le panel										
	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Nombre d'individus entrant à <i>t</i>	32 749	16 156	16 022	16 955	17 151	16 643	17 030	16 599	16 947	17 157	17 869
Nombre présents à <i>t</i> + 1	16 246	10 294	10 652	11 229	11 236	10 899	10 973	10 543	10 897	10 492	10 995
Nombre présent à <i>t</i> + 2	6 539	7 022	7 409	7 795	7 851	7 355	7 455	7 255	7 015	6 938	7 204
Âge	38,7	38,1	38,2	38,4	38,6	38,8	38,9	39,2	39,2	39,4	39,2
Âge de fin d'études	17,4	17,6	17,7	17,9	18,0	18,1	18,2	18,3	18,4	18,6	18,6

Dans notre cas, y_i est un vecteur d'observations décrivant une trajectoire individuelle de revenus. La vraisemblance conditionnelle $f(y_i | x_i; \theta)$ s'écrit donc comme un produit de probabilités conditionnelles que la transformation logarithmique de la vraisemblance ne change pas en une somme parce que la nécessité de moyenner, par rapport à la variable latente x_i , les vraisemblances individuelles place une somme entre deux produits.

Exemple : $y_i = (y_{i1}, y_{i2})$ est un couple de variables dichotomique i.i.d. suivant un probit :

$$\Pr\{y_{i1} = 1 | x_i = k\} = \Pr\{y_{i2} = 1 | x_i = k\} = \Phi(b_k).$$

Alors, $\theta = (b_1, \dots, b_K)$ et

$$f(y_i | x_i = k; \theta) = \Phi(b_k)^{y_{i1} + y_{i2}} [1 - \Phi(b_k)]^{1 - y_{i1} + 1 - y_{i2}}.$$

Alors

$$L(\theta, p_1, \dots, p_K) = \sum_{i=1}^N \ln \left[\sum_{k=1}^K p_k \Phi(b_k)^{y_{i1} + y_{i2}} [1 - \Phi(b_k)]^{1 - y_{i1} + 1 - y_{i2}} \right].$$

On constate la présence de l'opérateur $\sum_{k=1}^K$ qui empêche au logarithme de décomposer les termes multiplicatifs de $f(y_i | x_i = k; \theta)$.

L'idée de l'algorithme EM découle d'un principe important de la statistique (*selfconsistency*) inventé par Efron. Pour inventer un estimateur, imaginons que l'on connaisse les lois, et posons-nous la question : comment estimerait-on les paramètres?

Connaissant les paramètres, on peut produire une classification automatique des individus selon leur type le plus probable grâce au calcul des probabilités *a posteriori* :

$$\Pr\{x_i = k | y_i\} = \frac{f(y_i | x_i = k; \theta) p_k}{\sum_{m=1}^K f(y_i | x_i = m; \theta) p_m} = p_k(y_i) \text{ (disons).}$$

C'est l'étape E de l'algorithme EM.

Ensuite, connaissant $\Pr\{x_i = k | y_i\}$ pour tout k , il est naturel de vouloir estimer θ en maximisant la log vraisemblance conditionnelle en pondérant chaque version possible (correspondant à un type différent) par la probabilité *a posteriori* :

$$(\hat{\theta}, \hat{p}_1, \dots, \hat{p}_K) = \arg \max_{\theta, p_1, \dots, p_K} \sum_{i=1}^N \sum_{k=1}^K p_k(y_i) \ln [p_k f(y_i | x_i = k; \theta)], \tag{6}$$

sous la contrainte : $p_1 + \dots + p_K = 1$. Soit :

$$\hat{\theta} = \arg \max_{\theta} \sum_{i=1}^N \sum_{k=1}^K p_k(y_i) \ln f(y_i | x_i = k; \theta)$$

et

$$\hat{p}_k = \frac{1}{N} \sum_{i=1}^N p_k(y_i).$$

C'est l'étape M de l'algorithme EM. Il est très facile de démontrer que si on connaît exactement $p_k(y_i)$ alors maximiser cette pseudo-vraisemblance produit un estimateur asymptotiquement équivalent à l'estimateur de la vraisemblance (5).

Exemple (suite) : Dans l'exemple précédent, cela donne :

$$\ln f(y_i | x_i = k; \theta) = \sum_{t=1}^2 \{y_{it} \ln \Phi(b_k) + (1 - y_{it}) \ln [1 - \Phi(b_k)]\}.$$

Dans ce cas, c'est très simple. On estimera les paramètres en estimant K modèles Probit, un pour chaque type, en pondérant les observations par les probabilités *a posteriori*¹⁰.

Considérons maintenant une version simplifiée de notre modèle :

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

où $\eta_i = \sigma \underline{\eta}(x_i)$, avec $\underline{\eta}(k) = \underline{\eta}_k$. Oublions la possibilité de se retrouver sans emploi. Alors,

$$f(y_{i1}, y_{i2} | \eta_i; \theta) = f(y_{i1} - \eta_i; \alpha) f(y_{i2} - \eta_i; \alpha) c(F(y_{i1} - \eta_i; \alpha), F(y_{i2} - \eta_i; \alpha); \beta)$$

où α rassemble les paramètres de la densité marginale de ε_{it} (c'est-à-dire $f(y - \eta; \alpha)$) et β ceux de la copule¹¹. On voit que α apparaît à la fois dans la densité marginale de ε_{it} et dans la copule à travers les rangs. Un argument développé par Arcidiacono et Jones (2003) permet de simplifier l'étape M en procédant séquentiellement de la façon suivante.

D'abord, on maximise par rapport à α et σ la pseudo-vraisemblance marginale :

$$(\hat{\alpha}, \hat{\sigma}) = \arg \max_{\alpha, \sigma} \sum_{i=1}^N \sum_{t=1}^2 \sum_{k=1}^M p_k(y_i) \ln f(y_{it} - \sigma \underline{\eta}_k; \alpha). \tag{7}$$

Pratiquement, nous utilisons le principe du traitement des variables discrètes latentes de l'algorithme EM aussi pour simplifier l'estimation des paramètres de la densité marginale de ε_{it} . Nous l'avons spécifiée comme un mélange de lois normales. La loi marginale peut donc se voir comme un premier tirage du numéro de la composante dans le mélange, puis un tirage dans la loi de cette composante. L'étape (7) se réduit alors à une régression pondérée très simple, la pondération

10. Sans variables explicatives observées, on a même une solution analytique.

11. Avec trois observations, on aurait :

$$f(y_{i1}, y_{i2}, y_{i3} | x_i; \theta) = f(y_{i1} - \eta_i | x_i; \alpha) f(y_{i2} - \eta_i | x_i; \alpha) f(y_{i3} - \eta_i | x_i; \alpha) \\ \times c(F(y_{i1} - \eta_i; \alpha), F(y_{i2} - \eta_i; \alpha) | x_i; \beta) c(F(y_{i2} - \eta_i; \alpha), F(y_{i3} - \eta_i; \alpha) | x_i; \beta).$$

étant égale au produit de la probabilité *a posteriori* du type déterminant η_i et de la probabilité du numéro de la composante du mélange.

Enfin, dans un second temps, on maximise par rapport à β la pseudo-vraisemblance des rangs en utilisant les valeurs précédemment obtenues pour α et σ ($\hat{\alpha}$ et $\hat{\beta}$) :

$$\hat{\beta} = \arg \max_{\beta} \sum_{i=1}^N \sum_{k=1}^M p_k(y_i) \ln c(F(y_{i1} - \hat{\sigma}\underline{\eta}_k; \hat{\alpha}), F(y_{i2} - \hat{\sigma}\underline{\eta}_k; \hat{\alpha}); \beta). \quad (8)$$

L'algorithme EM consiste à itérer les étapes E et M qu'on vient de décrire dans les grandes lignes. Un exposé précis de l'algorithme d'estimation figure à l'annexe.

4. RÉSULTATS D'ESTIMATION

4.1 Hétérogénéité inobservée

Le graphique 1 montre la distribution η_i estimée pour différentes valeurs de K , le nombre supposé de points de support. Lorsqu'on augmente K , on voit une convergence vers une distribution normale. Ce résultat rejoint celui précédemment obtenu par Horowitz and Markatou (1996) sur données américaines. C'est bien ce à quoi il faut s'attendre si l'effet fixe individuel est en fait la somme d'un grand nombre de traits individuels indépendants (théorème limite-centrale).

4.2 Effet des variables exogènes

Nous reportons dans le tableau 3 les résultats d'estimation des paramètres des variables exogènes (x_{it}) de l'équation de salaire et des probabilités de non-emploi.

La moyenne des salaires est concave dans l'expérience (l'âge), le maximum étant atteint après 31 ans d'activité. Le paramètre de la copule croît avec l'expérience. Partant, la mobilité salariale décroît donc avec l'âge. Le risque de non-emploi diminue avec l'expérience jusqu'à 25 ans puis augmente à nouveau. L'inactivité augmente en effet significativement passé 50 ans. Enfin, la volatilité des salaires diminue d'abord, puis augmente avec l'ancienneté après huit années d'activité. L'éducation et l'effet fixe η_i augmentent la moyenne et la variance des salaires.

Le graphique 2 représente la densité de la loi marginale de ε_{it} . Cette distribution est fortement *non gaussienne*, avec des queues épaisses (kurtosis de 7,5). Cette *non-gaussianité* du choc transitoire contraste avec la *gaussianité* de l'effet individuel.

4.3 Décomposition de la variance des salaires

Le graphique 3 décrit l'évolution des trois composantes de la variance en coupe des logs des salaires : la variance de la composante déterministe, $\text{var}(\beta_j x_{it})$ la variance permanente, σ_j^2 , et la variance transitoire, $\exp(2\alpha_j x_{it})$ (le carré de $\exp(\alpha_j x_{it})$).

