ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 20/21 — 1991

Durée du chômage des jeunes en France

Antoine MOREAU, Michael VISSER *

RÉSUMÉ. – Nous estimons un modèle stationnaire de recherche d'emploi sur un échantillon de jeunes chômeurs. Toutes les restrictions de la théorie sont utilisées pour obtenir les estimateurs tant des paramètres structurels (durée moyenne entre deux offres, probabilités d'acceptation) que de la distribution des salaires de réserve. L'estimation d'un modèle de Cox avec un hasard de base constant par morceaux ne permet pas de rejeter l'hypothèse de stationnarité.

Duration of Youth Unemployment in France

ABSTRACT. – We estimate a stationary job search model on a sample of young french unemployed. We give estimates for the structural parameters (mean duration between offers, acceptance probabilities) and the reservation wages distribution, using all the restrictions implied by the search theory. Estimation results of a Cox Model with piece-wise constant hazard don't allow to reject the stationarity assumption.

---

* A. MOREAU : INSEE; M. VISSER : ENSAE-CREST. Nous remercions nos collègues de l'INSEE et de l'ENSAE-CREST pour leurs commentaires et leurs suggestions.

1 Introduction

Ces dernières années, la hausse du chômage s'est accompagnée d'une forte augmentation de la durée moyenne du chômage et du pourcentage des chômeurs de longue durée (ancienneté supérieure à un an). En France, la durée moyenne a augmenté de 7,6 mois en 1975 jusqu'à 15,7 mois en 1986; le pourcentage des chômeurs de longue durée a augmenté de 16,9 % à 46,1 % durant la même période (INSEE [1986, 1987]). Par contre, le nombre d'entrées au chômage rapporté à l'emploi total n'a pas augmenté avec la même ampleur: 0,5 % en 1976 à 0,6 % en 1986. Autrement dit, l'augmentation du taux de chômage résulte principalement des difficultés grandissantes de reclassement. (THÉLOT [1988]). Les épisodes de chômage sont en général plus courts pour les jeunes que pour les personnes plus âgées. Pourtant, l'augmentation de la durée moyenne du chômage et du pourcentage des chômeurs de longue durée est également importante parmi les jeunes. Pour eux, la durée moyenne du chômage était égale à 6,0 mois en 1975 et 11,4 mois en 1986, tandis que le pourcentage de chômeurs de longue durée valait 10,8 % en 1975 et 34,8 % en 1986 (INSEE [1986, 1987]).

Nous essayons ici d'analyser et d'estimer les déterminants de la durée du chômage des jeunes en France. En particulier, nous calculons les élasticités du salaire de réserve et de la durée par rapport à l'indemnité de chômage.

Il existe deux approches dans la littérature sur la durée du chômage. La première approche, ou approche structurelle, est caractérisée par le fait que la théorie du job search est utilisée explicitement. La probabilité conditionnelle de sortie du chômage (la fonction de hasard) est égale au produit du taux d'arrivée d'offres d'emploi et de la probabilité qu'une offre soit acceptable. Une offre de salaire est acceptable si elle dépasse le salaire de réserve (déterminé par la théorie du job search). C'est cette approche qu'utilisent, entre autres, KIEFER et NEUMANN [1979], LANCASTER et CHESHER [1983], NARENDRANATHAN et NICKELL [1985] et RIDDER et GORTER [1986]. La seconde approche, ou approche réduite, consiste simplement à spécifier et à estimer directement la fonction de hasard. Contrairement à l'approche précédente, la théorie du job search n'est pas utilisée. L'avantage évident est que l'on n'impose de cette manière aucune restriction *a priori*. LANCASTER [1979] et LANCASTER et NICKELL [1980] utilisent l'approche réduite. Dans ce papier, les deux approches sont adoptées.

Le modèle structurel est le modèle stationnaire de recherche d'emploi standard. L'idée de décomposer la fonction de hasard en un taux d'arrivée et une probabilité d'acceptation est bien sûr criticable. A l'appui de cette idée, on peut noter que 12 % des chercheurs d'emploi dans notre échantillon indiquent qu'ils ont reçu une ou plusieurs offres pendant le mois qui précédait l'enquête. 5 % de ces offres ont été refusées parce que le salaire était considéré comme insuffisant. Les autres offres ne sont pas acceptées pour des raisons diverses : conditions de travail jugées défavorables, distance travail-maison trop importante, etc.

La seconde approche est précédée par l'application d'un test récemment proposé par Jayet et Moreau [1988] pour tester l'hypothèse d'un hasard proportionnel (HP). Comme l'hypothèse HP n'est pas rejetée nous procédons à l'estimation de ce modèle par la méthode de Cox [1972]. Le modèle HP est ensuite estimé avec l'approche de Prentice et Gloeckler [1978]. Les résultats sont ensuite utilisés pour tester si le hasard est constant.

