

# Dynamique des salaires dans une cohorte

Thierry Magnac\*

Sébastien Roux<sup>§</sup>

Révision : 5 mars 2007

## Résumé

Dans cet article, nous examinons les données de salaires annuels, de 1976 à 1998, d'une cohorte d'hommes, entrés sur le marché du travail en 1976. La richesse des données nous permet d'analyser la dynamique des salaires, au cours d'une partie du cycle de vie de cette cohorte, à la différence de la quasi-totalité des autres études sur ce sujet, qui doivent mêler des cohortes différentes, pour avoir un nombre suffisant d'observations. Nous estimons, par pseudo-maximum de vraisemblance, différents modèles dynamiques de panel, de type ARMA, en présence d'effets individuels et de variances hétérogènes temporellement. Les résultats montrent que la variance des salaires, et donc l'inégalité, croît significativement, au cours du début de cycle de vie, mais semble stagner, avant l'arrêt de la croissance des salaires. L'estimation montre que les ordres des processus sont significativement plus élevés, que ceux avancés dans la littérature et sont du type ARMA(4,2). Ces estimations mettent en avant le poids de la composante permanente, liée aux effets individuels, qui rend compte de 60% de la variance des résidus de salaires.

---

\*Université de Toulouse 1 (IDEI et GREMAQ), Aile Jean-Jacques Laffont, Manufacture des Tabacs, 21 Allée de Brienne, 31000, Toulouse, Email : magnac@cict.fr

<sup>§</sup>CREST-INSEE, Paris.

# 1 Introduction<sup>1</sup>

Les modèles dynamiques de revenus, contrôlant des effets transitoires, permettent de construire la distribution des revenus permanents. Cette distribution donne une indication de l'inégalité structurelle de revenus sur tout le cycle de vie (Gottschalk et Moffitt, 1994). Les trajectoires de salaire sont néanmoins très hétérogènes au cours du temps, et entre cohortes, et l'analyse de la formation des inégalités de salaires, au cours de la carrière, se doit de contrôler les effets de cohorte.

Les modèles linéaires de panel à effets individuels, développés dès les années 70 (Hause, 1977, Lillard et Willis, 1979), ont débouché rapidement sur l'étude de dynamiques plus complexes (MacCurdy, 1982, Abowd et Card, 1989). L'ensemble des études sur ce sujet (utilisant essentiellement les données américaines du Panel Study of Income Dynamics, PSID) est résumé par Alvarez, Browning et Eyrnæs (2002). Le consensus passé indiquait que le processus simple qui caractérise le mieux les données de salaires, est un processus ARMA(1,2) (Baker, 1997, par exemple). Des études récentes remettent en cause ce consensus, en étudiant des dynamiques non-linéaires qui apporte des éclairages intéressants et différents (Alvarez et al., 2002, Meghir et Pistaferri, 2004), ou en étudiant d'autres aspects de l'inégalité, de la mobilité des revenus et de la pauvreté (Geweke et Keane, 2000, Hirano, 2002, Bonhomme et Robin, 2004).

Dans cet article, nous étudions la dynamique des salaires d'une seule cohorte, observée sur très longue période, de son entrée sur le marché du travail en 1976 jusqu'en 1998. Au delà du seul constat de l'existence d'inégalités, nous cherchons à caractériser le processus de formation des salaires, au cours de la carrière des individus, pour mieux comprendre comment celles-ci se forment. La cohorte que nous étudions, est définie, de manière peut-être inhabituelle, en fonction de *sa date* d'entrée sur le marché du travail (voir Le Minez et Roux, 2002). La période d'observation nous permet d'analyser l'inégalité intra-cohorte, au cours d'un peu plus que la première moitié de la vie active des individus de cette cohorte. Nous utilisons les données du panel DADS (Déclarations Annuelles de Données Sociales) parce que leur grande richesse permet de décrire, de manière précise, la dynamique des salaires dans le secteur privé. Nous choisissons la date de 1976, à cause de la rupture de tendance dans l'augmentation des salaires et revenus, apparue à la fin des années 1970 (Piketty, 2001) et parce que la période de temps à analyser est suffisamment longue, pour identifier la dynamique sous-tendant les trajectoires observées.

---

<sup>1</sup>Les auteurs ont bénéficié de nombreuses discussions avec Stéphane Grégoir, et des remarques de deux rapporteurs anonymes. Toutes les estimations reportées dans cet article ont été effectuées à l'aide du logiciel R (<http://www.r-project.org/>). Les programmes sont disponibles auprès des auteurs. Les auteurs restent néanmoins responsables des insuffisances et erreurs de cet article.

À la différence de Koubi (2004), qui a étudié avec attention les disparités de profil salarial liées aux caractéristiques observables des individus dans plusieurs cohortes, nous nous intéressons *seulement* à la variance des salaires, des individus de la cohorte sélectionnée, conditionnellement à certaines variables explicatives (expérience et éducation). Nous faisons également abstraction du débat sur la forte augmentation de l'inégalité des salaires, intervenue depuis le début des années 70 aux États-Unis et à la fin des années 70 au Royaume-Uni et sur la baisse ou la stabilité de la mobilité (Gottschalk, 1997). En France d'ailleurs au contraire, c'est à une décroissance des inégalités de salaires que l'on assiste, au moins entre 1967 et 1991 et, semble-t-il, à une stagnation ensuite (Bayet, 1996, Bonhomme et Robin, 2004, Koubi, 2004). Celle-ci s'est accompagnée d'une décroissance de la mobilité entre différents déciles de revenu dans les années 70 (Buchinsky, Fougère et Kramarz, 1999).

L'objectif de cet article est de mettre en évidence la façon dont les salaires se différencient au cours de la carrière, en caractérisant leur processus de formation, de la façon la plus simple et parcimonieuse possible. Pour cette raison, nous choisissons un cadre linéaire pour décrire ce processus. Nous travaillons sur l'écart des logarithmes de salaire, à leur moyenne par an, âge d'entrée et éducation, dont le profil a été examiné par Koubi (2004). Nous modélisons ces écarts, ou résidus, ainsi obtenus en prenant en compte des effets individuels en niveau, et une modélisation ARMA (Auto-Regressive Moving Average) pour la dynamique, mais en permettant l'hétéroscédasticité temporelle des innovations. Une telle modélisation nous permet de décomposer la variance de ces résidus, en une composante permanente liée aux caractéristiques inobservables fixes des individus et une composante transitoire liée à la dynamique des salaires. Ceci nous permet d'examiner la dépendance entre ces composantes et les conditions initiales du processus.

La décomposition que nous retenons, *in fine*, s'appuie sur des processus stationnaires. De fait, on peut tester la non-stationnarité de ces processus, en utilisant nos estimations, et elle est systématiquement rejetée. On peut ainsi interpréter la dynamique des salaires de la cohorte étudiée, comme une convergence progressive vers un processus stationnaire, à partir de conditions initiales. La corrélation des salaires à long terme correspond alors à la composante permanente, liée aux caractéristiques fixes des salariés.

En ce qui concerne les méthodes d'estimation, Alvarez et Arellano (2004) ont récemment remis en cause l'utilisation de méthodes d'estimation GMM, en utilisant des données très similaires aux nôtres, issues du PSID, en arguant du fait que les biais à distance finie et les pertes d'information peuvent être importantes. Des méthodes de maximum de vraisemblance ou de pseudo-maximum de vraisemblance, sous hypothèse de normalité, semblent plus robustes, et ce sont celles que nous

utilisons. De plus, l'utilisation de méthodes de vraisemblance est plus aisée que des méthodes de moment (comme les GMM), lorsque les données contiennent de nombreuses valeurs manquantes, comme c'est le cas ici. Les méthodes de moment, pour être simplement utilisées, requièrent, en effet, que les séquences des observations soient ininterrompues. Par rapport à la littérature existante, une innovation de cet article est qu'il propose des algorithmes, permettant de prendre en compte les observations manquantes dans les modèles dynamiques, mais en restant dans un cadre où l'attrition, c'est à dire la disparition du panel, est exogène.

Parmi les résultats principaux que nous obtenons, nous retiendrons d'abord que l'inégalité (approchée ici par la variance des résidus de salaire) augmente jusqu'à 35-37 ans, dans cette cohorte, et reste stable ensuite. Le processus de formation des salaires se décompose en une composante permanente et une composante transitoire, suivant un ARMA, d'un ordre plus élevé que dans la littérature, de type ARMA(4,2) pour l'ensemble de la population, et de type ARMA(4,1) pour la sous-population des non-qualifiés. L'augmentation des inégalités, au cours de la carrière, s'explique par la convergence des salaires vers une trajectoire "stationnaire", correspondant au processus ARMA estimé. L'écart à cette trajectoire, en début de carrière, correspond aux conditions initiales, qui font partir l'ensemble des salariés, à des niveaux de salaires plus concentrés que les niveaux atteints après dix ans de carrière. Des tests de rupture ont été conduits, afin d'examiner la stabilité de ces représentations dans le temps : ils rejettent l'hypothèse de stabilité, quel que soit l'ordre examiné, la statistique de test étant, néanmoins, moins élevée pour les processus finalement retenus. Les représentations de la dynamique salariale, examinées ici, doivent donc être considérées comme des approximations imparfaites du "vrai" processus générateur des salaires.

À partir du processus estimé, il est possible d'estimer la variance stationnaire, atteinte par le processus au bout de suffisamment d'années de carrière, et de la décomposer, entre celle liée à la composante permanente fixe et celle liée à la dynamique salariale propre. La part de variance individuelle des salaires, dans la solution de long terme, est estimée à 60% et il y a peu de différences dans ces variances, entre groupes d'éducation, quand celle-ci est approchée par la catégorie socio-professionnelle du premier emploi.

La Section 2 présente la modélisation et les méthodes d'estimation. Nous décrivons les données dans la Section 3. La Section 4 reporte les résultats d'estimation.

## 2 Processus dynamiques et estimation

Nous partons de l'idée qu'il y a beaucoup plus de diversité dans les trajectoires de salaires que celle qui avait été étudiée auparavant. Cela se traduit d'abord par le choix de la base de données qui ne comporte qu'une seule cohorte, ensuite par la manière dont nous modélisons la dynamique des salaires. Nous nous intéressons d'abord aux différents types de modèles qui ont été proposés dans la littérature puis aux méthodes d'estimation qui ont été utilisées.

### 2.1 Modèles

Le modèle le plus simple de décomposition de la variance en effets permanents et transitoires est (Hause, 1980, Lillard et Willis, 1978) :

$$y_{it} = \gamma_i + \varepsilon_{it}, \quad (1)$$

où  $\gamma_i$  est un effet individuel. Il faut préciser néanmoins quelles sont les propriétés postulées pour le choc transitoire. C'est MaCurdy (1982) qui a proposé l'un des modèles qui reste parmi les plus populaires dans la littérature sur la dynamique des salaires. En utilisant les données du PSID, il trouve que l'innovation  $\varepsilon_{it}$  semble bien approximée par un ARMA(1,2) :

$$\varepsilon_{it} = \rho\varepsilon_{it-1} + \eta_{it} + \theta_1\eta_{it-1} + \theta_2\eta_{it-2} \quad (2)$$

où  $\eta_{it}$  est supposé être un processus bruit blanc.

La littérature a pris différentes directions pour généraliser ce modèle de base.

Une première question tourne autour de la présence ou non d'hétérogénéité inobservable dans les taux de croissance des salaires et donc dans les rendements de l'expérience (Baker, 1997). Les taux de croissance des salaires des plus éduqués sont plus forts que les taux de croissance de ceux des moins éduqués par exemple. Baker suppose ainsi qu'il y a un deuxième effet individuel :

$$y_{it} = \gamma_i + \lambda_i.t + \varepsilon_{it}, \quad (3)$$

qui crée une corrélation sérielle persistante dans les taux de croissance puisqu'en supposant l'indépendance de  $\lambda_i$  et  $\Delta\varepsilon_{it}$  par ailleurs, on a :

$$Cov(\Delta y_{it}, \Delta y_{it-k}) = V\lambda_i \quad (4)$$

si  $k$  est suffisamment grand (supérieur à l'ordre de corrélation sérielle dans  $\Delta\varepsilon_{it}$ ).