GRAPHIQUE 1

DISTRIBUTION DE η_1 POUR DIFFÉRENTES VALEURS DE K , LE NOMBRE DE POINTS DE SUPPORT

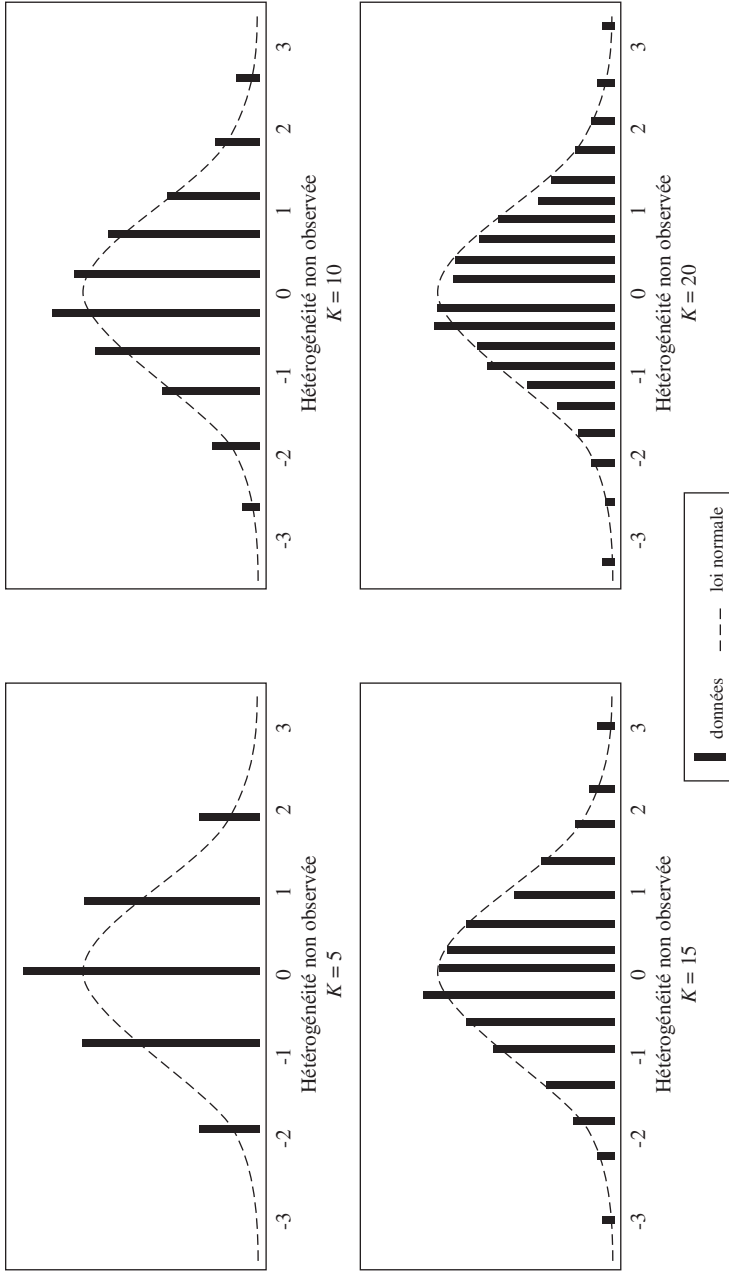


TABLEAU 3
PARAMÈTRES ESTIMÉS POUR LE PANEL 1990-1992

	Log de salaires			Probabilité de non-emploi	
	Moyenne (β , σ)	Volatilité (α)	Mobilité (γ)	Initiale (δ_0 , ζ_0)	Transition (δ , ρ , ζ_1)
Expérience	0,03776 (0,00112)	- 0,01172 (0,0191)	0,06154 (0,00713)	- 0,1104 (0,00351)	- 0,05234 (0,00538)
(100 ×) expérience ²	- 0,06113 (0,00237)	0,07065 (0,0380)	- 0,04738 (0,0144)	0,2154 (0,00751)	0,1049 (0,0110)
Éducation II	0,1850 (0,00808)	0,1206 (0,187)	0,1026 (0,0462)	- 0,5928 (0,0259)	- 0,2576 (0,0347)
Éducation III	0,4269 (0,0104)	0,7416 (0,194)	0,7007 (0,0586)	- 0,8137 (0,0374)	- 0,4092 (0,0518)
Éducation IV	0,5869 (0,0137)	0,5230 (0,264)	0,6368 (0,0766)	- 1,282 (0,0636)	- 0,4968 (0,0770)
Éducation V	0,8952 (0,0134)	1,086 (0,206)	0,9898 (0,0747)	- 1,315 (0,0633)	- 0,6853 (0,0863)
η_i	0,2304 (0,00321)	- -	- -	- 0,7775 (0,0137)	- 0,3334 (0,0167)
$e_{i,t-1}$	- -	- -	- -	- -	1,744 (0,0384)
Constante	8,173 (0,0130)	- 3,486 (0,266)	1,107 (0,0869)	- 0,1165 (0,0354)	- 1,115 (0,0612)

Ces trois contributions à la variance sont sensiblement égales. La contribution de l'éducation et de l'expérience augmente en 1995-1997. C'est dû pour moitié à un changement de composition de l'échantillon et pour une autre moitié à un accroissement du rendement de l'expérience; le rendement de l'éducation diminuant faiblement au cours des années 1990 (voir tableau 4 et graphique 4). L'effet le plus intéressant est sans doute l'opposition des évolutions des variances permanentes et transitoires. La variance permanente augmente quand la conjoncture se détériore et diminue quand elle s'améliore, alors que c'est l'inverse pour la variance transitoire.

5. AJUSTEMENT

Nous évaluons maintenant la capacité du modèle à reproduire quelques caractéristiques essentielles des données.

GRAPHIQUE 2

DENSITÉ MARGINALE DE ϵ_{it} POUR LE PANEL 1990-1992

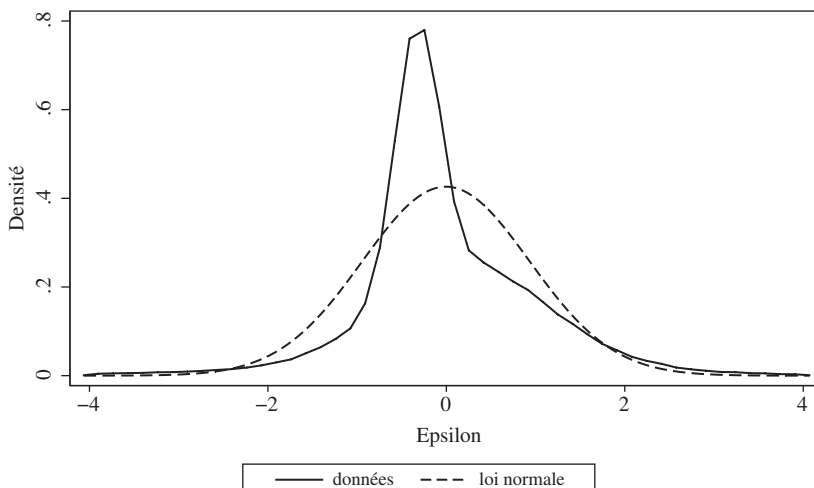


TABLEAU 4

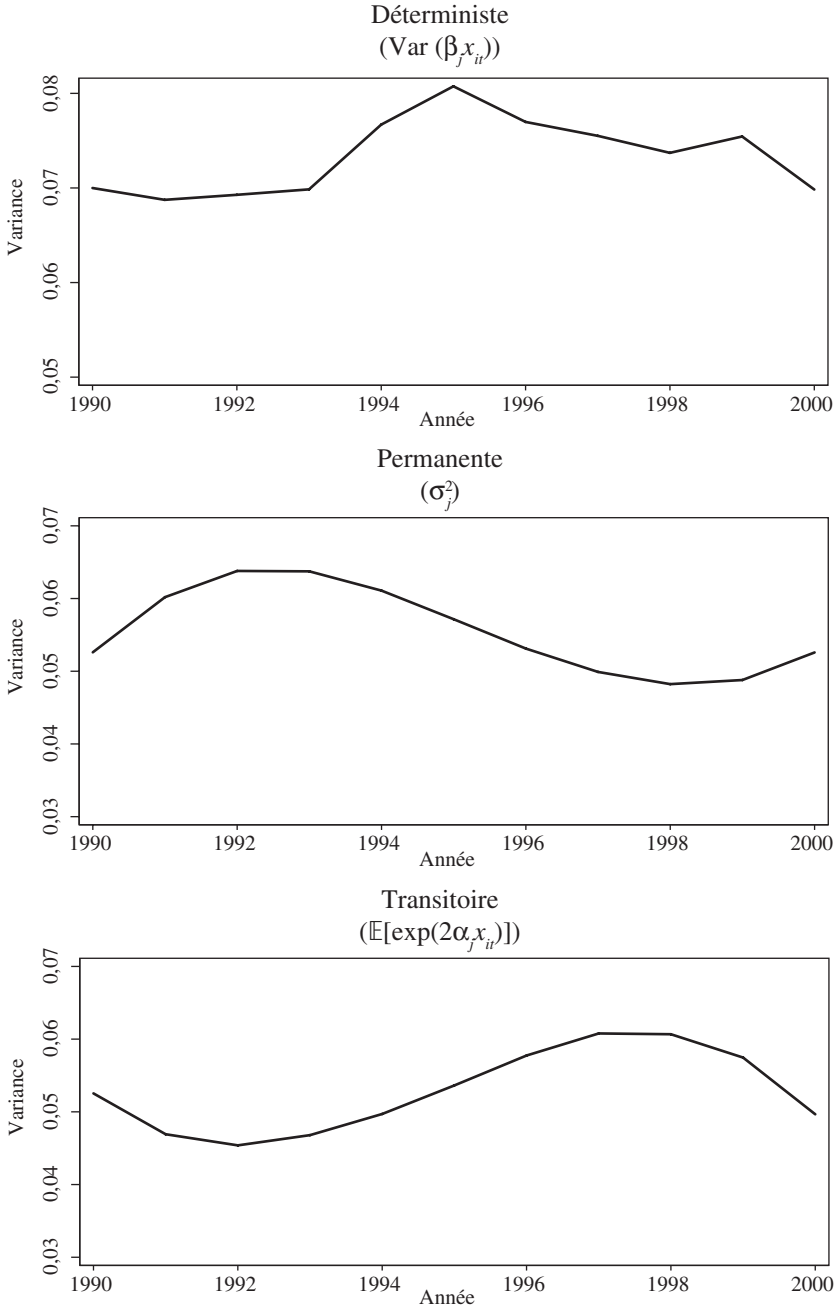
DÉCOMPOSITION DE OAXACA DE LA COMPOSANTE DÉTERMINISTE DE LA VARIANCE DES SALAIRES (référence = 1990-1992)