Le paragraphe 2 contient la description des données. La partie principale est constituée par le paragraphe 3, qui présente les modèles et les estimations. Finalement, le paragraphe 4 apporte quelques éléments de conclusion.

2 Les données

Les données proviennent de « L'enquête sur l'emploi 1986 » de l'INSEE. Cette enquête annuelle est spécialement destinée à étudier le marché du travail. Environ 64 000 ménages (approximativement 1/300 de la population totale) sont enquêtés; il s'agit d'un échantillon représentatif de la population vivant en France. Le questionnaire détaillé permet de classer les individus en chômeurs et non-chômeurs selon les critères du Bureau International du Travail (B.I.T.).

Chaque année l'INSEE étudie de plus près un groupe particulier d'individus à l'aide du questionnaire complémentaire. En 1986, il s'agissait des jeunes entre 15 et 26 ans. Ce sous-échantillon contient 18 435 observations. Comme cet article s'intéresse uniquement à l'entrée sur le marché du travail, une première sélection a consisté à retenir les jeunes qui avaient fini leurs études. Cette sélection a réduit l'échantillon à 5 725 individus qui ont fourni de l'information sur leur niveau d'éducation, leur situation sociale, leur service militaire (s'il a eu lieu) et surtout sur leur expérience sur le marché du travail. En particulier, si au moment de l'enquête (mars 1986), l'individu interrogé avait passé le cap de la période de recherche du premier emploi, ou n'était plus dans son premier emploi, des questions additionnelles permettaient de reconstituer son itinéraire.

En bref, nous travaillons avec deux sortes d'observations :

- Durées censurées : il s'agit d'individus qui sont au chômage, pour la première fois, en mars 1986.

- Durées non-censurées : il s'agit d'individus qui sont sortis de leur première période de chômage.

La table 1 contient des statistiques descriptives sur nos données.

On observe fréquemment dans notre échantillon que les jeunes interrompent leur recherche d'emploi pendant une certaine période (pour effectuer leur service militaire ou pour d'autres raisons). Le calcul de la durée du chômage est basé sur le dernier épisode de recherche ininterrompu. Le

TABLEAU 1

Statistiques descriptives

Épisodes Censurés	Minimum	Maximum	Moyenne	Écart-type
Durée (mois)	1	118	19.43	21.49
Allocation (FFr)	700	2447	1268.90	294.44
Sexe (1 = femme)	0	1	0.59	
$15 \leqslant$ âge < 18	0	1	0.07	
$18 \leqslant$ âge < 20	0	1	0.30	
$20 \leqslant$ âge < 22	0	1	0.33	
$22 \leqslant$ âge $\leqslant 26$	0	1	0.30	
Diplôme:				
1	0	1	0.34	
2	0	1	0.17	
3	0	1	0.33	
4	0	1	0.03	
5	0	1	0.06	
6	0	1	0.07	
$n = 191$				
Épisodes non-censurés				
Durée (mois)	1	70	9.61	10.15
Sexe (1 = femme)	0	1	0.54	
$15 \leqslant$ âge < 18	0	1	0.19	
$18 \leqslant$ âge < 20	0	1	0.42	
$20 \leqslant$ âge < 22	0	1	0.27	
$22 \leqslant$ âge $\leqslant 26$	0	1	0.12	
Diplôme:				
1	0	1	0.15	
2	0	1	0.10	
3	0	1	0.52	
4	0	1	0.04	
5	0	1	0.12	
6	0	1	0.07	
$n = 611$				

pourcentage des chômeurs de longue durée est de 40 %, ce qui correspond assez bien aux chiffres nationaux (voir également annexe 1).

Le système français de prestations sociales est relativement simple pour des individus entre 15 et 25 ans et sans expérience professionnelle (voir Liaisons Sociales (1988)). Ils peuvent bénéficier d'une allocation de 40 francs par jour pendant une période de 12 mois. Les femmes qui sont veuves, divorcées ou qui ont des enfants à charge peuvent recevoir une allocation d'environ 87 francs par jour. Parmi les 191 personnes au chômage dans notre échantillon, seulement 34 d'entre eux touchent des allocations. Pourtant, quelques chômeurs qui pourraient à priori recevoir une allocation déclarent ne pas en toucher. Un problème, également rencontré par LYNCH [1985], est que beaucoup de jeunes vivent toujours avec leurs parents (84 %). Ces jeunes ne prennent pas ceci en considération lorsqu'ils déclarent le montant de leur revenu pendant la période de recherche d'emploi. Nous avons tenté de résoudre ce problème en introduisant de l'hétérogénéité sur l'allocation de chômage (voir paragraphe 3).