Néanmoins, une autre question ouverte tourne autour de la racine auto-régressive. Le modèle est-il stationnaire,  $\rho < 1$  ? ou possède-t-il une racine unitaire,  $\rho = 1$  ? La résolution de cette question est difficile puisque Baker (1997) montre que les deux modèles sont extrêmement proches sur les données dont on dispose habituellement pour analyser les salaires, c'est-à-dire des périodes de temps d'une vingtaine d'années comme dans les données dont nous disposons ici. Autrement dit, ce sont des alternatives locales et le gain de modélisation apporté par l'équation (3) par rapport à la modélisation initiale s'estompe. Notons qu'il y a aussi des arguments économiques avancés par Baker et Solon (1999) pour expliquer la présence d'une racine unitaire. Des chocs comme des licenciements, qui ne sont pas explicitement modélisés ici, peuvent avoir des effets de très long terme s'ils ont pour conséquence la destruction irréversible de capital humain spécifique.

La direction prise par Geweke et Keane (2000) ou Hirano (2002) est différente. Elle consiste à remettre en cause l'utilisation de l'hypothèse de normalité pour les bruits blancs,  $\eta_{it}$ , considérés plus haut. Si l'objectif est de construire une matrice de mobilité entre quantiles de salaire, l'hypothèse de normalité peut se révéler en effet une source de biais très importante. Horowitz et Markatou (1996) adoptent une approche similaire en se proposant d'estimer semi-paramétriquement les différentes distributions des effets individuels et des bruits blancs. Un autre intérêt des deux articles de Geweke et Keane, et Hirano, est d'utiliser des méthodes bayésiennes puissantes comme l'échantillonnage de Gibbs pour construire les distributions a posteriori des paramètres et pas seulement leurs distributions asymptotiques. Ceci permet d'atteindre une meilleure précision à distance finie – si la spécification retenue est la bonne – quand la mobilité par quantiles doit être prédite. Néanmoins ces améliorations sont obtenues au prix d'une simplification de la dynamique du modèle de base en un simple AR(1).

C'est avec la même intention que Bonhomme et Robin (2004) procèdent en modélisant l'évolution des salaires par des copules qui permettent de représenter de façon partiellement non paramétrique les distributions marginales des salaires et de façon flexible la distribution jointe des salaires à des périodes successives. Là aussi, mais cela est dû en partie à la faible longueur de l'échantillon tiré des enquêtes Emploi (3 ans), la dynamique est modélisée de manière plus restreinte. Il est aussi difficile d'intégrer de manière générale l'hétérogénéité inobservable au delà de l'approximation de sa distribution par un nombre fini de points de support.

La troisième direction prise par la littérature porte sur d'autres non-linéarités. La première possibilité est d'explorer des processus générateurs des données comme des modèles ARCH (Meghir et Pistaferri, 2004) dans un modèle où il y a des chocs permanents et transitoires. Ils trouvent que la variance des chocs de revenu est persistente dans certains groupes d'éducation ce qui a

des conséquences sur l'épargne de précaution. Ces motivations sont aussi au cœur des méthodes proposées par Alvarez et al. (2002). Ces auteurs proposent de rendre compte de la difficulté de modéliser les processus de revenu en supposant que les paramètres des processus ARMA sont hétérogènes entre individus. Ainsi, la racine autorégressive peut être unitaire pour certains individus mais pas pour d'autres. De même, les variances et racines de la moyenne mobile peuvent différer entre individus. Des modèles de plus en plus complets sont estimés par des méthodes de simulation dérivées de méthodes d'inférence indirecte.

L'objectif que nous suivons dans cet article est plus modeste. D'abord, nous nous contentons de vouloir mesurer des paramètres liés aux variances et covariances et non aux distributions et nous ne visons pas à utiliser nos résultats pour estimer la mobilité entre quantiles de revenu, seulement à examiner l'évolution, dans une cohorte, de l'inégalité des salaires (une fois les caractéristiques observables contrôlées) au cours des premières années du cycle de vie. Ensuite, nous adoptons une forme réduite en ne cherchant pas à modéliser l'hétérogénéité individuelle des variances des termes d'erreur mais uniquement leurs propriétés de séries temporelles. Nous nous contentons d'estimer des processus ARMA dont la variance varie au cours du temps mais en tentant de le faire d'une façon efficace, cette estimation restant robuste à la présence d'hétérogénéité dans les variances. Enfin, nous considérons dans un premier temps que la dynamique des salaires est la même pour tous les individus composant la cohorte étudiée, tout en reportant dans un deuxième temps, des résultats pour des groupes de qualification différente. Nous estimons les paramètres gouvernant cette dynamique commune de manière à ce que l'hétérogénéité des trajectoires salariales qu'elle génère corresponde à celle observée dans les données. Ce sont ces méthodes que nous présentons maintenant.

## 2.2 Estimation

Pour estimer la représentation théorique la plus adaptée du processus générateur des données, il faut pouvoir mesurer l'adéquation du modèle théorique à la matrice de variance empirique d'observations longitudinales  $y_{it}$ . La méthode d'estimation des paramètres d'Abowd et Card (1989) consiste à rendre la variance théorique du processus générateur des données, résumée par  $\mathbf{cov}$ , la plus proche possible de la matrice de variance empirique,  $\widehat{\mathbf{cov}}$ . La méthode des moindres carrés asymptotiques ou de distance minimum s'applique naturellement dans de tels cas (Gouriéroux et Monfort, 1989).

Soit  $b$  le vecteur de paramètres définissant le processus statistique que nous estimons, l'estimateur

donné par les m.c.a est celui minimisant le critère :

$$\hat{b} = \arg \min_b N (\widehat{\mathbf{cov}} - \mathbf{cov}(b))' \hat{V}_{\mathbf{cov}}^{-1} (\widehat{\mathbf{cov}} - \mathbf{cov}(b)) \quad (5)$$

où  $N$  est le nombre d'observations. Sous l'hypothèse nulle, le critère minimisé suit un Chi deux à un nombre de degrés de liberté égal à  $rg(\mathbf{cov}(b)) - rg(b)$ , soit  $T(T+1)/2 - rg(b)$ . C'est cette statistique que l'on peut considérer pour accepter ou refuser le modèle théorique.

Néanmoins, à distance finie, Altonji et Segal (1996) ont montré que cet estimateur appelé par eux, Optimal Minimum Distance (OMD) était sérieusement biaisé par rapport à celui qui utilise la matrice identité comme matrice de poids, dit Equally Weighted Minimum Distance (EWMD) :

$$\hat{b}_{EWMD} = \arg \min_b \{N (\widehat{\mathbf{cov}} - \mathbf{cov}(b))' (\widehat{\mathbf{cov}} - \mathbf{cov}(b))\} \quad (6)$$

C'est pourquoi le deuxième estimateur est souvent préféré. Il existe des moyens de corriger partiellement les biais du premier estimateur à distance finie soit par des techniques de jackknife (Kezdi, Hahn et Solon, 2002) ou de bootstrap (Horowitz, 1998). Ces méthodes sont néanmoins numériquement très coûteuses et ne semblent pas améliorer de façon substantielle les résultats de l'estimation par EWMD. La littérature s'est aussi employée à proposer d'autres estimateurs, de vraisemblance empirique, d'"exponential tilting" ou de GMM avec révision continue ("continuously updating"). Les résultats d'une analyse complète par Monte Carlo de ces différentes méthodes d'estimation sont présentés dans Ramalho (2005).

Nous utilisons en fait une autre méthode qui est celle du maximum de vraisemblance. L'argument principal à l'appui de cette dernière méthode d'estimation est celui d'efficacité dans le cas où le modèle est bien spécifié et de robustesse dans le cas où le modèle est mal spécifié en utilisant un cadre de pseudo-vraisemblance (Gouriéroux et Monfort, 1989).

A la suite d'Alvarez et Arellano (2004), on considère l'équation, valide de  $t = 1$  à  $T$  seulement :

$$y_{it} = \alpha_1 y_{i(t-1)} + \dots + \alpha_p y_{i(t-p)} + \sigma_\eta \eta_i + v_{it} \quad (7)$$

où  $\eta_i$  est un effet individuel centré et réduit. Les conditions initiales sont supposées données ( $y_{i0}, \dots, y_{i(1-p)}$ ) et le terme d'erreur est :

$$v_{it} = \sigma_t w_{it} \quad (8)$$

où  $w_{it}$  est une moyenne mobile d'ordre  $q$  :

$$w_{it} = \zeta_{it} - \psi_1 \zeta_{it-1} - \dots - \psi_q \zeta_{it-q}. \quad (9)$$



Le modèle est donc fonction des variables aléatoires  $z_i = (y_{i0}, \dots, y_{i(1-p)}, \eta_i, \zeta_{i(1-q)}, \dots, \zeta_{iT})$ . Nous supposons que ce vecteur est distribuée selon une loi normale multivariée :

$$z_i \rightsquigarrow N(0, \Omega_z)$$

où  $\Omega_z$  a une structure plus précisément décrite dans l'annexe A.1 mais que nous pouvons résumer de la façon suivante. Les corrélations entre conditions initiales et effet individuel sont non restreintes alors que les innovations  $\zeta_{it}$  sont supposées orthogonales à toute variable passée, y compris aux conditions initiales. Par contre, la condition initiale,  $y_{i0}$ , peut être corrélée avec des chocs antérieurs comme  $\zeta_{i0}, \zeta_{i(-1)}, \dots$

On déduit de ces restrictions les moments des variables observables :

$$y_i \rightsquigarrow N(0, \Omega_y).$$

où  $\Omega_y$  est une fonction des paramètres du modèle ( $\{\alpha_k\}_{k=1, \dots, p}, \{\psi_k\}_{k=1, \dots, q}, \sigma_\eta, \{\sigma_t\}_{t=1, \dots, T}, \Sigma$ ). (Voir Annexe A.1).

L'hypothèse de normalité peut paraître forte de prime abord mais elle peut être réinterprétée comme une hypothèse auxiliaire sur laquelle s'appuie la méthode du pseudo-maximum de vraisemblance. Les estimateurs qui maximisent la vraisemblance auxiliaire sont convergents et asymptotiquement normaux et leur matrice de variance-covariance se calcule de manière robuste aux violations de l'hypothèse de normalité.

Enfin, un des avantages de la méthode de maximum de vraisemblance est de permettre de traiter de façon aisée, de manière conceptuelle tout au moins, les valeurs manquantes dans les séries individuelles. En effet, écrivons  $y_i^{obs}$  le vecteur, de taille,  $T_i \leq T$ , des valeurs observées du vecteur  $y_i$  et notons  $S_i$  la matrice de sélection, de taille  $[T_i, T]$  qui est définie par :

$$y_i^{obs} = S_i y_i. \tag{10}$$

A condition de supposer que la loi de  $S_i$  ne dépend pas des paramètres du modèle structurel, nous aurons :

$$y_i^{obs} \rightsquigarrow N(0, S_i \Omega_y S_i')$$

et la contribution de cette observation à la vraisemblance s'en déduit.

Si le traitement conceptuel des valeurs manquantes est simplement décrit ci-dessus, la mise en œuvre pratique est plus complexe. En effet, il faut tenir compte pour chaque individu de la séquence des années pour lesquelles il est observé dans les données pour écrire sa contribution à la

vraisemblance. Étant donné le nombre d'années pour lesquelles les individus peuvent être présents ou absents cela correspond à  $2^T$  combinaisons possibles<sup>2</sup>.

## 2.3 Tests

Pour choisir la meilleure représentation des données, deux critères sont utilisés. Le premier repose sur la comparaison des vraisemblances entre deux modèles emboîtés l'un dans l'autre en mettant en œuvre le test du rapport de vraisemblance. Le second consiste à vérifier si la dynamique générant les trajectoires salariales est stable. Pour cela, nous mettons en œuvre un test de rupture des coefficients : l'hypothèse nulle que les paramètres gouvernant la dynamique salariale restent identiques avant et après 1987.

### Test du rapport de vraisemblance

Le parti pris dans ce papier est de se limiter à une classe de processus linéaires de type ARMA pour représenter les trajectoires salariales. La question de l'ordre du processus reste toutefois non résolue. Pour le déterminer, nous estimons des modèles différents en considérant des ordres AR allant de 1 à 4 et MA de 1 à 3. L'examen des vraisemblances permet de comparer un modèle d'ordre  $(p, q)$  donné à un modèle d'ordre supérieur, par exemple  $(p + k, q)$ . Le test du rapport de vraisemblance consiste alors en un test de spécification du modèle de type ARMA( $p, q$ ) contre la spécification ARMA( $p + k, q$ ). La statistique de test est la suivante :

$$S^L = 2(L(p + k, q) - L(p, q)) \quad (11)$$

où  $L(p, q)$  est la valeur de la log-vraisemblance atteinte par le modèle d'ordre  $(p, q)$ . Sous l'hypothèse que le modèle ARMA( $p, q$ ) est bien spécifié, cette statistique suit une loi du  $\chi_2^2$ , à  $k$  degrés de liberté. On peut aussi utiliser les critères d'Akaike ou BIC pour comparer les modèles non emboîtés (Gouriéroux et Monfort, 1989).