	1990	1993	1995	1997	2000
Effet de composition	0,0700	0,0704	0,0741	0,0734	0,0742
+ rendements de l'éducation	0,0700	0,0700	0,0733	0,0696	0,0666
+ rendements de l'expérience	0,0700	0,0699	0,0807	0,0755	0,0698

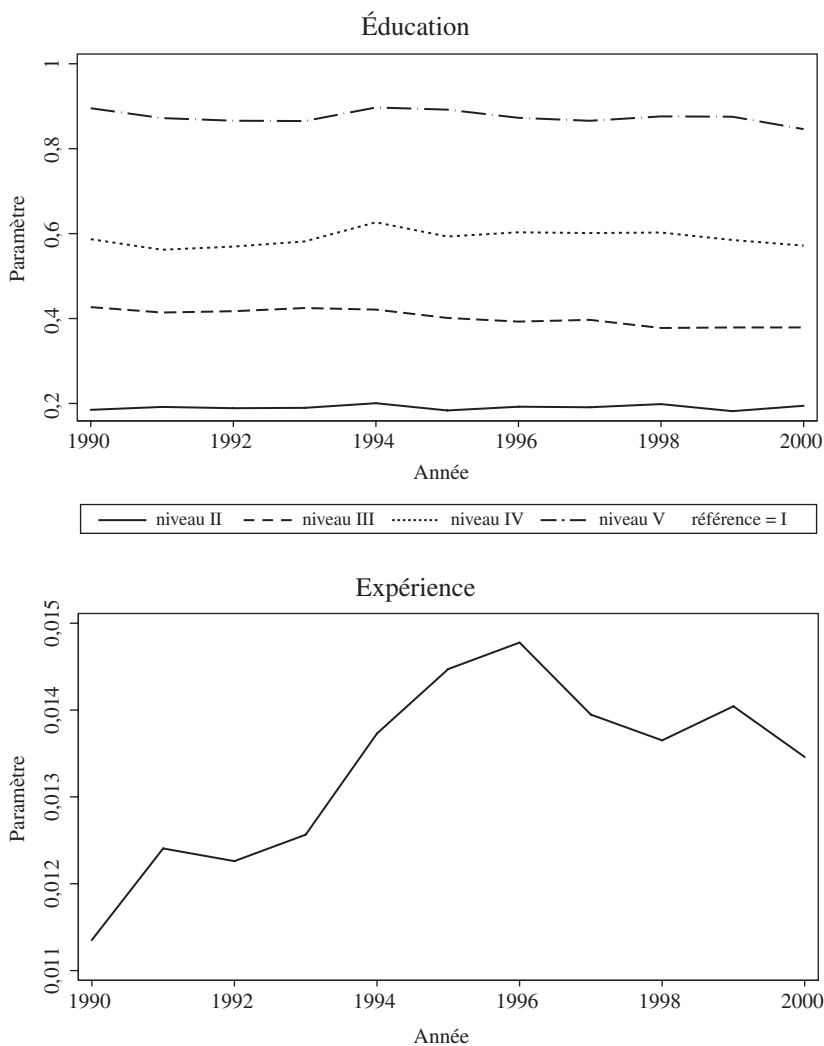
Les graphiques 5 et 6 montrent les distributions marginales pour 1990-1992, observées et prédites par le modèle, de certaines variables d'intérêt construites à partir des logs de salaires y_{it} et des résidus $u_{it} = y_{it} - \beta_j x_{it}$ (premières et secondes différences et valeurs moyennes sur trois ans). L'ajustement est bon en général, bien qu'un peu moins bon lorsqu'il s'agit de reproduire la dynamique à deux ans. Puis nous considérons les probabilités de transition entre quintiles de salaires entre t et $t + 1$ et entre t et $t + 2$ (tableau 5). Ici encore, l'ajustement est bon, bien que le modèle montre une certaine tendance à surestimer la mobilité dans la partie supérieure de la distribution. Enfin, nous vérifions que l'ajustement ne se détériore pas à certaines dates plus qu'à d'autres. Le graphique 7 montre ainsi l'évolution de quelques moments importants tels qu'on les observe et tels qu'on les prédit.

GRAPHIQUE 3

DÉCOMPOSITION DE LA VARIANCE DES LOGS DE SALAIRES



GRAPHIQUE 4
RENDEMENTS DE L'ÉDUCATION ET DE L'EXPÉRIENCE

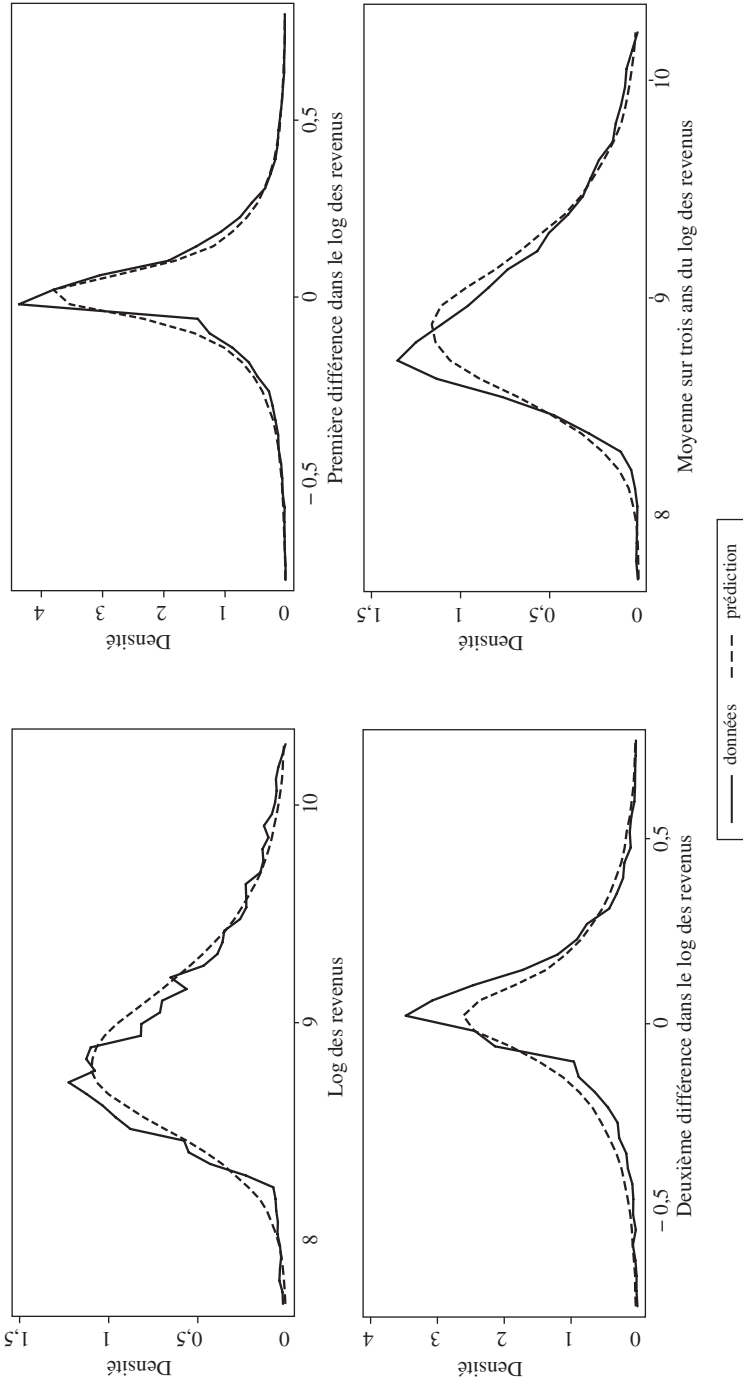


Il apparaît que le qualité du modèle ne change pas significativement d'une date à l'autre.

En résumé, le modèle décrit correctement les données à la fois dans la dimension de coupe et dans la dimension temporelle. On note cependant une légère détérioration de l'ajustement lorsqu'on cherche à prédire la dynamique à deux ans. En particulier, le modèle surestime un peu la mobilité malgré la présence d'hétérogénéité inobservée (effet mobile-stable contrôlé).