### Test de rupture

Le principe du test de rupture des coefficients consiste à vérifier si les paramètres du modèle restent constants au cours du temps. On considère ainsi un modèle plus général que celui estimé, dans lequel les coefficients gouvernant la dynamique ARMA, à savoir les vecteurs de paramètre  $\alpha$  et  $\psi$ , peuvent être différents entre la période 1976-1987 et 1988-1998. Formellement, cela revient à

---

<sup>2</sup>Plutôt que d'écrire toutes les combinaisons possibles, nous listons l'ensemble des combinaisons présentes dans les données, soit l'ensemble des  $S_i$  possibles, et estimons pour chaque groupe d'individus correspondant au même  $S_i$  leur contribution à la vraisemblance.

réécrire la dynamique du modèle comme dépendant de paramètres  $\alpha^1$  et  $\psi^1$  en première période et  $\alpha^2$  et  $\psi^2$  en deuxième période. Le test de rupture consiste à tester de façon jointe les égalités  $\alpha^1 = \alpha^2$  et  $\psi^1 = \psi^2$ . Les autres paramètres relatifs aux conditions initiales restent les mêmes.

Comme nous estimons de toute façon le modèle contraint (sous l’hypothèse d’absence de rupture), nous disposons des estimations de  $\alpha$  et  $\psi$  dans ce cadre et des autres paramètres du modèle. Le test du score consiste à vérifier si ces valeurs maximisent bien la vraisemblance du modèle non contraint (c’est-à-dire celui autorisant la rupture des coefficients). Ce test consiste à examiner le gradient du logarithme de la vraisemblance non contrainte, également appelé le vecteur du score, et vérifier qu’il n’est pas différent de zéro. Si on appelle  $s(\alpha, \psi)$  le vecteur du score, la statistique de test est la suivante :

$$S^S = \frac{1}{N} s(\hat{\alpha}, \hat{\psi})' V(s_i)^{-1} s(\hat{\alpha}, \hat{\psi}) \quad (12)$$

où  $V(s_i)$  est la matrice de variance des vecteurs du score par individu. Le nombre de relations d’égalité testées est égal à  $p+q$ , correspondant aux dimensions des vecteurs  $\alpha$  et  $\psi$ . Sous l’hypothèse nulle, la statistique de test  $S^S$  suit donc une loi du  $\chi_2$  à  $p+q$  degrés de liberté.

### 3 Données et Description

La richesse de la base de données que nous utilisons nous permet d’estimer la dynamique des salaires dans une cohorte unique définie comme l’ensemble des individus rentrés sur le marché du travail en 1976. Dans cette section, après avoir présenté cette base, nous expliquons comment nous construisons les variances résiduelles des salaires après avoir retiré de leur moyenne les effets de période, d’âge d’entrée et d’éducation. Nous décrivons ensuite la structure des variances résiduelles.

#### 3.1 Sélection de l’échantillon

Nous utilisons la même base de données que celle de Le Minez et Roux (2002) en nous restreignant aux hommes rentrés sur le marché du travail en 1976 et travaillant dans le secteur privé en 1998. Cette base est extraite de l’échantillon des DADS de l’INSEE (Roux, 2003) et reporte les salaires des travailleurs du secteur privé et semi-public tels qu’ils sont déclarés par les employeurs. Comme Baker et Solon (1999) qui utilisent des données similaires sur le Canada, cette mesure des salaires a plusieurs avantages par rapport aux données d’enquêtes qui ont été utilisées jusqu’à présent. Comme ces revenus sont utilisés pour des raisons administratives, ces observations devraient être libres d’erreurs de codage, d’arrondi et autres erreurs de mesure même si peuvent subsister des problèmes d’adéquation entre la mesure annuelle que nous utilisons et le vrai modèle économique

de détermination des salaires. De plus, l'échantillonnage des individus nés en octobre d'une année paire, permet une bonne représentativité des données et surtout accroît le nombre d'individus présents dans la base.

Il y a un certain nombre de contreparties à ces avantages. D'abord, peu de variables décrivent l'hétérogénéité observée entre les individus et seules trois variables sont utilisables à part le sexe : la date de naissance, la période et un indicateur de catégorie socio-professionnelle. Ce dernier est mesuré lors du premier emploi et nous regroupons ces catégories en trois groupes des moins aux plus éduqués. Ensuite, comme l'univers n'est pas clos, il y a dans les histoires des individus de nombreux exemples d'absences. Sans exclure que certaines de ces sorties temporaires soient de pures erreurs, d'autres absences "structurelles" se réfèrent à des passages hors-champ de l'observation, comme des passages par l'inactivité, le chômage, un travail dans le secteur public ou comme non salarié. Il y a aussi des valeurs manquantes dues à l'absence de données DADS en 1981, 1983 et 1990 à cause des recensements ainsi que des anomalies, qui sont maintenant bien connues, en ce qui concerne l'année 1993.

La richesse de ces données nous permet de nous concentrer sur une seule cohorte, définie par sa date d'entrée sur le marché du travail. La littérature que nous avons présentée dans la section précédente montre, en effet, que l'hétérogénéité dans les processus de revenu est très importante (Alvarez et al., 2001). Considérer un échantillon trop hétérogène conduit alors à des problèmes de mauvaise spécification qui peuvent conduire à des biais dans la modélisation de la dynamique et en particulier dans le nombre de retards dans les processus ARMA.

Ce souci d'homogénéité nous a aussi conduit à la sélection des hommes ayant tous occupé leur premier emploi à temps plein de plus de 6 mois en 1976 et occupant toujours un emploi à temps plein en 1998 (cf. Le Minez et Roux, 2002, pour une description plus précise de la construction des cohortes). Cette sélection en fonction de la présence aux dates initiales et finales et du temps passé en emploi, peut être endogène mais notre intention n'est que descriptive. Notre analyse se restreint à la dynamique des salaires dans cette population qui nous paraît une des populations les plus à même de décrire, dans le champ du salariat du secteur privé, la façon dont les inégalités se forment pendant les 20 premières années de la carrière compte tenu de toute interruption due à l'inactivité, au chômage ou à d'autres activités. Cette analyse n'est donc pas une tentative de construction de salaires potentiels dans le secteur privé ou dans l'économie puisque nous devrions alors corriger des biais de sélection dus à l'attrition (Guillotin et Sevestre, 1994). Ceci est hors du champ de cette étude mais en est un des développements possibles.

Notre échantillon est composé de 5236 hommes dont la première entrée dans le champ des DADS

est 1976. Il faut noter que, malgré le caractère administratif des données, la sélection retenue ne représente qu'environ le tiers d'une génération annuelle rentrant sur le marché du travail. D'autres rentrent dans le secteur public ou deviennent indépendants dès la sortie du système scolaire ou après une période passée dans le secteur privé. D'autres connaissent une période temporaire hors du secteur privé en 1998. Leur âge et éducation en 1976 sont décrits dans le Tableau 1. Plus de la moitié de l'échantillon est entrée à 18 ans sur le marché du travail (ou plus précisément, dans le champ couvert par nos données) et un autre gros quart y est entré à 20 ou 22 ans.<sup>3</sup> Le reste se répartit entre des entrées précoces à 16 ans (7%) et des entrées tardives de 24 à 34 ans (10%). En termes d'éducation, la population ayant occupé un premier emploi très qualifié (i-e de cadre ou de profession intermédiaire) reste faible (7%), les individus ayant débuté par un emploi qualifié ou non qualifié (i-e d'employé ou d'ouvrier) se partageant le reste de l'échantillon. Ne disposant pas directement du niveau de diplôme atteint par les individus échantillonnés, nous interprétons la catégorie socio-professionnelle au premier emploi comme un niveau d'éducation : les individus occupant un premier emploi très qualifié sont considérés comme étant passés dans l'enseignement supérieur, ceux occupant un premier emploi qualifié comme étant passés par l'enseignement secondaire.

La structure des valeurs manquantes est décrite par le tableau 2. Ce tableau donne la fréquence des présences à deux dates  $t$  et  $t'$  pour toutes les périodes entre 1976 et 1998 sauf pour les années 1981, 1983 et 1990 puisque les DADS ne furent pas traitées ces années-là. Comme, par hypothèse de construction, tout le monde est présent en 1976 et 1998, la taille de l'échantillon pour chaque année se lit dans les lignes ou colonnes qui font référence à ces deux périodes (la ligne de 1976 ou la colonne de 1998). Elle atteint son minimum en 1985, année de la période entre 1976 et 1998 dont la conjoncture macro-économique est une des plus mauvaises avant la remontée des années 1987 à 1990. En 1985, l'attrition atteint son maximum de 17%, qui reste une valeur modérée par rapport à des données d'enquête sur les revenus ou salaires. Cette mesure de l'attrition ne résume qu'une partie de la structure des valeurs manquantes. Le Tableau 2 donne des renseignements sur la probabilité d'observation de paires de périodes. Par exemple, 70,6% de l'échantillon est présent à la fois en 1985 et 1994, cette dernière année étant également une année de mauvaise conjoncture économique. C'est la valeur la plus basse que prend cette fréquence d'observations. Il y a peu d'autres caractéristiques à noter dans ce tableau. La sortie de l'échantillon entre les années 1976 et 1985 semble plus rapide que la remontée des taux entre 1985 et 1998 ce qui est peut être un indice d'une stabilisation de la relation au travail salarié privé mais cela reste une conjecture. On peut noter la baisse relative du

---

<sup>3</sup>Comme seules sont concernées des personnes nées en octobre des années paires, la structure par âge en 1976 ne comporte que des âges pairs.

taux de présence en 1994 par rapport aux autres années qui est due à une erreur d'échantillonnage des individus dans l'enquête DADS pour cette année.

### 3.2 Construction des données

Le salaire que nous utilisons est le salaire net journalier perçu. Il est construit comme le salaire total divisé par le nombre de jours rémunérés par l'employeur. S'il y a plusieurs employeurs, nous reconstruisons une mesure pondérée de ce salaire (voir Le Minez et Roux, 2002, pour une explication plus détaillée). Ne sont conservées parmi les observations que celles correspondant à des emplois à temps complet et pour lesquelles les salaires journaliers ne sont pas dans les centiles extrêmes de la distribution annuelle des salaires. Ces observations sont alors considérées comme manquantes, comme si les individus concernés n'étaient pas présents les années correspondant à ces observations. Même lorsqu'une de leur observation est manquante, les individus sont conservés dans l'échantillon, les autres observations étant source d'information.

Nous avons déjà expliqué que nous ne disposons que de peu de variables explicatives individuelles ou familiales. Les régressions de salaire à la Mincer ne peuvent donc apporter d'informations que sur les effets de l'âge ou de l'expérience. Introduire des variables explicatives fixes au cours du temps, comme l'éducation, n'a pas beaucoup de sens puisque l'effet de ces variables est confondu avec les effets individuels constants au cours du temps et que l'orthogonalité entre elles et les effets individuels est douteuse. On peut néanmoins effectuer l'analyse économétrique par groupes d'éducation et c'est ce que nous proposerons dans la partie empirique.

Nous utilisons d'abord l'information sur l'âge d'entrée, l'éducation et la période. Si  $E_{it}$  est le revenu salarial de l'individu  $i$  au temps  $t$ , nous étudions l'écart du logarithme de cette valeur à la moyenne par âge d'entrée croisée avec l'éducation et par période :

$$y_{it} = \log(E_{it}) - \overline{\log(E_{it})_{aet}} \quad (13)$$

où  $\bar{x}_{aet}$  désigne la moyenne de  $x$  pour un âge d'entrée  $a$ , avec une éducation  $e$ , au temps  $t$ . En cela, on caractérise alors le résidu à un moment du temps et les positions relatives des salaires à l'intérieur d'un groupe d'expérience et d'éducation donné (Koubi, 2003). Nous contrôlons donc des effets de l'expérience de manière très générale, en différenciant ces effets en fonction de l'éducation, sans imposer un profil particulier, comme cela est le cas dans les équations de Mincer. Nous ne reporterons pas ici les profils des salaires moyens au cours de la carrière puisqu'ils sont très bien connus maintenant (Koubi, 2003, et les références qui y sont citées). Les salaires augmentent fortement jusqu'à 40 ans environ, évoluent plus faiblement entre 40 et 50 ans et stagnent ou décroissent

à partir de 50 ans.

### 3.3 L'estimation de la matrice de variance-covariance des salaires

Ce sont donc les variances résiduelles qui nous intéressent. La matrice de variance-covariance des résidus des équations de salaire,  $y_{it}$ , peut être estimée par :

$$\widehat{V}(\mathbf{y}_i) = \frac{\sum_{i=1}^N (y_{i1}, \dots, y_{iT}) (y_{i1}, \dots, y_{iT})'}{N-1}. \quad (14)$$

Les modèles présentés dans la section précédente se proposent de comparer cette matrice de variance-covariance empirique à celle correspondant au processus générateur des données postulé. Par exemple, si les salaires sont générés par un bruit blanc, on aura :  $V(y_{it}) = \sigma^2$ , et  $cov(y_{it}, y_{is}) = 0$  pour  $t \neq s$ . Dans ce cas, la matrice de variance du processus théorique est l'identité multipliée par la variance.

Puisque cette matrice est symétrique, elle ne compte que  $\frac{T(T+1)}{2}$  éléments indépendants que sont les covariances théoriques entre  $y_{it}$  et  $y_{is}$ ,  $cov_{t,s} = E((y_{it} - Ey_{it})(y_{is} - Ey_{is}))$ . On réécrit les éléments de cette matrice de variance-covariance sous forme d'un vecteur de taille  $T(T+1)/2$ , appelé  $\mathbf{cov} = (cov_{1,1}, \dots, cov_{T,1}, cov_{2,2}, \dots, cov_{T,2}, \dots, cov_{T,T})'$ . On peut alors appliquer le théorème central limite aux contreparties empiriques de ces moments théoriques, sous des conditions portant sur l'existence de moments d'ordre supérieur et on obtient :

$$\sqrt{N}(\widehat{\mathbf{cov}} - \mathbf{cov}) \Rightarrow N(0, V_{\mathbf{cov}}) \quad (15)$$

où  $\Rightarrow$  désigne la convergence en distribution et  $\widehat{\mathbf{cov}}$  est la contrepartie empirique de  $\mathbf{cov}$ , c'est-à-dire la représentation vectorielle de la matrice  $\widehat{V}(y_i)$ .  $V_{\mathbf{cov}}$  est la matrice de variance-covariance de  $\mathbf{cov}$ , et est donc de taille  $T(T+1)/2 * T(T+1)/2$ . Une estimation convergente de cette matrice sera donnée par les moments d'ordre 4, en utilisant les mêmes indices que pour  $\mathbf{cov}$  :

$$\begin{aligned} \widehat{V}_{\mathbf{cov}}(s, t, s', t') &= \frac{1}{N} \sum_{i=1}^N y_{is} y_{it} y_{is'} y_{it'} \\ &\quad - \left[ \frac{1}{N} \sum_{i=1}^N y_{is} y_{it} \right] \left[ \frac{1}{N} \sum_{i=1}^N y_{is'} y_{it'} \right] \end{aligned} \quad (16)$$

### 3.4 Evolutions des variances et corrélations des salaires

Les tableaux reportant ces évolutions sont assez difficile à lire et il est beaucoup plus parlant de les représenter sous forme de graphiques. On représente d'abord dans la Figure 1, les variances à

chaque période, ainsi que leurs intervalles de confiance à 95%. Ces variances décrivent l'évolution des inégalités intracohorte des salaires au cours de la carrière. Les corrélations estimées à partir des résidus de salaires définis par l'équation (13) sont reportées dans le Tableau 3 et dans les Figures 2 et 3. Le tableau peut être lu de deux façons : en l'étudiant par ligne ce qui revient à raisonner par rapport à une année donnée, auquel cas on examine les corrélations avec les résidus des autres années et à des ordres différents ; en l'examinant en diagonale, c'est à dire à ordre de retard constant, ce qui permet d'examiner l'évolution de la persistance de court terme des résidus de salaire au cours de la carrière. Les Figures 2 et 3 offrent ces représentations, plus parlante nous l'espérons, que les résultats présentés dans le tableau 3. Les Figures 2 reportent les auto-corrélations à différents ordres pour certaines années d'observation prises régulièrement (tous les 5 ans : 1982, 1987, 1992 et 1997) sur la période 1976 à 1998 (le chiffre entre parenthèses correspond au nombre d'années séparant 1976 de l'année servant de base pour représenter les corrélations). Les Figures 3 reportent quand à elles les auto-corrélations à des retards compris entre 1 et 6 pour toutes les années de la période (le chiffre entre parenthèses correspond au retard pour lesquelles les auto-corrélations sont estimées).

Les résultats peuvent être résumés en trois grands traits. Tout d'abord, la variance des salaires augmente au cours du temps comme le montre la Figure 1. Cette évolution a une forme concave, à deux exceptions près au début de la période. Les années 1977 et 1978 se distinguent par des variances très importantes et nous n'avons pas d'explication convaincante à ce phénomène. Après 1993, la croissance des variances s'arrête. On sait par d'autres indices que l'année 1993 se caractérise dans les données par des erreurs de mesure importantes (Le Minez et Roux, 2002) mais l'arrêt de croissance se vérifie les années suivantes. En supposant que les erreurs de mesure sont suffisamment faibles pour que l'on puisse conclure que la variance augmente jusqu'en 1993 puis reste constante, cela voudrait dire que la croissance de l'inégalité au sein d'une cohorte cesse 17 ans après l'entrée sur le marché du travail, c'est à dire vers l'âge de 35 ans, donc avant la phase de ralentissement de la croissance de la moyenne des salaires (20-25 ans après l'entrée ou vers 40 ans, Koubi, 2004). Cette évolution n'est pas compatible avec une représentation du processus de salaire sous la forme d'une marche aléatoire selon laquelle la variance augmenterait indéfiniment.

Deuxièmement, les résidus sont fortement persistants. Leur corrélation est égale à 0.8 à un an d'intervalle à la fin des années 1990 mais reste égale à 0.2 à 22 ans d'écart entre 1976 et 1998, évidence empirique d'effets individuels inobservables persistants dans le cadre choisi ici<sup>4</sup>.

Troisièmement, l'auto-corrélation des résidus augmente au cours du temps. La première sur-

---

<sup>4</sup>Cette forte corrélation pourrait également être le signe de l'existence de racine unitaire, le processus de formation des salaires suivant alors une marche aléatoire, mais cette interprétation n'est pas confirmée par l'examen de l'évolution des variances des résidus.



diagonale du Tableau 3 en est une illustration puisqu'à l'ordre 1, l'auto-corrélation augmente de 0.2 à 0.8. La Figure 3(1) reporte la même information et montre en outre que ces coefficients sont très précisément estimés. Les Figures 3(2-6) montrent que ce phénomène n'est pas limité à l'ordre 1 mais qu'il s'étend au moins jusqu'à l'ordre 6. Dans une dimension différente qu'est l'auto-corrélation à période fixée et non plus à ordre donné, cela donne les allures caractéristiques des auto-corrélations des figures 2 aux périodes 1978, 1982, 1987, 1992 et 1997. Une croissance linéaire de l'auto-corrélation depuis le début de la période d'observation en 1976 jusqu'à la date d'analyse puis une décroissance lente de l'auto-corrélation jusqu'à la dernière période d'observation. L'interprétation en est claire. L'inégalité des salaires semble se figer au cours du temps puisqu'elle est de plus en plus persistante. Cette évolution est compatible avec l'idée d'une convergence des salaires vers un processus stationnaire. En début de carrière, les conditions initiales limiteraient les inégalités salariales. Puis, les inégalités de trajectoire se cristalliseraient au fur et à mesure du déroulement de la carrière. Les modèles que nous allons estimer se doivent de représenter de manière adéquate ces faits stylisés.

## 4 Résultats

Nous avons estimé par maximum de vraisemblance les modèles ARMA décrits dans la Section 2.2 en utilisant l'échantillon complet (5236 observations) et le sous échantillon composé des hommes de moindre qualification uniquement (2340 observations). Comme nous l'avons expliqué, nous tenons compte de la structure compliquée des valeurs manquantes.

### 4.1 Choix de modèles

Les valeurs de la vraisemblance pour de modèles ARMA( $p, q$ ) à différents ordres  $p = 1$  à 4 et  $q = 1$  à 3 dans l'échantillon principal sont reportées dans le Tableau 4. Le nombre de degrés de liberté de chaque modèle  $y$  est aussi reporté.

Les résultats sont peu ambigus. Alors que les niveaux de vraisemblance atteints sont significativement (au sens des tests de rapport de vraisemblance pour les modèles emboîtés, cf. la statistique de test définie par l'équation (11)) croissants avec  $p$  ou  $q$ , le niveau atteint est d'autant plus élevé que  $p$  l'est. Quand  $p = 3$ , un test formel de  $q = 2$  contre l'alternative  $q = 3$  ne permet pas de rejeter l'hypothèse nulle à tout niveau usuel de test. De plus, un modèle ARMA(2,3) est sans ambiguïté rejeté au profit d'un ARMA(3,2) par les critères d'Akaike ou BIC. Dans l'échantillon global, le modèle ARMA(4,2) semble donc être le plus à même de décrire les données.

Un autre élément pouvant intervenir dans le choix du modèle concerne la stabilité de la dynamique sur la période considérée. Nous calculons la statistique de test définie dans l'équation (12) pour chaque modèle estimé. L'hypothèse d'absence de rupture est dans tous les cas rejetée. Toutefois, la statistique de test est la plus faible pour le modèle décrivant le mieux les données, à savoir un modèle ARMA(4,2) pour l'ensemble de la population.

Les résultats, dans le sous-échantillon des individus les moins éduqués, dont sont reportés dans le Tableau 5. Il est impossible de rejeter l'hypothèse nulle  $q = 1$  contre  $q = 2$  quand  $p \geq 3$ . Comme dans l'échantillon complet, les gains de vraisemblance sont les plus importants lorsqu'on augmente l'ordre AR.

De manière générale, on voit que la dynamique est beaucoup plus longue que celle qui est typiquement avancée dans la littérature. Or notre analyse porte sur un échantillon qui semble plus homogène que celui que ceux habituellement utilisés et les hypothèses que nous faisons sur les variances sont plus générales. Les biais conduisant à trouver des longueurs de retard trop importantes dans les modèles ARMA à cause de mauvaises spécifications devraient donc être probablement plus faibles.

Il faut cependant regarder de façon détaillée l'estimation des coefficients dans chaque modèle pour se faire une idée d'éventuelles sources de mauvaise spécification.

## 4.2 Coefficients estimés

Ce sont les estimations pour l'ensemble de la population qui sont principalement exposées dans cette section. Les coefficients décrivant la dynamique des salaires, puis les conditions initiales et les écarts types des innovations, font l'objet des commentaires. Les résultats obtenus sur la sous-population des moins éduqués sont présentés à la fin de cette section. Ils illustrent certaines difficultés inhérentes à la méthode d'estimation.

### **Coefficients des processus ARMA**

Les résultats des différents modèles estimés sur l'ensemble de la population sont présentés dans le Tableau 6. Nous commençons par les résultats portant sur les coefficients autorégressifs puis nous abordons l'analyse des coefficients estimés de la partie moyenne mobile. Nous analysons les traits communs à l'ensemble des modèles estimés, avec une mise en perspective par rapport à la littérature.

Le premier coefficient autorégressif est dans tous les cas supérieur à 0.56 (excepté 0.39 pour l'ARMA(2,3)) et toujours inférieur à 0.98. La stationnarité de la série peut être appréciée en examinant la somme des coefficients AR. Cette somme est comprise entre 0.839 et 0.888 et peut

s'interpréter comme un effet de long terme. Sa valeur élevée reflète une très forte persistance des chocs mais n'est pas suffisamment proche de 1 pour penser qu'une racine unitaire est présente. Le processus est bien stationnaire et une statistique formelle de test est présentée dans le tableau pour démontrer l'absence de racine unitaire. Ce résultat correspond à celui obtenu par Alvarez et Arellano (2004) sur des données espagnoles ou américaines (PSID). Nos coefficients sont légèrement plus élevés.

La comparaison des différents modèles montre que les valeurs des coefficients dépendent grandement de l'ordre de l'ARMA estimé. Cela illustre l'importance de considérer une dynamique d'ordre suffisamment important pour décrire le processus. Ainsi, l'effet du deuxième retard autorégressif (AR) dépend beaucoup de la considération ou non d'un deuxième retard dans la partie moyenne mobile (MA). Alors que le deuxième retard AR est négatif dans les modèles où  $q = 1$ , il est positif et très significatif dans les modèles où le retard de la moyenne mobile est supérieur ou égal à 2. Ceci pourrait expliquer pourquoi le modèle préféré dans la littérature est ARMA(1,2) si les auteurs n'ont pas testé la présence de termes supérieurs à cause du manque de précision ou de biais à distance finie de méthodes GMM par rapport aux méthodes de vraisemblance (Alvarez et Arellano, 2004).