GRAPHIQUE 5

AJUSTEMENT DES DISTRIBUTIONS MARGINALES DES NIVEAUX DE SALAIRES (1990-1992)



GRAPHIQUE 6

AJUSTEMENT DES DISTRIBUTIONS MARGINALES DES RÉSIDUS (1990-1992)

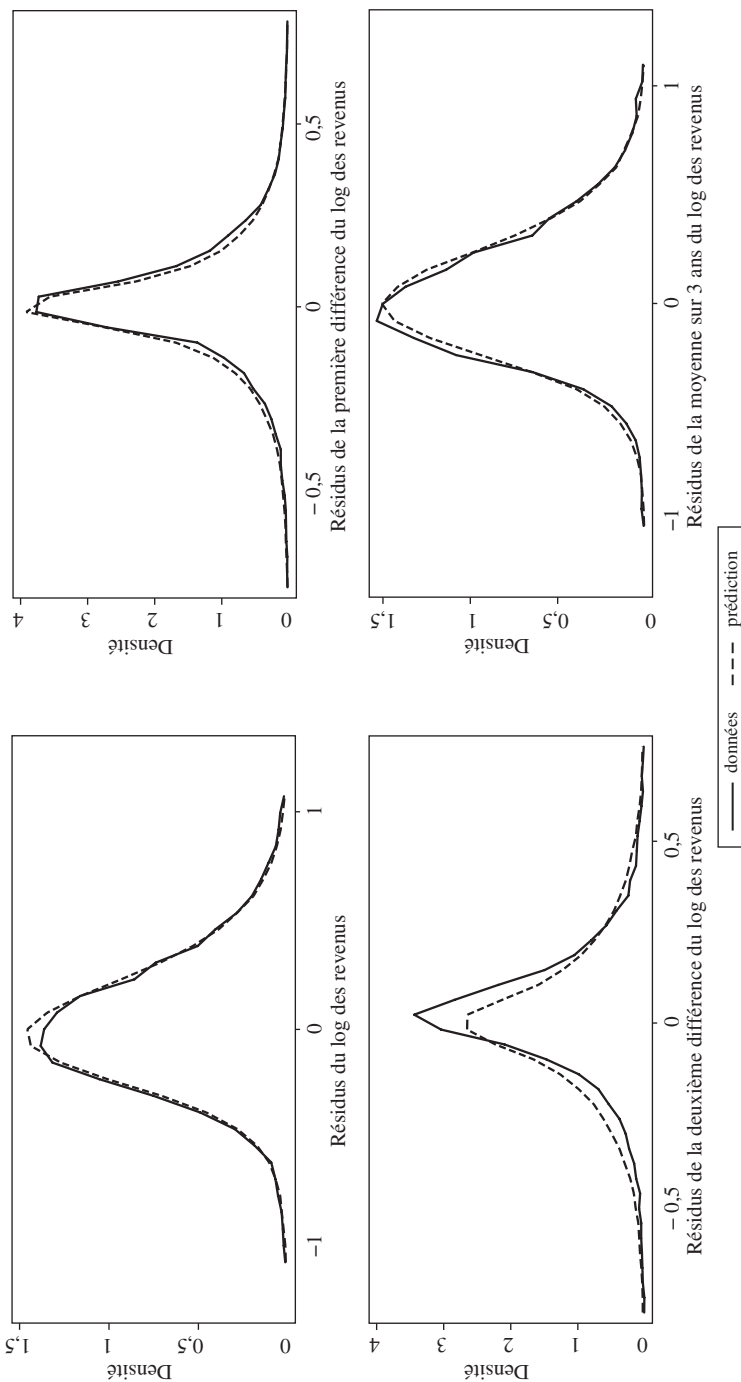


TABLEAU 5
AJUSTEMENT DES MATRICES DE TRANSITION (1990-1992)

	Observée					Prédite				
1990/ 1991	0,70	0,20	0,07	0,02	0,01	0,76	0,18	0,04	0,01	0,00
	0,21	0,54	0,19	0,04	0,01	0,19	0,56	0,19	0,05	0,01
	0,06	0,22	0,53	0,17	0,02	0,04	0,20	0,53	0,19	0,03
	0,02	0,03	0,18	0,63	0,14	0,01	0,05	0,20	0,57	0,16
	0,01	0,01	0,02	0,14	0,82	0,00	0,01	0,03	0,16	0,79
1990/ 1992	0,67	0,23	0,08	0,02	0,01	0,68	0,22	0,07	0,03	0,01
	0,23	0,49	0,22	0,04	0,01	0,22	0,46	0,22	0,08	0,02
	0,07	0,23	0,50	0,18	0,03	0,07	0,22	0,43	0,22	0,06
	0,02	0,04	0,19	0,59	0,16	0,02	0,08	0,23	0,47	0,20
	0,01	0,01	0,02	0,16	0,80	0,01	0,02	0,05	0,20	0,72

6. INÉGALITÉ DE LONG TERME

À l'aide du modèle nous pouvons poursuivre les trajectoires individuelles et calculer des revenus permanents sur les données simulées. L'algorithme EM donne comme sous-produit une estimation des probabilités *a posteriori* du type latent, $p_k(y_i)$. Pour chaque individu i de chaque panel j nous prédisons $\eta_i = \underline{\eta}_k$ pour la valeur de k telle que $p_k(y_i) = \max\{p_k(y_i), \dots, p_k(y_i)\}$. Sachant η_i , le modèle n'est guère différent d'un modèle autorégressif et il est aisé de prédire y_{it} au delà de l'horizon des données ($t_j + T$). En pratique, nous simulerons des trajectoires sur divers horizons de un an à l'âge de la retraite (arbitrairement fixé à 65 ans). Lorsque le modèle prédit un passage par le chômage, un revenu de remplacement de 60 % du salaire précédent est affecté à l'individu¹².

Pour un horizon de simulation H donné, nous calculons un revenu permanent égal à l'annuité équivalente au flux des revenus des H périodes futures pour un taux d'actualisation β :

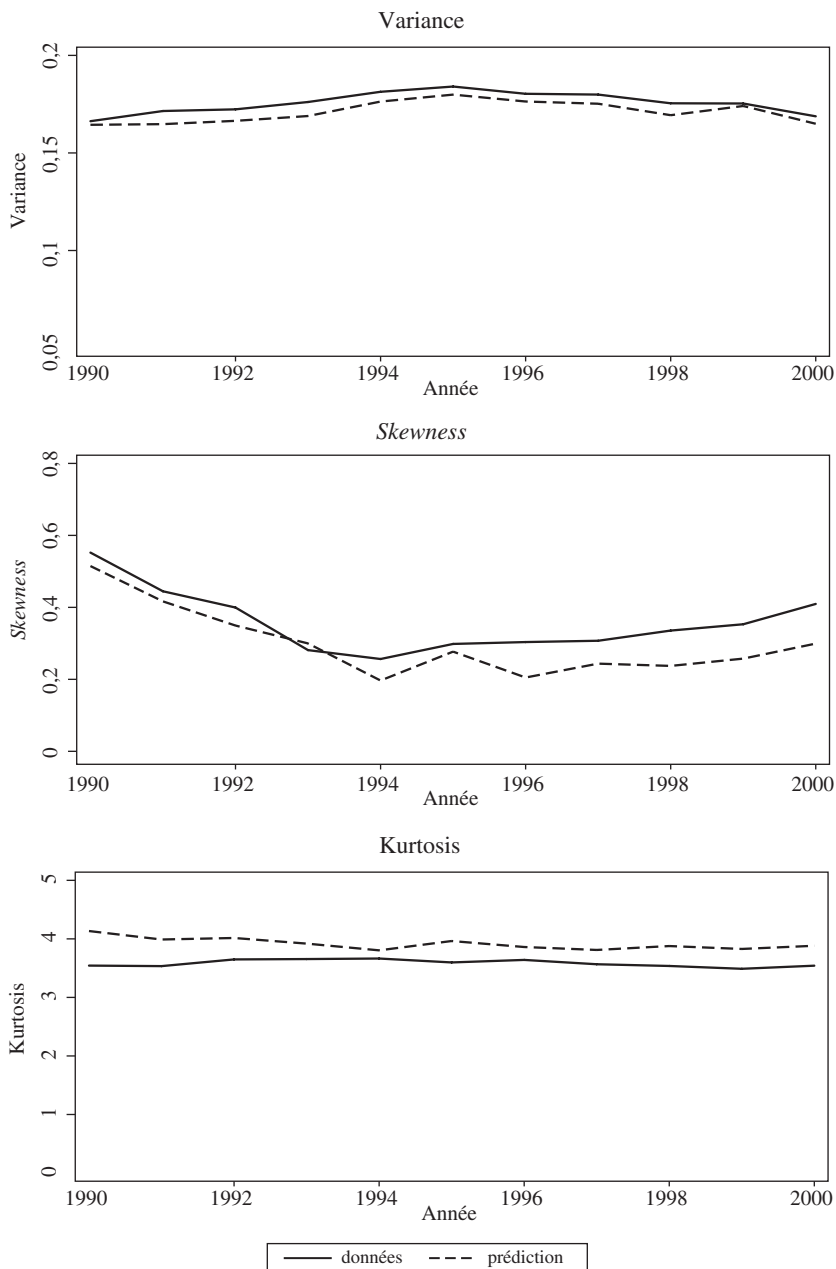
$$Y_{it}^P = \frac{Y_{it} + \beta Y_{i,t+1} + \dots + \beta^{H_{it}-1} Y_{i,t+H_{it}-1}}{1 + \beta + \dots + \beta^{H_{it}-1}}$$

où Y_{it}^P est une notation pour le niveau de revenu perçu à la période t (salaire ou revenu de remplacement), $H_{it} = \min\{H, 65 - \text{Age}_{it}\}$ et $\beta = 0,95$. Le calcul d'annuités permet de comparer des portefeuilles de maturités différentes.

12. 60 % est le taux de remplacement moyen sur la période. Voir Martin (1996).

GRAPHIQUE 7
 AJUSTEMENT, ÉVOLUTION SUR 1990-2000

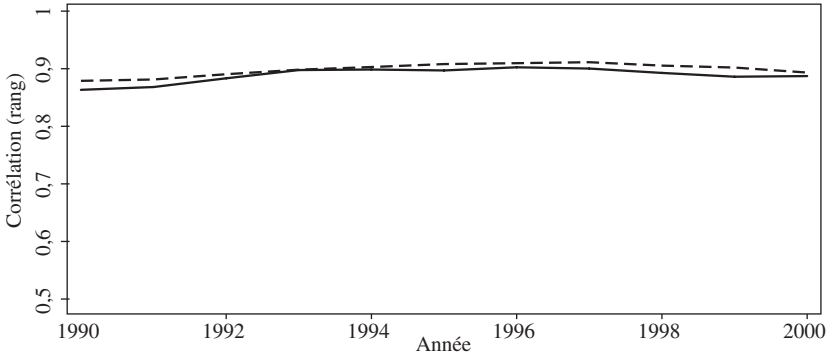
Distributions marginales



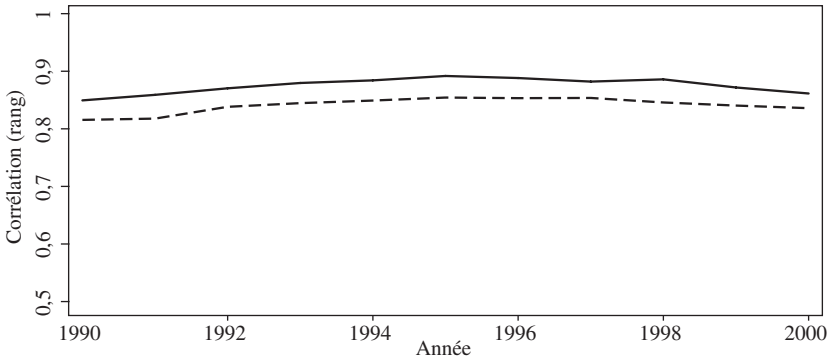
GRAPHIQUE 7 (suite)

Mobilité des salaires

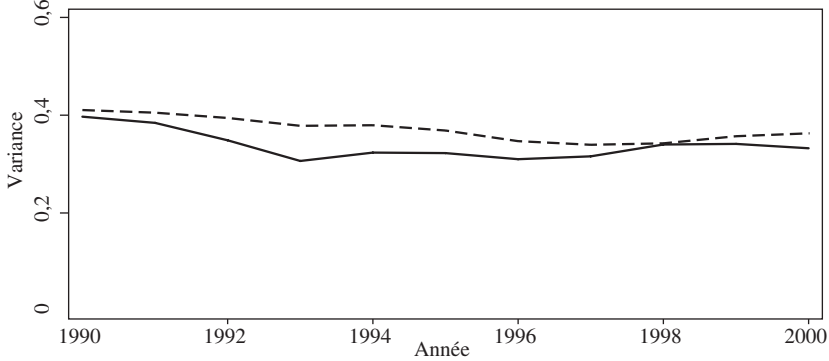
Spearman, 1 an



Spearman, 2 ans

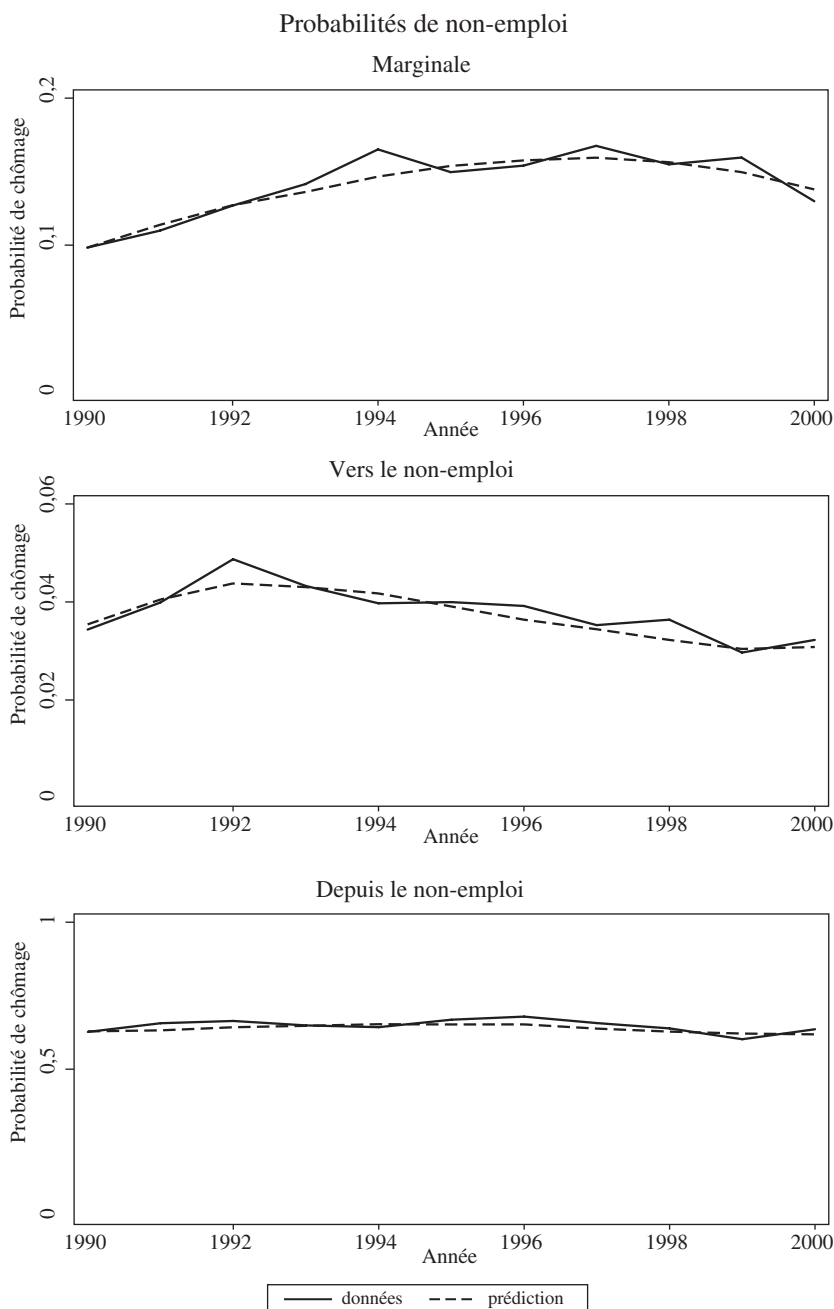


Variance des premières différences



— données - - - prédiction

GRAPHIQUE 7 (suite)



Le tableau 6 reproduit différents indices d'inégalités de long terme, la variance de $y_{iH}^p = \ln Y_{iH}^p$, le Gini, le rapport interdécile $D9/D1$, pour les maturités $H = 1, 5, 10, \infty$. L'objectif des paragraphes qui suivent est de commenter ces résultats, si possible à l'aide de graphiques.

6.1 L'effet du chômage sur les inégalités

Nous commençons par étudier l'impact du chômage sur les inégalités. Le graphique 8 vise à comparer les inégalités de revenus calculées dans la population des employés à l'inégalité de revenu dans la population totale en t_j et ceci pour chaque panel j . Plus précisément, le graphique 8 décrit l'évolution du rapport de la variance de y_{iH}^p dans la population des employés à la variance dans la population totale, pour $H = 1, 5, 10, \infty$.

Comme on pouvait s'y attendre, la différence entre les deux échantillons s'amointrit avec l'éloignement de l'horizon, la condition de chômeur étant une condition transitoire. L'écart ne se réduit jamais complètement, d'un part à cause

TABLEAU 6

INÉGALITÉ DE LONG TERME POUR DIFFÉRENTES MATURITÉS, 1990-2000
(taux de remplacement au chômage = 60 %, taux d'escompte = 5 %)

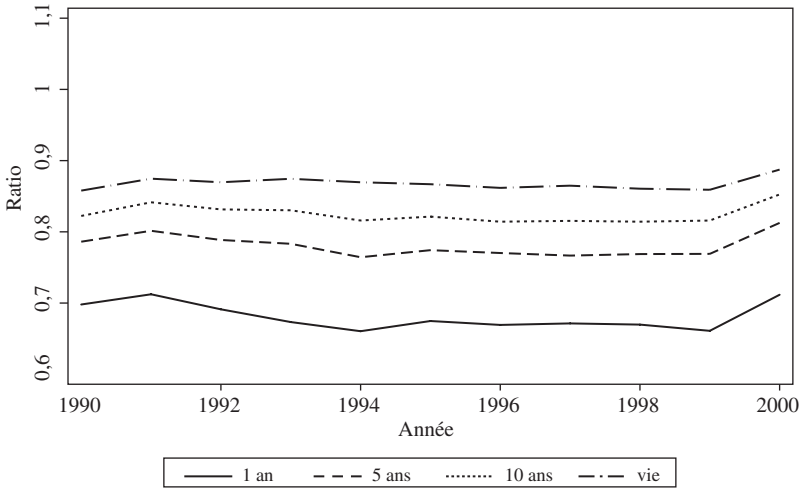
Ma- turité	Échantillon complet				Employés seulement			
	1 an	5 ans	10 ans	Vie	1 an	5 ans	10 ans	Vie
	Variance							
1990	0,2449	0,2012	0,1858	0,1753	0,1708	0,1594	0,1539	0,1512
1991	0,2435	0,2036	0,1869	0,1765	0,1728	0,1626	0,1569	0,1544
1992	0,2494	0,2064	0,1902	0,1791	0,1718	0,1622	0,1576	0,1552
1993	0,2621	0,2170	0,1991	0,1864	0,1792	0,1717	0,1670	0,1647
1994	0,2758	0,2259	0,2072	0,1943	0,1830	0,1745	0,1705	0,1694
1995	0,2720	0,2276	0,2085	0,1958	0,1849	0,1778	0,1717	0,1694
1996	0,2729	0,2258	0,2054	0,1913	0,1856	0,1758	0,1703	0,1677
1997	0,2748	0,2249	0,2048	0,1907	0,1858	0,1750	0,1690	0,1661
1998	0,2647	0,2180	0,1995	0,1877	0,1785	0,1702	0,1642	0,1622
1999	0,2659	0,2151	0,1954	0,1830	0,1778	0,1663	0,1600	0,1579
2000	0,2511	0,2060	0,1886	0,1784	0,1778	0,1655	0,1596	0,1575

TABLEAU 6 (suite)

Ma- turi-té	Échantillon complet				Employés seulement			
	1 an	5 ans	10 ans	Vie	1 an	5 ans	10 ans	Vie
	Gini							
1990	0,2689	0,2490	0,2416	0,2367	0,2409	0,2317	0,2275	0,2255
1991	0,2675	0,2499	0,2421	0,2374	0,2410	0,2332	0,2290	0,2273
1992	0,2686	0,2501	0,2429	0,2380	0,2390	0,2317	0,2285	0,2270
1993	0,2740	0,2558	0,2485	0,2433	0,2419	0,2369	0,2345	0,2333
1994	0,2792	0,2598	0,2526	0,2476	0,2435	0,2387	0,2368	0,2365
1995	0,2774	0,2603	0,2522	0,2470	0,2440	0,2400	0,2365	0,2352
1996	0,2796	0,2599	0,2514	0,2452	0,2453	0,2389	0,2358	0,2341
1997	0,2803	0,2594	0,2506	0,2444	0,2453	0,2388	0,2350	0,2332
1998	0,2755	0,2557	0,2472	0,2421	0,2415	0,2356	0,2317	0,2306
1999	0,2757	0,2538	0,2448	0,2389	0,2410	0,2335	0,2292	0,2274
2000	0,2708	0,2506	0,2425	0,2373	0,2424	0,2339	0,2298	0,2277
	Ratio interdécile, D9/D1							
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

GRAPHIQUE 8

EFFET DU CHÔMAGE SUR LES INÉGALITÉS



du taux d'escompte et d'autre part parce que l'hétérogénéité inobservée détermine le risque de chômage, si bien que la population des employés en t n'est pas représentative.