Le coefficient MA(1) est toujours positif et varie entre .15 et .57 (à l'exception du modèle ARMA(2,3) pour lequel le coefficient est négatif). Néanmoins, comme les modèles où  $q = 1$  sont toujours rejetés contre les modèles  $q = 2$  et que ces derniers ne sont pas rejetés contre  $q = 3$ , il semble plus raisonnable de ne regarder que les coefficients estimés pour  $q = 2$ . Dans ce cas, la variabilité des coefficients MA(1) est réduite mais reste conséquente, entre 0.15 et 0.46. Les termes supérieurs ( $q = 3$ ) sont à la limite de la significativité, ou la perdent, dès que le retard de la partie autorégressive dépasse 3.

Le choix de modèles a montré sans ambiguïté que faire croître l'ordre de la partie autorégressive s'adaptait mieux aux données que faire croître les retards dans la partie moyenne mobile quand on utilise des procédures de sélection de modèles soit par emboîtement soit par des critères d'Akaike et de BIC. Or, quand on compare le modèle ARMA(4,2), qui semble préférable à un modèle plus parcimonieux ARMA(3,2), il apparaît que le coefficient du terme autorégressif d'ordre 4,  $\alpha_4$ , n'est pas significatif. Ce qui change principalement entre ces deux ensembles de résultats est la structure de corrélation entre les conditions initiales et le facteur individuel. Ce qui semble donc rejeté par un test formel est, non pas l'ordre du processus mais la structure de corrélation initiale qu'un ARMA(3,2) impose. Il devrait donc suffire de relâcher les contraintes sur les corrélations initiales pour que l'ordre  $p = 3$  devienne plus acceptable. On notera que l'on ne peut pas faire le même raisonnement en comparant le modèle ARMA(3,2) aux modèles plus parcimonieux concurrents de

type ARMA(3,1) puisque le coefficient  $\psi_2$  est significatif dans (3,2) ou à ceux de type ARMA(2,2) puisque le coefficient  $\alpha_3$  est significatif dans (2,2). Néanmoins, le modèle ARMA(4,2) s'adapte le mieux aux données et prend en compte de façon la plus satisfaisante les conditions initiales. Il ne faut simplement pas oublier que la dynamique stationnaire limite de ce processus est plutôt de type ARMA(3,2).

### Composante permanente et conditions initiales

Nous avons vu que les conditions initiales sont un élément important caractérisant le modèle. Elles sont ici modélisées par un vecteur de variables aléatoires correspondant aux valeurs antérieures à la date du début du processus. Ainsi, les variables  $(y_{i,1-p}, \dots, y_{i0})$  ne suivent pas un processus restreint. De même, les chocs  $(\zeta_{i,1-q}, \dots, \zeta_{i0})$ , qui s'interprètent comme les innovations antérieures au début du processus peuvent être corrélées avec les valeurs initiales de  $y$ , de façon très libre. La seule restriction sur ces chocs est qu'ils sont indépendants entre eux et non corrélés aux chocs antérieurs.

La corrélation entre  $y_0$  et le choc  $\zeta_0$  du processus bruit blanc générateur des données (ou entre  $y_{-1}$  et  $\zeta_{-1}$ ) est extrêmement forte, quoique significativement inférieure à 1. Cela signifie que les individus bénéficiant d'un tirage favorable de leurs valeurs initiales de salaires  $y_0$  ont un tirage de leur composante transitoire  $\zeta_0$  également élevé, ce qui affecte négativement leur salaire dans le court terme (correspondant au signe négatif du MA, cf. équation 9). En effet, dans toutes les estimations, le coefficient MA(1) estimé est positif et le plus souvent significatif. Ainsi, un choc  $\zeta_{it}$  élevé se traduit à la période suivante par un effet de retour à la moyenne, plus ou moins important selon le niveau du coefficient MA(1),  $\psi_1$ . Cette très forte corrélation entre ces salaires explique le phénomène de croissance de la variance des salaires au cours du temps. Ces conditions initiales conduisent à avoir une variance relativement faible au début du processus, dans la mesure où la composante cumulative AR s'oppose à l'effet de la composante transitoire MA. Au fur et à mesure de l'évolution de la carrière, le processus converge vers ses valeurs stationnaires. Ainsi, quel que soit le processus considéré, la corrélation entre  $y_t$  et  $\zeta_t$ , diminue pour converger vers une valeur stationnaire proche de 0.5<sup>5</sup>.

Les conditions initiales renseignent aussi sur la corrélation de l'effet persistant  $\eta_i$  avec les valeurs initiales du processus,  $(y_0, \dots, y_{-p})$ . Dans tous les cas, ces corrélations sont positives significatives, entre 0.1 et 0.3. Le modèle préféré est aussi un de ceux qui donne la valeur  $\sigma_\eta$  la plus forte, 0.055. Nous reviendrons sur la signification de ces valeurs quand nous aborderons la décomposition de la variance qu'entraîne une telle modélisation.

---

<sup>5</sup>Cette valeur est estimée par le rapport de l'écart-type de l'innovation sur l'écart-type de  $y_t$ ,  $\frac{\sigma_\zeta}{V(y_t)^{1/2}}$ .

### Écart-types des innovations

L'estimation des écart-types des innovations de la partie moyenne mobile sont reportées dans le Tableau 7 pour certains modèles seulement puisque les résultats varient peu. Notons que même si les données sont absentes pour les années 1981, 1983 et 1990, ces paramètres sont identifiés puisqu'ils affectent la vraisemblance globale i.e. à d'autres dates. Les écart-types de ces périodes sont cependant, comme il est attendu, moins bien précisément estimés et ont tendance à s'écarter des valeurs dans les périodes voisines.

Sans prendre en compte les estimations des premières années qui sont relatives aux conditions initiales et non aux erreurs du modèle, on s'aperçoit que ces écart-types sont approximativement constants au cours du temps puisqu'ils varient entre .209 et .245 (à l'exception de 1981 qui est une année manquante et 1979, qui correspond à la première année effective d'estimation) pour le modèle ARMA(3,2). Si on teste l'absence de variabilité temporelle de cet écart-type, elle est rejetée principalement à cause de la très grande précision de l'estimation de certains paramètres comme l'écart-type en 1993. C'est d'ailleurs celui-ci qui est le plus élevé, une confirmation de l'importance des erreurs de mesure cette année-là (Roux, 2003). Mais il ne semble pas y avoir d'effet temporel marqué et qui suive, en particulier, le rythme des affaires.

### Résultats pour les moins éduqués

L'interprétation des résultats pour le sous-échantillon des moins éduqués est assez similaire. Les résultats, présentés dans le tableau 8, sont cependant moins significatifs.

Tout d'abord, l'effet de long terme est un peu plus élevé que dans l'échantillon global puisqu'il peut atteindre la valeur de 0.94 (ARMA(2,2)), ce qui est une valeur exceptionnelle par rapport aux autres processus pour qui elles restent comprises entre 0.83 et 0.90. Dans cet échantillon, il semble y avoir deux racines de vraisemblance dont l'une n'apparaît que dans le modèle (2,2) et, dans une moindre mesure, dans le modèle (2,1). Il nous a été impossible de la retrouver dans des modèles d'ordre inférieur en faisant varier la valeur initiale des paramètres. Néanmoins il est possible de trouver un maximum local pour un ARMA(2,2) de vraisemblance significativement plus faible et dont les résultats restent proches des autres colonnes. Cette solution ne permet pas d'ailleurs de rejeter la validité du modèle ARMA(2,1).

Le *maximum maximorum* dans le cas ARMA(2,2) se caractérise par des coefficients élevés et très significatifs des parties autorégressive et moyenne mobile aux ordres 1. Le coefficient AR est ainsi supérieur à 1 et celui du MA est proche de 1. Les estimations de ce processus se caractérisent aussi par un niveau de variance de l'effet individuel pratiquement deux fois plus faible que pour les autres processus. Cet effet peut être illusoire puisque nous verrons que la décomposition de la

variance en partie individuelle et en partie processus stationnaire dépend de l'inverse de la somme des coefficients autorégressifs et de la variance de l'effet individuel. Néanmoins, ce maximum se distingue de l'autre par une caractéristique très peu convaincante. Les coefficients  $\alpha_2$  et  $\psi_2$  ne sont pas significativement différents de zéro. Cela implique qu'une représentation ARMA(1,1) pourrait aussi bien décrire ce processus que celle qui est estimée. Surtout, le test de racine unitaire ne rejette plus dans ce cas l'hypothèse d'égalité à 1 de la somme des coefficients de l'AR. Or, il s'agit du seul cas dans l'ensemble des estimations où ce test est accepté.

Pour toutes ces raisons, comme dans l'ensemble de la population, nous retenons comme processus décrivant le mieux la dynamique des salaires celui correspondant à l'ARMA(4,2), avec une dynamique stationnaire correspondant plutôt à un ARMA(3,2), comme dans l'ensemble de la population. Le reste des coefficients est relativement comparable entre les échantillons et les variances estimées qui ne sont pas reportées ici sont en particulier très peu différentes, surtout pour les dernières années. L'hétéroscédasticité en fonction du niveau d'éducation ne semble pas être un phénomène majeur dans cette cohorte.

### 4.3 Fidélité des prédictions

Les graphes de variances et d'autocorrélations reportent les prédictions obtenues dans l'échantillon global, en utilisant le modèle retenu, ARMA(4,2).

On s'aperçoit d'abord dans la Figure 1 que les processus considérés ne permettent pas de saisir complètement les fortes variances des années 1977 et 1978, de même que la tendance à la hausse des variances des salaires au cours du temps. Les années 81 et 83 ne sont pas directement observées dans les données. C'est pourquoi les variances estimées ces années-là sont significativement plus élevées.

Dans les figures 2, les autocorrélations prédites sont présentées comme dans la partie descriptive. Si on prend comme critère que les autocorrélations prédites appartiennent à l'intervalle de confiance à 95% des autocorrélations observées, les allures sont assez fidèlement prédites quelle que soit l'année considérée.

Si on s'intéresse maintenant à la prédiction des ordres de corrélation dans les Figures 3, on s'aperçoit là que les allures sont respectées mais qu'il arrive parfois que les valeurs prédites n'appartiennent pas aux intervalles de confiance. Le résultat notable est que les courbes prédites prédisent une autocorrélation trop forte. C'est une des conclusions les plus importantes portant sur la mauvaise spécification dans cette analyse.

C'est dans la Figure 4 que nous reportons les graphes de normalité des salaires. L'hypothèse

de normalité n'est visiblement pas acceptée. La distribution semble dissymétrique et les queues de distribution sont semble-t-il plus fines que les queues d'une distribution normale tout au moins depuis les années 1980. <sup>6</sup> La méthode d'estimation est robuste à une telle erreur de spécification mais perd ses propriétés d'efficacité. L'absence de normalité pourrait être un signe de rigidités asymétriques à la hausse et à la baisse sur les évolutions de salaire.

#### 4.4 Décomposition de la variance

Les modèles dynamiques de panel saisissent une trajectoire non stationnaire des données en jouant sur les différences entre les conditions initiales et la solution de long terme ou asymptotique du processus. Cette dernière solution est stationnaire et est donnée par :

$$\alpha(L).y_{it} = \sigma_{\eta}\eta_i + \sigma_t.\psi(L).\zeta_{it}. \quad (17)$$

En décomposant  $y_{it} = y_i + \sigma_t.z_{it}$ , on obtient :

$$\alpha(1).y_i = \sigma_{\eta}\eta_i,$$

$$\alpha(L).z_{it} = \psi(L).\zeta_{it},$$

et la décomposition de variance suivante :

$$V y_{it} = \frac{\sigma_{\eta}^2}{\alpha(1)} + \sigma_t^2 V z_{it}. \quad (18)$$

Nous reportons dans le Tableau 9 cette décomposition de la variance pour les différents modèles estimés en faisant l'hypothèse que  $\sigma_t$  correspond à sa valeur en 1998 (voir tableau 7). Les résultats montrent que la proportion de la variance expliquée par le terme d'hétérogénéité inobservable est environ égale à 60%, un chiffre plus élevé que celui que l'on peut obtenir quand on fait la décomposition simple de la variance "sans modèle" des variables observées et qui est environ égale à 40%. Le rôle des conditions initiales est en effet important puisque la corrélation entre l'hétérogénéité individuelle et les conditions initiales peut s'interpréter comme la proportion de variance des conditions initiales expliquée par le facteur individuel. Ces corrélations étaient de l'ordre de .25 (entre .15 et .30, voir Tableau 6). la part de variance de 40% se situe donc entre la corrélation de long terme (57%) et celle due aux conditions initiales (25%). Le processus des revenus "s'adapte" en effet à l'effet individuel au cours du temps.