6.2 Structure de l'inégalité de long terme

Nous décomposons la variance du log des annuités en la somme de quatre termes. Une première composante, dite déterministe, représente l'effet des variables exogènes (éducation et expérience). On la calcule comme la variance des moyennes de y_{iH}^P pour un niveau d'éducation et d'expérience donné (variance interéducation et expérience). Le deuxième effet, dit permanent, capture les différences dues à l'hétérogénéité inobservée. On le calcule comme la moyenne des variances pour un niveau d'éducation et d'expérience donné (variance intra-éducation et expérience) de y_{iH}^P simulé en annulant les chocs transitoires. Les deux dernières composantes mesurent la contribution des chocs transitoires. Un premier terme, dit simplement composante transitoire, est la variance intra aux groupes définis par le croisement de l'éducation, l'expérience et le type d'hétérogénéité inobservée (k), de y_{iH}^P simulé avec chocs transitoires i.i.d. (il suffit de forcer la densité-copule à être égale à 1 et de tirer indépendamment à chaque date dans la loi marginale). Enfin, une dernière composante, dite persistante, est la variance résiduelle :

$$\begin{aligned} \text{variance totale} = & \text{variance interéducation, expérience} \\ & + \text{variance intra-éducation et expérience, aucun choc} \\ & \text{transitoire} \end{aligned}$$

- + variance intra-éducation, expérience et k , choc i.i.d.
- + variance résiduelle

Le graphique 9 montre les résultats de cette décomposition pour l'année de base 1990 et pour différentes maturités. Les composantes déterministe et permanente expliquent approximativement, l'une et l'autre, 40 % de la variance totale. Il y a peu d'effet de la maturité. Il en va différemment des composantes transitoire et résiduelle. Initialement, la composante transitoire prend tous les 20 % restant. Puis, la dynamique des rangs prend peu à peu de l'ampleur jusqu'à expliquer 10 % de la variance totale.

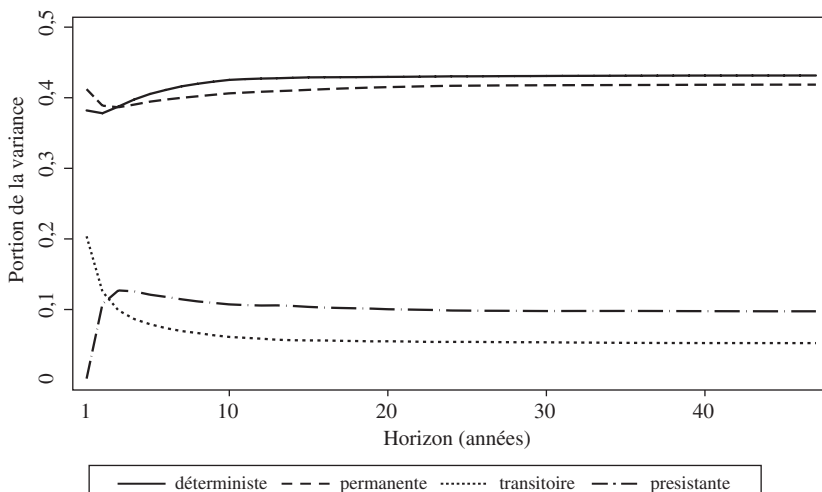
La raison pour laquelle la variance transitoire ne disparaît jamais tout à fait tient au choix d'un taux d'escompte de 5 % plutôt que 0 %. Si nous avions calculé des moyennes plutôt que des annuités, la mémoire de la condition initiale tendrait vers 0 avec H (poids de $1/H$) alors qu'avec $\beta < 1$, lorsque H tend vers l'infini, le poids de Y_{it} dans Y_{it}^p tend vers $1 - \beta$.

Nous voyons donc qu'il reste une source d'inertie des revenus dans la dynamique des rangs qui n'est pas entièrement capturée par l'hétérogénéité inobservée, comptant pour un quart de la contribution de l'hétérogénéité inobservée à l'inégalité de long terme.

D'autre part, il apparaît que la structure de l'hétérogénéité inobservée reste à peu près stable après cinq ans (Cf. note 2). On notera cependant que le niveau de l'inégalité de long terme continue de décroître avec la maturité au delà de 5 ans : 82 % à 5 ans (en rapport de l'inégalité de coupe), 76 % à 10 ans et 72 % pour tout le cycle de vie (Cf. tableau 6).

GRAPHIQUE 9

VARIANCE DE LONG TERME SELON LA MATURITÉ EN 1990



6.3 *Évolution, 1990-2000*

Le graphique 10 illustre les résultats relatifs à l'évolution des inégalités de long terme au cours des années 1990, pour l'échantillon des employés seulement à une date donnée et pour l'échantillon total des employés et des chômeurs. La partie gauche décrit l'évolution de la variance de y_{itH}^P pour différents horizons; la partie droite l'évolution du ratio $\text{var } y_{itH}^P / \text{var } y_{it1}^P$. Ce ratio peut s'interpréter comme un indice d'immobilité au sens de Shorrocks. Si les individus échangent beaucoup leurs positions dans les distributions marginales, alors l'inégalité de long terme sera faible. Plus l'inégalité de long terme est faible relativement à l'inégalité de coupe et plus la mobilité est dite égalisatrice (Shorrocks, 1978; Fields, 2005).

Nous voyons qu'en effet la mobilité est égalisatrice. Lorsqu'on simule des trajectoires individuelles jusqu'à 65 ans ($H = \infty$), on obtient un ratio $\text{var } y_{itH}^P / \text{var } y_{it1}^P$ de 70 %. Le graphique 11 expérimente une variation du taux de remplacement servant à calculer le revenu des chômeurs. Même avec un taux de 80 %, la réduction d'inégalité sur le long terme est encore très significative (ratio d'à peu près 75 %). Pour un ratio de remplacement de 40 % l'indice d'immobilité tombe à 60 %. La perspective de sortir du chômage est donc d'autant plus égalisatrice que l'écart de revenu entre les employés et les chômeurs est important.

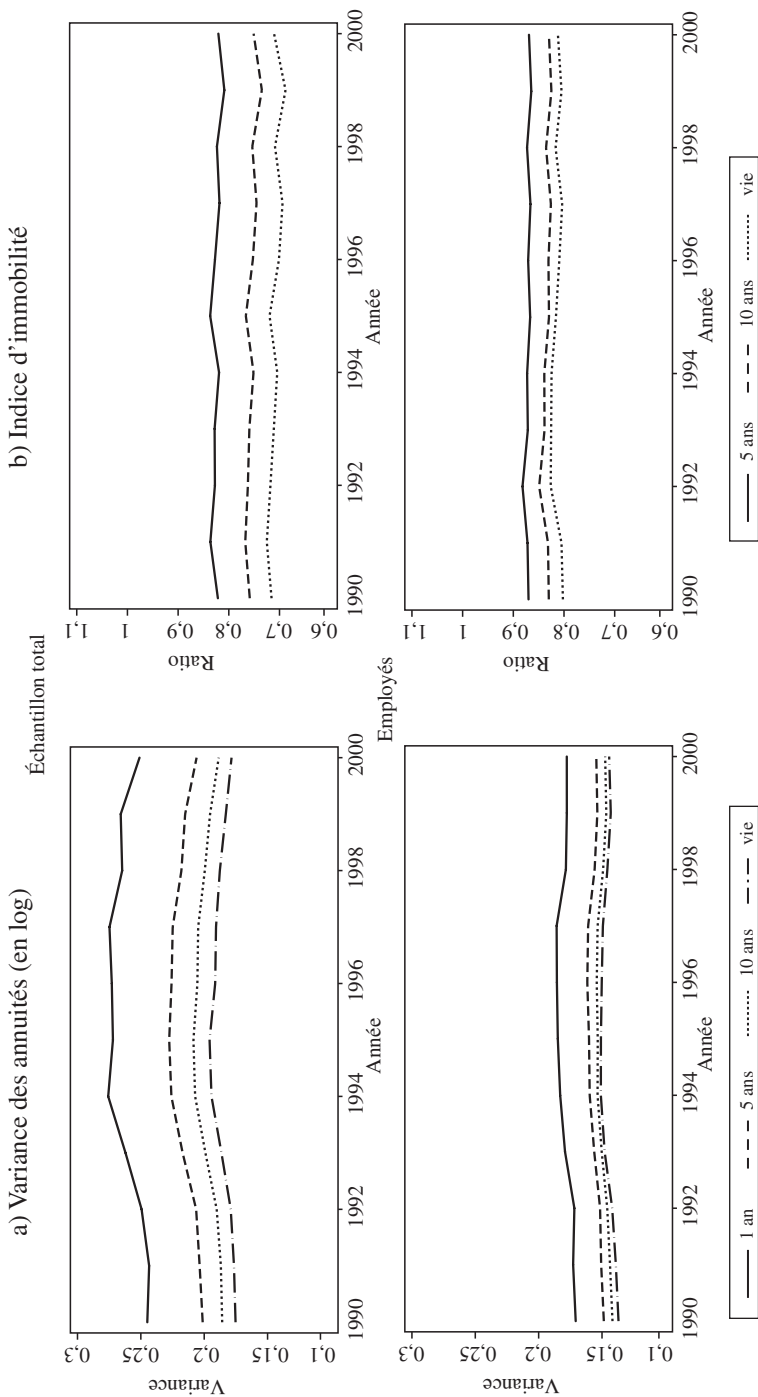
Considérons maintenant un échantillon d'employés. La réduction de l'indice d'inégalité de long terme vis-à-vis de l'indice de coupe est beaucoup plus forte pour l'échantillon total que pour l'échantillon des employés. Le ratio $\text{var } y_{itH}^P / \text{var } y_{it1}^P$ ($H = \infty$) fluctue ainsi autour de 90 % pour l'échantillon des employés. Pourtant, l'expérience qui est faite ici consiste à laisser vieillir une cohorte de travailleurs employés à une date donnée. Parmi ceux-ci, certains sont déjà âgés et leur salaire ne progressera plus que très peu. Les plus jeunes peuvent quant à eux espérer des augmentations de salaires. On peut donc trouver la réduction de long terme des inégalités un peu faible.

Toutefois, ce que mesure l'indice d'immobilité $\text{var } y_{itH}^P / \text{var } y_{it1}^P$ est l'effet total de la mobilité sur les inégalités. Or, les travailleurs font face à deux types de mobilité : une mobilité salariale qui les fait échanger des positions dans l'échelle des salaires, égalisatrice, et une mobilité emploi/chômage, fortement génératrice d'inégalités. Pour se faire une idée de la contribution respective des deux mobilités dans l'effet résultant, il suffit de comparer l'indice d'immobilité calculé pour les employés dans le graphique 10 et l'indice calculé dans l'expérience contre-factuelle où l'on élimine complètement le risque de chômage (graphique 12). Au lieu de 90 % on obtient un indice de 80 %. On peut donc dire que la mobilité salariale tend à réduire les inégalités de 20 % mais le risque de chômage limite cette réduction à 10 %.

On notera enfin que l'indice d'immobilité des employés répond un peu au cycle. Il est maximal en 1993-1994, au moment où le risque de chômage est le plus élevé. C'est à ce moment là que la mobilité est la moins égalisatrice.

GRAPHIQUE 10

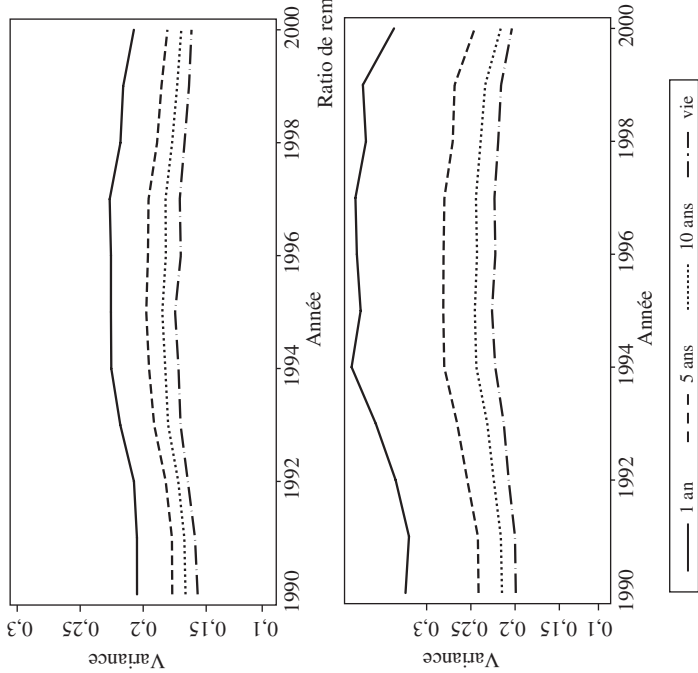
ÉVOLUTION DES INÉGALITÉS DE LONG TERME (taux de remplacement = 60 %)



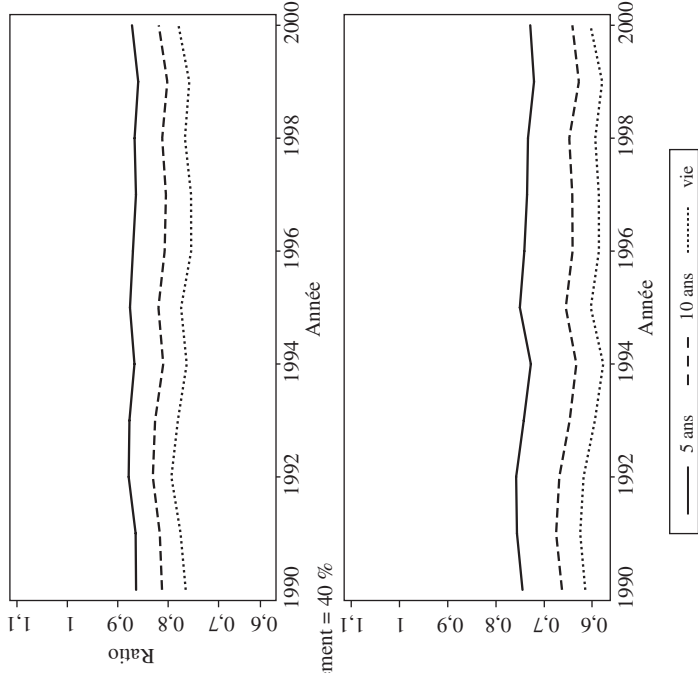
GRAPHIQUE 11

ÉVOLUTION DES INÉGALITÉS DE LONG TERME POUR DIFFÉRENTS RATIOS DE REMPLACEMENT (ÉCHANTILLON TOTAL)

a) Variance des annuités (en log)

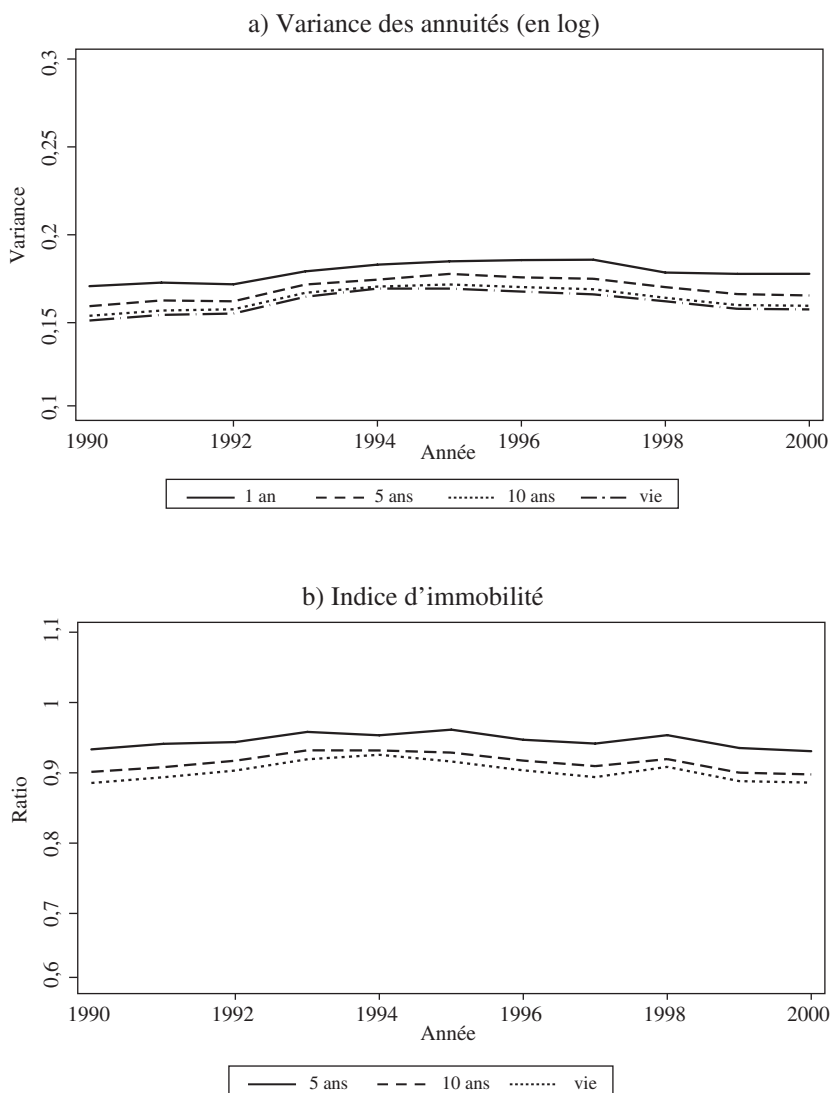


b) Indice d'immobilité



GRAPHIQUE 12

ÉVOLUTION DES INÉGALITÉS DE LONG TERME SANS RISQUE DE CHÔMAGE



CONCLUSION

Dans cette étude, nous proposons un modèle de la dynamique salariale adapté à une estimation à partir de panels courts comme l'enquête Emploi. nous utilisons le modèle pour simuler des trajectoires individuelles de salaires au-delà de la période d'enquête et calculer des revenus permanents. Nous mesurons le

rapport entre l'inégalité de revenus permanents (inégalité de long terme) et l'inégalité salariale de coupe. Ce rapport est inférieur à un, preuve que la mobilité des revenus est égalisatrice. Cependant, nous constatons le rôle essentiel joué par le risque de chômage dans cette mesure. La mobilité réduit les inégalités sur le long terme dans un échantillon représentatif de travailleurs employés ou au chômage, en grande partie parce que le chômage ne dure pas éternellement. À l'inverse, le risque de chômage est fortement générateur d'inégalité dans l'échantillon des employés. Nous mesurons qu'il annule ainsi la moitié du bénéfice égalisateur de la mobilité salariale.

Ces résultats nous enseignent que des politiques en faveur de la réduction des inégalités doivent en priorité s'attacher à réduire la durée du chômage, ou à tout le moins limiter au maximum les pertes de revenu associées au chômage, et à réduire le risque de chômage pour les employés.