---

<sup>6</sup>Ce résultat pourrait provenir de l'apurement données des deux centiles les plus extrêmes de la distribution des salaires.

Ainsi, en début de vie active, au moment où les conditions initiales pèsent fortement sur la dynamique salariale, une plus grande partie de la variance des salaires est liée aux conditions initiales, elles-même corrélées à la fois à la composante fixe inobservable et au processus ARMA. Les conditions initiales affectent le processus ARMA de sorte que sa contribution à la variance est plus faible que dans la situation stationnaire.

## 5 Conclusion

Cet article présente certaines méthodes d'estimation et de caractérisation de processus dynamiques décrivant les carrières salariales. Utilisant l'information apportée par le panel DADS, nous avons estimé diverses caractéristiques de la dynamique des salaires dans la cohorte rentrant en 1976 sur le marché du travail. L'examen de cette cohorte à l'aide de représentations de type ARMA apporte les enseignements suivants.

Tout d'abord, disposant d'une cohorte plus homogène que dans les autres articles de cette littérature, notamment en terme de progression dans la carrière, nous pouvons donner une interprétation aux conditions initiales comme étant celles sous lesquelles les individus observés débutent leur carrière. La disponibilité de ces conditions initiales permet de donner un certain crédit à la représentation adoptée, à savoir que le processus de carrière converge vers une trajectoire de long terme stationnaire. En particulier, la non-stationnarité de ce processus est testée (par un test de racine unitaire) et rejetée dans pratiquement toutes les spécifications examinées. Il n'existe pas non plus de différences trop marquées entre les prédictions du modèle et les observations.

Les processus retenus sont d'ordre plus élevé que dans la littérature, ce qui est peut-être le signe d'une hétérogénéité des processus gouvernant les dynamiques salariales. Une des limites de ce travail consiste en effet à considérer qu'une dynamique singulière suffit à représenter l'hétérogénéité des trajectoires salariales. Une autre limite consiste en le rejet systématique de l'hypothèse d'homogénéité des coefficients ARMA sur la période considérée. Là aussi, cette limite reflète probablement une hétérogénéité sous-jacente des processus en œuvre, peut-être liée à la modification sur la période des règles de formation des salaires.

Ce travail peut être répliqué sur d'autres cohortes et la série temporelle de ces paramètres pour celles-ci peut être étudiée. En particulier, il serait intéressant d'examiner l'évolution au cours du temps de l'effet des conditions initiales sur le processus de carrière. Une fois la représentation de ces carrières stabilisée, elle pourrait être utilisée comme génératrice de données dans le cadre des modèles de micro-simulation qui nécessitent de reconstruire sur une longue période les carrières



salariales.

Il reste plusieurs questions ouvertes. Il pourrait par exemple exister un modèle plus parcimonieux qui permettrait de faire aussi bien en termes de prédiction. L'introduction de modèles non linéaires permettrait peut-être d'améliorer cette représentation. Le développement de méthodes simples d'estimation qui permettrait d'analyser les entrées et sorties du panel permettrait aussi de mieux comprendre le mode de formation des salaires.

D'un point de vue économique, il reste également à construire une représentation théorique compatible avec les processus de carrière obtenus comme résultat de cette analyse. Un article de Postel-Vinay et Turon (2005) est une approche prometteuse en ce sens.

## BIBLIOGRAPHIE

**Abowd, J., et D., Card,** (1989), “On the Covariance Structure of Earnings and Hours Changes”, *Econometrica*, 57 : 411-445.

**Altonji, J.G. et L.M. Segal,** (1996), “Small Sample Bias in GMM estimation of Covariance Structures”, *Journal of Business & Economic Statistics*, 14(3) : 353-65.

**Alvarez, J., M., Arellano,** (2004), “Robust Likelihood Estimation of Dynamic Panel Data Models”, Working paper 0421, CEMFI.

**Alvarez, J., M., Browning et M., Ejrnaes,** (2002), “Modelling Income Processes with Lots of Heterogeneity” , Document de travail, Institute of Economics, Copenhagen.

**Baker, M.,** (1997), “Growth Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings”, *Journal of Labor Economics*, 15 : 338-75.

**Baker, M., et G., Solon,** (1999), “Earnings Dynamics and Inequality Among Canadian Men, 1976-1992”, NBER WP 7370.

**Bayet, A.** (1996), “Carrières continues, carrières incomplètes et salaires”, *Economie et Statistique*, 299 : 21-36.

**Bonhomme, S., et J.M., Robin,** (2004), “Modeling Individual Earning Trajectories Using Copulas with an Application to the Study of Earnings Inequality : France, 1990-2002.”, unpublished manuscript.

**Buchinsky, M., D.,Fougère et F., Kramarz,** (1999), “La mobilité salariale en France : 1967-1986”, *Revue Économique*, 49 : 879-90.

**Geweke, J., et M., Keane,** (2000), “An empirical analysis of earnings dynamics among men in the PSID : 1968-1989”, *Journal of Econometrics*, 96 : 293-356.

**Goux, D., et E., Maurin,** (1994), “Education, expérience et salaire : tendances récentes et évolutions de long terme”, *Économie et Prévision*, 116 : 155-178.

**Gottschalk,P.,** (1997), “Inequality, Income Growth and Mobility : The Basic Facts”, *Journal of Economic Perspectives*, 11(2) : 21-40.

**Gottschalk, P., et R., Moffitt,** (1994), “The Growth of Earnings Instability in the US Labor Market”, *Brookings Papers on Economic Activity*, 2 : 217-72.

**Gouriéroux, C., et A., Monfort,** (1989), *Statistique des modèles économétriques*, Economica : Paris.

**Guillot Y. et P. Sevestre,** (1994), “Estimations de fonctions de gains sur données de panel : endogénéité du capital humain et effets de la sélection”, *Économie et Prévision*, 116 : 119-35.

**Hause, J.C.**, (1980), “The Fine Structure of Earnings and On-the-job Training Hypothesis”, *Econometrica*, 48(4) : 1013-29.

**Hirano, K.** (2002), “Semiparametric Bayesian Inference in Autoregressive Panel Data Models”, *Econometrica*, 70.2 : 781-799.

**Horowitz, J.**, (1998), “Bootstrap Methods for Covariance Structures”, *Journal of Human Resources*, 33 : 39-61.

**Horowitz, J., et M., Markatou**, (1996), “Semiparametric Estimation of Regression Models for Panel Data”, *Review of Economic Studies*, 63 : 145-68.

**Kezdi, G., J., Hahn et G., Solon**, (2002), “Jackknife Minimum Distance Estimation” , *Economic Letters*, 76 : 35-45.

**Koubi, M.**, (2004), “Les carrières salariales par cohorte entre 1967 et 2000”, *Économie et Statistiques*, 369-70 :149-71.

**Le Minez, S., et S., Roux**, (2002), “Les différences de carrières salariales à partir du premier emploi”, *Économie et Statistiques*, 351 : 31-63.

**Lillard, L. et R., Willis**, (1978), “Dynamic Aspects of Earnings Mobility”, *Econometrica*, 46 : 985-1012.

**Lollivier, S., et J.F., Payen**, (1990), “L’hétérogénéité des carrières individuelles mesurée sur données de panel”, *Économie et Prévision*, 92-93 : 87-95.

**MaCurdy, T.**, (1982), “The Use of Time Series Processes to Model Structure of Earnings in a Longitudinal Analysis”, *Journal of Econometrics*, 18 : 83-114.

**Meghir, C., et L., Pistaferri**, (2004), “Income Variance Dynamics and Heterogeneity”, *Econometrica*, 72 : 1-32.

**Piketty, T.**, (2001), *Les hauts revenus en France au XX<sup>e</sup> siècle*, Grasset : Paris.

**Postel-Vinay, F. et H., Turon**, (2005), “On-the-job Search, Productivity Shocks and the Individual Earnings Process”, PSE Working papers, 2005-31.

**Ramalho, J.J.S.**, (2005), “Small Sample Bias of Alternative Estimation Methods for Moment Condition Models : Monte Carlo Evidence for Covariance Structures and Instrumental Variables”, *Studies in Nonlinear Dynamics and Econometrics*, Vol. 9, No. 1, Article 3.

**Roux, S.**, (2003), “Refonte du panel DADS : principes et premières estimations d’emploi et de salaire”, Document de travail INSEE.

# A Spécification de la vraisemblance.

## A.1 Le modèle

On développe ici de manière plus précise que dans le texte le modèle utilisé par Alvarez et Arellano (2004) :

$$y_{it} = \alpha_1 y_{i(t-1)} + \dots + \alpha_p y_{i(t-p)} + \sigma_\eta \eta_i + v_{it}$$

où  $v_{it}$  est un terme défini plus bas. Cette équation est supposée valide de  $t = 1$  à  $T$  seulement et nous supposons par simplicité que  $p \geq 1$  et  $q \geq 1$ .

Soit  $\alpha = (\alpha_1, \dots, \alpha_p)$  et  $M_T(\alpha)$  une matrice de dimension  $[T, T + p]$  où  $p = \dim(\alpha)$  :

$$M_T(\alpha) = \begin{pmatrix} -\alpha_p & \dots & -\alpha_1 & 1 & 0 & \dots & 0 \\ 0 & -\alpha_p & \dots & -\alpha_1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & -\alpha_p & \dots & -\alpha_1 & 1 \end{pmatrix}.$$

On note  $y_i^{[1,T]} = (y_{i1}, \dots, y_{iT})$  et  $y_i^{[1-p,0]} = (y_{i(1-p)}, \dots, y_{i0})$ . Nous avons :

$$\begin{pmatrix} (I_p \ 0) \\ M_T(\alpha) \end{pmatrix} \begin{pmatrix} y_i^{[1-p,0]} \\ y_i^{[1,T]} \end{pmatrix} = \begin{pmatrix} y_i^{[1-p,0]} \\ \sigma_\eta \eta_i \cdot j_T + v_i^{[1,T]} \end{pmatrix}$$

où  $j_T$  est le vecteur de dimension  $T$  dont tous les éléments sont égaux à 1 et où  $v_i^{[1,T]} = (v_{i1}, \dots, v_{iT})$ .

De façon équivalente :

$$\begin{pmatrix} I_p & 0 \\ M_T(\alpha) \end{pmatrix} \cdot y_i^{[1-p,T]} = \begin{pmatrix} y_i^{[1-p,0]} \\ \sigma_\eta \eta_i \cdot j_T + v_i^{[1,T]} \end{pmatrix} \quad (19)$$

Si :

$$v_{it} = \sigma_t w_{it}$$

où  $w_{it}$  est une moyenne mobile d'ordre  $q$  :

$$w_{it} = \zeta_{it} - \psi_1 \zeta_{it-1} - \dots - \psi_q \zeta_{it-q}.$$

On peut écrire

$$w_i^{[1,T]} = M_T(\psi) \cdot \zeta_i^{[1-q,T]}$$

où  $\zeta_i^{[1-q,T]} = (\zeta_{i1-q}, \dots, \zeta_{iT})$ .

On note  $\Lambda$  une matrice carrée dont la diagonale est  $(\sigma_1, \dots, \sigma_T)$  de telle manière que :

$$v_i^{[1,T]} = \Lambda \cdot M_T(\psi) \cdot \zeta_i^{[1-q,T]}.$$

En remplaçant dans (19) on obtient :

$$\begin{pmatrix} I_p & 0 \\ M_T(\alpha) \end{pmatrix} \cdot y_i^{[1-p,T]} = \begin{pmatrix} I_p & 0_{p,1} & 0_{p,T+q} \\ 0_{T,p} & \sigma_\eta \cdot j_T & \Lambda \cdot M_T(\psi) \end{pmatrix} \begin{pmatrix} y_i^{[1-p,0]} \\ \eta_i \\ \zeta_i^{[1-q,T]} \end{pmatrix}. \quad (20)$$

Nous allons maintenant construire la matrice de variance-covariance de la variable observée  $y_i^{[1-p,T]}$ . On construit d'abord la matrice de variance-covariance de  $\begin{pmatrix} y_i^{[1-p,0]} & \eta_i & \zeta_i^{[1-q,T]} \end{pmatrix}$ . Comme  $\eta_i, \zeta_i^{[1-q,T]}$  sont identiquement et indépendamment distribués et ont une variance unitaire, le coin sud-est de cette matrice est la matrice identité de taille  $(1+q+T)$ . Le coin nord-est est une matrice non restreinte  $V y_i^{[1-p,0]} = \Gamma_{00}$  comme l'est  $E y_i^{[1-p,0]} \eta_i = \gamma_{0\eta}$ . De plus, si nous supposons que  $E(y_{i\tau} \zeta_{it}) = 0$  pour tout  $\tau < t$ ,  $E(y_i^{[1-p,0]} \cdot (\zeta_i^{[1,T]})') = 0$ . La seule inconnue est alors  $E(y_i^{[1-p,0]} \cdot (\zeta_i^{[1-q,0]})')$ .