La méthodologie que nous développons ici dépasse le cadre de l'application que nous en faisons. C'est un outil de mesure des inégalités de salaires de long terme et de la mobilité des revenus salariaux facilitant la comparaison d'échantillons de trajectoires de revenus différents, par exemple les hommes par rapport aux femmes, différents pays, différents groupes sociaux.

L'outil n'est cependant pas parfait. Notre modèle ne décrit pas de façon complètement satisfaisante la dynamique transitoire des salaires. Différentes pistes méritent d'être explorées. Ainsi, il serait utile de tester des spécifications de copules plus sophistiquées comme la copule de Student, ou autorisant une dynamique markovienne du second ordre. Enfin, il serait aussi judicieux d'autoriser la présence d'erreurs de mesures.

ANNEXE

ESTIMATION

Fixons K , une grille $\underline{\eta}_1, \dots, \underline{\eta}_K$ correspondant à la distribution de la composante permanente η_i et M le nombre de composantes gaussiennes de la distribution marginale des chocs transitoires ε_{it} .

Soit $x_i = (x_{i1}, \dots, x_{iT})$, $y_i = (y_{i1}, \dots, y_{iT})$, et $e_i = (e_{i1}, \dots, e_{iT})$, les observations correspondantes à l'individu i . Soit $j(i) \in \{1, \dots, J\}$ l'index du panel auquel ces observations appartiennent (c'est-à-dire $j(i) = j$ tel que $i \in S_j$).

La vraisemblance d'une observation individuelle est le mélange discret suivant :

$$\ell_j(y_i, e_i | x_i) = \sum_{k=1}^K p_k \ell_j(y_i, e_i | x_i, k), \quad i \in S_j$$

où

$$\begin{aligned} \ell_j(y_i, e_i | x_i, k) &= P_j^0(e_{i1} | x_{i1}, k) \times \prod_{t=1}^{T-1} P_j(e_{i,t+1} | e_{it}, x_{i,t+1}, k) \\ &\times \prod_{t=1}^T \left[\frac{1}{\exp(\alpha_j x_{it})} f_j(\varepsilon(y_{it} | x_{it}, k) | x_{it}) \right]^{e_{it}} \\ &\times \prod_{t=1}^{T-1} c[F_j(\varepsilon(y_{it} | x_{it}, k) | x_{it}), F_j(\varepsilon(y_{i,t+1} | x_{i,t+1}, k)); \exp(\gamma_j x_{it}) - 1]^{e_{it} e_{i,t+1}} \end{aligned}$$

où c est la densité de la copule de Plakett de paramètre $\exp(\gamma_j x_{it}) - 1$, et où

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

Notez que nous permettons que les distributions de ε_{it} soient différentes dans chacun des sous-panels j .

Étape E

Supposons des valeurs initiales des paramètres p_k , $(\mu_{mj}, \pi_{mj}, \omega_{mj})$, (δ_{0j}, μ_{0j}) , $(\rho_j, \delta_{1j}, \mu_{1j})$, $(\beta_j, \sigma_j, \alpha_j)$, et γ_j , pour chaque individu dans le sous-panel S_j , $j = 1, \dots, J$, calculez la probabilité postérieure des types latents de η_i et ε_{it} :

$$p_j(k | y_i, e_i, x_i) = \frac{p_k \ell_j(y_i, e_i | x_i, k)}{\ell_j(y_i, e_i | x_i)}, \quad k = 1, \dots, K,$$

$$\pi_j(m | y_{it}, x_{it}, k) = \frac{\pi_{mj} f_j(\varepsilon(y_{it} | x_{it}, k) | x_{it}, m)}{f_j(\varepsilon(y_{it} | x_{it}, k) | x_{it})}, \quad m = 1, \dots, M.$$

Étape M

Partant de $p_j(k | y_i, e_i, x_i)$ et $\pi_j(m | \varepsilon_{ij}, x_{it}, k)$,

- Actualisez p_k en maximisant

$$\sum_{j=1}^J \sum_{i \in S_j} \sum_{k=1}^K p_j(k | y_i, e_i, x_i) \ln p_k,$$

par rapport à p_k et sous la contrainte :

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

Sur chaque sous-panel j , réalisez les opérations suivantes :

- Maximisez

$$\sum_{i \in S_j} \sum_{m=1}^M \sum_{k=1}^K p_j(k | y_i, e_i, x_i) \sum_{t | e_{it}=1} \pi_j(m | y_{it}, x_{it}, k) \ln \pi_{mj},$$

par rapport à π_{mj} et sous la contrainte :

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

- Maximisez

$$\sum_{i \in S_j} \sum_{k=1}^K p_j(k | y_i, e_i, x_i) \ln P_j^0(e_{i1} | x_{i1}, k; \delta_{0j}, \mu_{0j}),$$

par rapport à (δ_{0j}, μ_{0j}) , et

$$\sum_{i \in S_j} \sum_{k=1}^K \sum_{t=1}^{T-1} p_j(k | y_i, e_i, x_i) \ln P_j(e_{i,t+1} | e_{it}, x_{i,t+1}, k; \rho_j, \delta_{1j}, \mu_{1j}),$$

par rapport à $(\rho_j, \delta_{1j}, \mu_{1j})$.

- Minimisez

$$\sum_{i \in S_j} \sum_{k=1}^K p_j(k | y_i, e_i, x_i) (y_{it} - \beta_j x_{it} - \sigma_j \eta_k)^2,$$

par rapport à β_j et σ_j . Ensuite, calculez les résidus $u(y_{it} | x_{it}, k) = y_{it} - \beta_j x_{it} - \sigma_j \eta_k$ et minimisez

$$\sum_{i \in S_j} \sum_{k=1}^K p_j(k | y_i, e_i, x_i) \left[u(y_{it} | x_{it}, k)^2 - \exp(2\alpha_j x_{it}) \right]^2,$$

par rapport à α_j .

- Calculez les résidus standardisés $\varepsilon(y_{it} | x_{it}, k)$ et maximisez

$$\sum_{i \in S_j} \sum_{k=1}^K p_j(k | y_i, e_i, x_i) \sum_{m=1}^M \sum_{t/e_{it}=1} \pi_j(m | y_{it}, x_{it}, k) \ln \left[\frac{1}{\omega_{mj}} \varphi \left(\frac{\varepsilon(y_{it} | x_{it}, k) - \mu_{mj}}{\omega_{mj}} \right) \right],$$

par rapport à μ_{mj} et ω_{mj} . La solution de ce problème est une moyenne pondérée des variances des résidus standardisés.

- Calculez les rangs :

$$r(y_{it} | x_{it}, k) = F_j(\varepsilon(y_{it} | x_{it}, k) | x_{it}).$$

et maximisez

$$\sum_{i \in S_j} \sum_{k=1}^K \sum_{m=1}^M p_j(k | y_i, e_i, x_i) \sum_{t/e_{it}=e_{i,t+1}=1} \ln c[r(y_{it} | x_{it}, k), r(y_{i,t+1} | x_{i,t+1}, k); \exp(\gamma_j x_{it}) - 1],$$

par rapport à γ_j .

• Finalement, rassemblant tous les sous-panels, régressez les estimations des paramètres j -spécifiques sur une constante, j , j^2 et j^3 . Calculez les estimations contraintes des paramètres, et commencez une nouvelle étape E. Itérez jusqu'à la convergence.

Un code STATA de cet algorithme est disponible sur demande.

BIBLIOGRAPHIE

- AABERGE, R., A. BJØRKLUND, M. JÄNTTI, M. PALME, P. J. PEDERSEN, N. SMITH et 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.
- ABOWD, J. et D. CARD (1989), « On the Covariance Structure of Earnings and Hours Changes », *Econometrica*, 57 : 411-445.
- ARCIDIACONO, A. et J. JONES (2003), « Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm », *Econometrica*, 71 : 933-946.
- BONHOMME, S. et J.-M. ROBIN (2009), « Assessing the Equalizing Force of Mobility Using Short Panels : France, 1990-2000 », *Review of Economic Studies*, 76(1) : 63-92.
- BOWLUS, A. et J. M. ROBIN (2004), « Twenty years of Rising Inequality in US Lifetime labour Income Values », *Review of Economic Studies*, 71(3) : 709-742.
- BURKHAUSER, R. V. et 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.
- BURKHAUSER, R. V., D. HOLTZ-EAKIN et 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.
- FIELDS, G. S. (2005), « Does Income Mobility Equalize Longer-Term Incomes? New Measures of an Old concept », Cornell University.
- FLINN, C. (2002), « Labour Market Structure and Inequality : A Comparison of Italy and the US », *The Review of Economic Studies*, 69(3) : 611-645.
- HAIDER, S. (2001), « Earnings Instability and Earnings Inequality of Males in the United States : 1967-1991 », *Journal of Labor Economics*, 19(4) : 799-836.
- HALL, R. et F. MISHKIN (1982), « The Sensitivity of Consumption to Transitory Income : Estimates from Panel Data on Households », *Econometrica*, 50 : 461-481.
- HOROWITZ, J. L. et M. MARKATOU (1996), « Semiparametric Estimation of Regression Models for Panel Data », *Review of Economic Studies*, 63 : 145-168.
- JOE, H. (1997), *Multivariate Models and Dependence Concepts*, London, Chapman & Hall.
- MAASOUMI, E. et 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.
- MARTIN, J. P. (1996), « Measures of Replacement Rates for the Purpose of International Comparisons : A Note » *OECD Economic Survey*, 26 : 99-115.
- MOFFIT, R. et P. GOTTSCHALK (2002), « Trends in the Transitory Variance of Earnings in the US », *The Economic Journal*, 112 : C68-C73.
- NELSEN, R. B. (1998), *An Introduction to Copulas*, New-York, Springer Verlag.

- SHORROCKS, A. (1978), « Income Inequality and Income Mobility », *Journal of Economic Theory*, 19 : 376-393.
- SKLAR, A. (1959), « Fonctions de répartition à n dimensions et leurs marges », *Publ. Inst. Statist. Univ. Paris*, 8 : 229-231.
- 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.