Écrivons :

$$E(y_i^{[1-p,0]} \cdot (\zeta_i^{[1-q,0]})') = \Omega = [\omega_{rs}]$$

où  $r \in [1-p, 0]$  and  $s \in [1-q, 0]$  :

$$\begin{aligned} r < s : & \quad \omega_{rs} = 0 \\ r \geq s : & \quad \omega_{rs} \text{ est non contrainte} \end{aligned}$$

car l'innovation  $\zeta_i^s$  est postérieure à  $r$  et n'est pas corrélée à  $y_i^r$ .

En reprenant les notations du texte :

$$\Omega_z = V \begin{pmatrix} y_i^{[1-p,0]} \\ \eta_i \\ \zeta_i^{[1-q,T]} \end{pmatrix} = V \begin{pmatrix} y_i^{[1-p,0]} \\ \eta_i \\ \zeta_i^{[1-q,0]} \\ \zeta_i^{[1,T]} \end{pmatrix} = \begin{pmatrix} \Gamma_{00} & \gamma_{0\eta} & \Omega & 0 \\ \gamma'_{0\eta} & 1 & 0 & \vdots \\ \Omega' & 0 & \ddots & 0 \\ 0 & \dots & 0 & I_T \end{pmatrix}.$$

La matrice  $\Sigma$  qui, dans le corps du texte, correspond à la matrice de variance-covariance des conditions initiales s'écrit ainsi :

$$\Sigma = V \begin{pmatrix} y_i^{[1-p,0]} \\ \eta_i \\ \zeta_i^{[1-q,0]} \end{pmatrix} = \begin{pmatrix} \Gamma_{00} & \gamma_{0\eta} & \Omega \\ \gamma'_{0\eta} & & I_T \\ \Omega' & & \end{pmatrix}.$$

On a également,

$$\begin{aligned} R_T(\alpha) &= \begin{pmatrix} (I_p & 0) \\ M_T(\alpha) \end{pmatrix}^{-1} \\ S_{T,p}(\psi, \Lambda, \sigma_\eta) &= \begin{pmatrix} I_p & 0_{p,1} & 0_{p,T+q} \\ 0_{T,p} & \sigma_\eta \cdot j_T & \Lambda \cdot M_T(\psi) \end{pmatrix} \end{aligned}$$

et nous avons :

$$\Omega_y = V \left( y_i^{[1-p,T]} \right) = R_T(\alpha) \cdot S_{T,p}(\psi, \Lambda, \sigma_\eta) \cdot \Omega_z \cdot S_{T,p}(\psi, \Lambda, \sigma_\eta)' R_T(\alpha)'$$

## A.2 Paramétrisation

### A.2.1 Corrélations avec les conditions initiales

Les conditions initiales sont données par  $\zeta_i^{[1-q,0]}$ ,  $\eta$  et  $y_i^{[1-p,0]}$ . En utilisant la décomposition de Choleski (*i.e.* l'équivalent de la racine carrée matricielle étant une matrice triangulaire inférieure  $C$  telle que  $C'C = \Sigma$ ), et en normalisant la variance de chacune des composantes à 1, la matrice  $\Sigma$  peut être paramétrée à partir de la matrice suivante dans un système de coordonnées polaires.

$$\text{Transf} \left( V \begin{pmatrix} \zeta_{i0} \\ \vdots \\ \zeta_{i,1-q} \\ \eta_i \\ y_{i,1-p}/\sigma_{1-p} \\ \vdots \\ y_{i,0}/\sigma_0 \end{pmatrix} \right) = \begin{pmatrix} 1 & 0 & \dots & & & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & 0 & 1 & 0 & \ddots & & \\ & \dots & 0 & 1 & 0 & \ddots & \\ \vdots & & \theta_{1-q,1-p}^{(1)} & \theta_{\eta,1-p} & \theta_{1-p,1-p} & 0 & \ddots \\ 0 & & \vdots & \vdots & & \ddots & 0 \\ \theta_{0,0}^{(1)} & & \theta_{\eta,0} & \dots & & & \theta_{0,0} \end{pmatrix}$$

où  $\theta_{1-q,1-p}^{(1)} = 0$  si  $p > q$  et plus généralement  $\theta_{l,m}^{(1)} = 0$  si  $l > m$ .

### A.2.2 Racines des polynômes retard

Soient  $\lambda_1$  et  $\lambda_2$  deux racines hors du cercle unité qui sont réelles ou complexes conjuguées. Les représentations AR ou MA sont alors les produits d'éléments simples :

$$\begin{aligned} (1 - \lambda_1 L)(1 - \lambda_2 L) &= 1 - (\lambda_1 + \lambda_2)L + \lambda_1 \lambda_2 L^2 \\ &= 1 - \alpha_1 L - \alpha_2 L^2 \end{aligned}$$

où  $\alpha_1 = \lambda_1 + \lambda_2$  and  $\alpha_2 = -\lambda_1 \lambda_2$ . Comme les racines sont hors du cercle unité, cela est équivalent à :

$$\begin{aligned} \alpha_2 &\in ]-1, 1[ \\ | \alpha_1 | &< 1 - \alpha_2 \end{aligned}$$

C'est cette paramétrisation que nous retenons pour la modélisation.

TAB. 1 – Composition de l'échantillon en termes d'âge à l'entrée et de la nature du premier emploi.

Âge	Nombre	Fréquence (%)
16	344	6,57
18	2 914	55,65
20	960	18,33
22	479	9,15
24	250	4,77
26	139	2,65
28	71	1,36
30	34	0,65
32	18	0,34
34	27	0,52

Premier emploi occupé	Nombre	Fréquence(%)
Très qualifié	376	7,18
Qualifié	2 520	48,13
Non qualifié	2 340	44,69
Total	5 236	100,00

TAB. 2 – Structure des valeurs manquantes de la variable dépendante en fonction des périodes.

Périodes	1976	1977	1978	1979	1980	1982	1984	1985	1986	1987	1988	1989	1991	1992	1993	1994	1995	1996	1997	1998
1976	100.00	92.82	85.31	85.54	84.34	84.74	83.63	82.89	84.24	84.21	83.82	84.72	86.10	87.49	86.77	82.91	89.15	91.50	93.51	100.00
1977		92.82	81.13	80.52	79.43	79.49	78.44	77.60	78.76	78.72	78.44	79.20	80.20	81.40	80.81	77.24	82.91	85.12	86.94	92.82
1978			85.31	76.70	74.71	74.20	73.03	72.15	73.09	73.13	72.96	73.51	74.20	75.42	74.64	71.35	76.60	78.53	80.14	85.31
1979				85.54	77.58	75.90	74.20	73.68	74.47	73.93	73.26	74.05	74.83	75.92	75.15	71.33	76.89	78.82	80.50	85.54
1980					84.34	76.30	74.12	73.42	74.31	73.82	73.05	73.55	74.52	75.38	74.58	70.74	75.96	77.77	79.30	84.34
1982						84.74	76.60	75.15	75.96	75.57	74.64	74.89	75.52	76.36	75.34	71.77	76.78	78.57	80.16	84.74
1984							83.63	77.12	77.12	76.18	75.27	75.27	75.31	75.97	74.92	71.35	76.43	77.87	79.14	83.63
1985								82.89	77.88	76.43	75.12	75.06	75.06	75.67	74.62	70.55	75.82	77.27	78.46	82.89
1986									84.24	79.14	77.41	76.66	76.60	77.14	75.96	71.93	77.18	78.50	79.87	84.24
1987										84.21	78.93	77.35	76.89	77.16	75.92	71.75	77.12	78.61	79.81	84.21
1988											83.82	78.48	77.20	77.43	75.86	72.29	77.35	78.44	79.51	83.82
1989												84.72	79.18	78.55	76.95	72.88	78.11	79.30	80.50	84.72
1991													86.10	81.53	78.99	74.79	79.68	80.86	82.16	86.10
1992														87.49	81.26	76.41	81.40	82.56	83.54	87.49
1993															86.77	77.02	81.30	82.22	82.95	86.77
1994																82.91	79.60	79.43	79.62	82.91
1995																	89.15	86.12	85.89	89.15
1996																		91.50	88.35	91.50
1997																			93.51	93.51
1998																				100.00

Note : Ce tableau donne la fréquence d'observation des personnes aux périodes  $t$  et  $t'$ . Ainsi la cellule (1976, 1977) montre que l'on connaît le salaire de 92.8% des personnes de l'échantillon à ces deux dates. De plus, comme toutes les personnes sont présentes en 1976 et en 1998, les cellules (1976, 1976), (1976, 1998) et (1998, 1998) sont égales à 100%.



TAB. 3 – Matrice d'autocorrélation des résidus de salaires.

Périodes	1976	1977	1978	1979	1980	1982	1984	1985	1986	1987	1988	1989	1991	1992	1993	1994	1995	1996	1997	1998		
1976	1.000																					
1977	0.265	1.000																				
1978	0.267	0.287	1.000																			
1979	0.368	0.300	0.268	1.000																		
1980	0.532	0.420	0.391	0.330	1.000																	
1982	0.534	0.430	0.388	0.400	0.334	1.000																
1984	0.629	0.551	0.516	0.484	0.475	0.447	1.000															
1985	0.719	0.662	0.630	0.598	0.563	0.533	0.533	1.000														
1986	0.721	0.670	0.670	0.623	0.593	0.552	0.552	0.552	1.000													
1987	0.759	0.707	0.707	0.660	0.610	0.621	0.621	0.610	0.610	1.000												
1988	0.793	0.710	0.710	0.652	0.652	0.659	0.659	0.652	0.652	0.652	1.000											
1989	0.776	0.776	0.776	0.705	0.702	0.702	0.702	0.705	0.702	0.702	0.702	1.000										
1991	1.000	1.000	1.000	1.000	0.741	0.741	0.741	0.741	0.741	0.741	0.741	0.741	1.000									
1992	0.830	0.830	0.830	0.830	0.731	0.731	0.731	0.731	0.731	0.731	0.731	0.731	0.731	1.000								
1993	0.793	0.793	0.793	0.793	0.777	0.777	0.777	0.777	0.777	0.777	0.777	0.777	0.777	0.777	1.000							
1994	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800	0.800	1.000						
1995	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	0.758	1.000					
1996	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	0.829	1.000				
1997	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	0.831	1.000			
1998	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	0.833	1.000		
1999	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	0.835	1.000	

TAB. 4 – Vraisemblances des modèles : Échantillon complet.

ARMA( $p, q$ )	$q=1$	2	3
$p=1$	75513.74 (28)	75557.25 (30)	75571.69 (32)
2	75588.58 (31)	75592.68 (34)	75703.23 (37)
3	75780.77 (35)	75815.19 (38)	75817.29 (42)
4	76012.31 (40)	76032.90 (43)	76032.97 (47)

Note : La première ligne contient les valeurs de la log-vraisemblance atteinte pour chaque modèle ARMA. La seconde ligne contient, entre parenthèses, les degrés de liberté du modèle.

TAB. 5 – Vraisemblances des modèles : groupe des non qualifiés..

ARMA( $p, q$ )	$q=1$	2	3
$p=1$	33480.62 (28)	33509.59 (30)	33513.89 (32)
2	33524.27 (31)	33535.65 (34)	33543.30 (37)
3	33609.71 (35)	33613.64 (38)	33613.79 (42)
4	33727.70 (40)	33730.69 (43)	33730.71 (47)

Note : cf. table 4.

TAB. 6 – Résultats d'estimation des modèles ARMA, ensemble.

ARMA	1-1	2-1	3-1	4-1	1-2	2-2	3-2	4-2	1-3	2-3	3-3	4-3
$\alpha_1$	0.839 (0.008)	0.977 (0.037)	0.946 (0.049)	0.941 (0.054)	0.869 (0.009)	0.562 (0.110)	0.612 (0.128)	0.787 (0.056)	0.884 (0.009)	0.390 (0.025)	0.599 (0.106)	0.788 (0.053)
$\alpha_2$		-0.095 (0.024)	-0.072 (0.022)	-0.067 (0.024)		0.251 (0.091)	0.268 (0.101)	0.147 (0.048)		0.444 (0.021)	0.265 (0.081)	0.147 (0.052)
$\alpha_3$			-0.007 (0.013)	-0.015 (0.011)			-0.043 (0.013)	-0.050 (0.012)			-0.028 (0.017)	-0.052 (0.055)
$\alpha_4$				0.014 (0.008)				-0.006 (0.011)				-0.004 (0.024)
$\psi_1$	0.445 (0.013)	0.571 (0.036)	0.531 (0.051)	0.497 (0.056)	0.456 (0.012)	0.155 (0.108)	0.185 (0.127)	0.336 (0.057)	0.471 (0.012)	-0.047 (0.029)	0.171 (0.106)	0.336 (0.054)
$\psi_2$					0.051 (0.009)	0.180 (0.047)	0.219 (0.055)	0.162 (0.033)	0.047 (0.009)	0.316 (0.015)	0.209 (0.045)	0.162 (0.039)
$\psi_3$									0.029 (0.007)	0.059 (0.008)	0.013 (0.009)	-0.002 (0.061)
$\sigma_\eta$	0.051 (0.002)	0.039 (0.004)	0.045 (0.005)	0.042 (0.006)	0.043 (0.002)	0.061 (0.006)	0.055 (0.008)	0.042 (0.005)	0.039 (0.002)	0.056 (0.003)	0.056 (0.007)	0.042 (0.006)
$Corr(y_0, \eta)$	0.198 (0.016)	0.148 (0.016)	0.156 (0.017)	0.282 (0.019)	0.184 (0.017)	0.174 (0.025)	0.114 (0.019)	0.283 (0.020)	0.179 (0.017)	0.171 (0.020)	0.114 (0.019)	0.283 (0.022)
$Corr(y_{-1}, \eta)$		0.246 (0.017)	0.276 (0.020)	0.203 (0.021)		0.214 (0.020)	0.253 (0.022)	0.182 (0.022)		0.194 (0.018)	0.252 (0.021)	0.182 (0.024)
$Corr(y_{-2}, \eta)$			0.249 (0.017)	0.275 (0.021)			0.246 (0.017)	0.278 (0.021)		0.244 (0.017)	0.277 (0.025)	0.244 (0.025)
$Corr(y_{-3}, \eta)$				0.251 (0.018)			0.249 (0.018)	0.249 (0.018)				0.249 (0.018)
$Corr(y_0, \zeta_0)$	0.976 (0.004)	0.960 (0.004)	0.947 (0.006)	0.890 (0.008)	0.976 (0.003)	0.957 (0.005)	0.959 (0.005)	0.892 (0.008)	0.965 (0.005)	0.746 (0.025)	0.959 (0.005)	0.892 (0.011)
$Corr(y_0, \zeta_{-1})$					0.097 (0.014)	0.229 (0.016)	0.202 (0.021)	0.258 (0.020)	0.179 (0.024)	0.550 (0.028)	0.205 (0.019)	0.258 (0.069)
$Corr(y_{-1}, \zeta_{-1})$						0.972 (0.005)	0.940 (0.047)	0.949 (0.010)		0.764 (0.033)	0.943 (0.006)	0.948 (0.223)
$Corr(y_0, \zeta_{-2})$									0.044 (0.011)	0.332 (0.029)	-0.156 (0.016)	0.002 (3.512)
$Corr(y_{-1}, \zeta_{-2})$										-0.609 (0.039)	-0.209 (0.016)	-0.029 (3.262)
$Corr(y_{-2}, \zeta_{-2})$											-0.966 (0.005)	0.029 (13.595)
Test de rupture en 1987	210.91	301.92	290.63	42.51	270.96	386.86	254.47	40.49	268.56	302.22	334.03	98.63
Test de racine unitaire	410.15	65.54	52.88	46.38	227.34	72.54	39.99	65.39	171.75	179.17	53.01	32.35

Note :  $\alpha_p$  = coefficients autorégressifs ;  $\psi_q$  = coefficients moyenne mobile ;  $\sigma_\eta$  = écart-type de l'effet individuel ;  $corr(y_{1-p}, \eta)$  = corrélation entre conditions initiales et effet individuel ;  $corr(y_{1-p}, \zeta_{1-q})$  = corrélation entre conditions initiales et bruit blanc.

TAB. 7 – Estimation des écart-types temporels

$\sigma_t$	4-1	3-2	4-2	2-3
1976	0.296*	0.296*	0.296*	0.293*
	(0.004)	(0.004)	(0.004)	(0.004)
1977	0.408*	0.405*	0.408*	0.396*
	(0.009)	(0.010)	(0.009)	(0.008)
1978	0.407*	0.389*	0.405*	0.378
	(0.006)	(0.008)	(0.006)	(0.006)
1979	0.343*	0.328	0.345*	0.319
	(0.005)	(0.009)	(0.005)	(0.006)
1980	0.286	0.287	0.284	0.291
	(0.005)	(0.006)	(0.005)	(0.005)
1981	0.286	0.325	0.306	0.333
	(0.010)	(0.021)	(0.013)	(0.027)
1982	0.211	0.215	0.210	0.231
	(0.013)	(0.019)	(0.012)	(0.008)
1983	0.251	0.282	0.263	0.299
	(0.011)	(0.022)	(0.015)	(0.025)
1984	0.213	0.212	0.211	0.221
	(0.014)	(0.014)	(0.014)	(0.008)
1985	0.233	0.234	0.234	0.236
	(0.006)	(0.006)	(0.006)	(0.007)
1986	0.244	0.242	0.244	0.242
	(0.006)	(0.006)	(0.006)	(0.006)
1987	0.235	0.235	0.236	0.235
	(0.005)	(0.006)	(0.005)	(0.005)
1988	0.221	0.218	0.220	0.220
	(0.005)	(0.005)	(0.005)	(0.005)
1989	0.233	0.235	0.235	0.236
	(0.006)	(0.006)	(0.006)	(0.006)
1990	0.219	0.230	0.203	0.265
	(0.015)	(0.046)	(0.029)	(0.021)
1991	0.228	0.223	0.233	0.219
	(0.010)	(0.015)	(0.013)	(0.006)
1992	0.210	0.204	0.207	0.203
	(0.006)	(0.007)	(0.006)	(0.007)
1993	0.241	0.247	0.244	0.249
	(0.006)	(0.007)	(0.007)	(0.007)
1994	0.237	0.232	0.235	0.229
	(0.006)	(0.007)	(0.006)	(0.006)
1995	0.223	0.223	0.223	0.224
	(0.005)	(0.006)	(0.005)	(0.006)
1996	0.220	0.220	0.221	0.221
	(0.006)	(0.007)	(0.006)	(0.007)
1997	0.217	0.217	0.217	0.218
	(0.005)	(0.006)	(0.006)	(0.006)
1998	0.211	0.209	0.211	0.208
	(0.005)	(0.005)	(0.005)	(0.005)

Note : (\*) Les écart-types ainsi reportés sont ceux des conditions initiales et non ceux des chocs  $w_{it}$ .(voir texte). Leur nombre est égal à  $p$ .

TAB. 8 – Résultats d'estimation des modèles ARMA, Non qualifiés au premier emploi.

ARMA	1-1	2-1	3-1	4-1	1-2	2-2	3-2	4-2	1-3	2-3	3-3	4-3
$\alpha_1$	0.836	1.036	0.987	0.948	0.874	1.352	0.820	0.868	0.885	0.417	0.821	0.867
	0.013	0.043	0.056	0.060	0.013	0.264	0.136	0.077	0.012	0.055	0.154	0.085
$\alpha_2$		-0.134	-0.102	-0.081		-0.409	0.076	0.041		0.411	0.080	0.043
		0.029	0.027	0.029		0.227	0.116	0.074		0.044	0.123	0.073
$\alpha_3$			-0.005	-0.003			-0.029	-0.023			-0.034	-0.025
			0.015	0.013			0.016	0.016			0.023	0.025
$\alpha_4$				0.013			-0.002	-0.002				-0.002
				0.011			0.002	0.002				0.003
$\psi_1$	0.442	0.626	0.566	0.494	0.459	0.972	0.392	0.410	0.470	-0.026	0.393	0.410
	0.020	0.044	0.060	0.062	0.018	0.305	0.140	0.078	0.018	0.063	0.160	0.088
$\psi_2$					0.064	-0.168	0.118	0.097	0.060	0.302	0.123	0.099
					0.013	0.143	0.070	0.050	0.013	0.026	0.070	0.050
$\psi_3$									0.023	0.053	-0.004	-0.002
									0.011	0.011	0.018	0.024
$\sigma_\eta$	0.050	0.032	0.039	0.038	0.040	0.020	0.043	0.037	0.037	0.055	0.043	0.037
	0.004	0.004	0.006	0.007	0.003	0.011	0.008	0.006	0.003	0.005	0.009	0.006
$Corr(y_0, \eta)$	0.182	0.140	0.139	0.262	0.170	0.150	0.108	0.260	0.167	0.184	0.108	0.261
	0.025	0.024	0.024	0.029	0.024	0.030	0.033	0.030	0.024	0.033	0.034	0.030
$Corr(y_{-1}, \eta)$		0.229	0.295	0.179		0.251	0.287	0.167		0.176	0.288	0.167
		0.026	0.031	0.034		0.034	0.031	0.036		0.027	0.031	0.037
$Corr(y_{-2}, \eta)$			0.229	0.298			0.228	0.302			0.229	0.302
			0.026	0.032			0.026	0.032			0.025	0.033
$Corr(y_{-3}, \eta)$				0.238			0.239	0.239				0.239
				0.027			0.027	0.027				0.027
$Corr(y_0, \zeta_0)$	0.982	0.965	0.956	0.910	0.977	0.966	0.962	0.913	0.971	0.744	0.961	0.912
	0.005	0.005	0.008	0.011	0.004	0.005	0.009	0.011	0.006	0.040	0.010	0.012
$Corr(y_0, \zeta_{-1})$					0.115	0.207	0.206	0.263	0.159	0.553	0.209	0.263
					0.020	0.028	0.032	0.030	0.027	0.042	0.030	0.029
$Corr(y_{-1}, \zeta_{-1})$						0.965	0.937	0.956		0.728	0.937	0.954
						0.019	0.009	0.008		0.052	0.009	0.073
$Corr(y_0, \zeta_{-2})$									0.037	0.324	-0.140	0.001
									0.016	0.049	0.025	0.417
$Corr(y_{-1}, \zeta_{-2})$										-0.658	-0.182	-0.053
										0.056	0.044	1.404
$Corr(y_{-2}, \zeta_{-2})$											-0.970	0.078
											0.007	1.933
Test de rupture en 1987	99.07	142.62	135.20	21.10	119.00	118.68	132.29	25.35	135.41	217.48	265.29	32.04
Test de racine unitaire	171.42	38.24	32.36	29.59	101.01	2.24	24.15	32.99	91.06	69.44	21.15	34.67

TAB. 9 – Décomposition de la variance : Échantillon complet

ARMA( $p, q$ )	$q=1$	2	3
$p=1$	0.604 (0.168)	0.613 (0.176)	0.626 (0.183)
2	0.630 (0.180)	0.616 (0.173)	0.626 (0.185)
3	0.626 (0.180)	0.630 (0.182)	0.630 (0.183)
4	0.599 (0.184)	0.612 (0.191)	0.612 (0.190)

Note : Proportion de la variance due à l'hétérogénéité constante *i.e.*  $\sigma_u^2/(\sigma_u^2 + \sigma_\varepsilon^2)$ . L'écart type temporel  $\sigma_t$  est supposé égal au niveau atteint en 1998. Entre parenthèses on reporte la variabilité totale ( $\sigma_u^2 + \sigma_\varepsilon^2$ ).

TAB. 10 – Décomposition de la variance : Non qualifiés au premier emploi occupé.

ARMA( $p, q$ )	$q=1$	2	3
$p=1$	0.567 (0.160)	0.580 (0.170)	0.590 (0.175)
2	0.604 (0.178)	0.656 (0.187)	0.589 (0.174)
3	0.594 (0.176)	0.598 (0.177)	0.598 (0.177)
4	0.549 (0.178)	0.559 (0.182)	0.559 (0.182)

Note : cf. note du tableau 9