Dans notre échantillon, les entrées dans le premier emploi s'étagent entre 1976 et 1986; il nous fallait donc rendre comparables les salaires déclarés lors de l'enquête, en les ramenant en francs constants 1986. La question posée dans le questionnaire complémentaire porte sur le premier salaire perçu, mais il est clair que les réponses données sont sujettes à caution. En effet, une autre question de l'enquête portait sur le salaire en mars 1986. En comparant les deux réponses des individus qui sont toujours dans leur premier emploi en mars 1986, on voit que moins de 19 % des deux salaires déclarés sont différents, alors que 40 % des jeunes interrogés sont depuis au moins 3 ans dans cet emploi. Par ailleurs, pour ceux qui ont changé d'emploi, 11 % indiquent que leur salaire a baissé et 10 % qu'il est resté le même, en francs courants. En déflatant les salaires pour les ramener en francs 1986, on trouve que 29 % des individus auraient connu une baisse de salaire, ce qui est irréaliste.

Nous avons choisi la solution suivante:

– Pour les individus qui sont encore dans leur premier emploi en mars 1986, nous n'utilisons que l'information sur le salaire au moment de l'enquête.

– Pour les autres, il n'y a pas de bonne solution. La moins mauvaise consiste à considérer que la réponse ne porte pas sur le premier salaire, mais sur le salaire à la fin du premier emploi.

Les six catégories de diplômes correspondent au plus haut niveau de diplôme par l'individu. Elles sont définies de la manière suivante: [1]

(1) Abandon en cours de 1^er^ cycle ou études techniques courtes.

(2) Niveau CAP-BEP.

(3) CAP-BEP.

(4) Niveau 3^e^, 2^e^, 1^re^ sans diplôme professionnel.

(5) Bac et niveau Bac (général ou technique).

(6) Diplôme supérieur au Bac.

L'échantillon final ne comporte que 802 observations. En effet, nous avons supprimé celles pour laquelle les variables suivantes n'étaient pas renseignées:

– Le montant d'allocation.

– Le salaire.

– Le diplôme.

Enfin, tous les itinéraires présentant une incohérence (par exemple, durée de recherche d'emploi très différente de la différence entre date de sortie

1. Nous remercions J. Affichard pour son aide dans la définition de ces catégories.

des études et date d'entrée dans le premier emploi) ont été éliminés, ainsi que les individus ne répondant pas aux critères du B.I.T.

Avec ces sélections, nous perdons beaucoup d'observations. Ceci est surtout dû au fait que l'enquête n'était pas destinée à cette étude. La

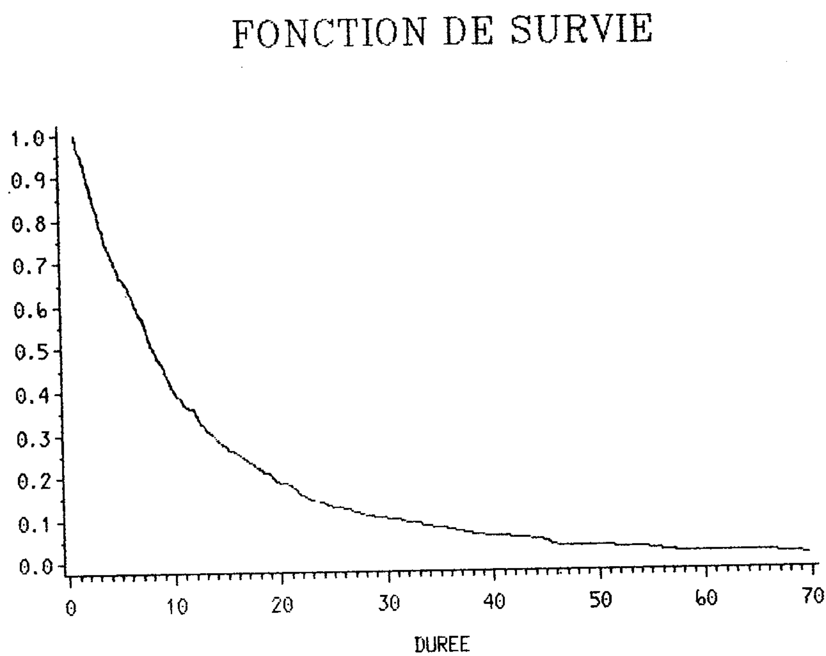

FIGURE 1

durée de recherche d'emploi est calculée en combinant plusieurs questions rétrospectives. Il suffit qu'une réponse soit manquante ou incohérente pour que l'observation soit inutilisable.

L'estimateur de Kaplan-Meier (voir COX et OAKES [1984]) de la fonction de survie est représenté en figure 1 et l'estimateur de la fonction de hasard estimée par la méthode de LIU et van RYZIN [1985] (voir annexe 2) en figure 2.

Les deux estimateurs sont non paramétriques; ils ne prennent pas en compte l'hétérogénéité observée ou non-observée. On voit d'après la figure 2 que le hasard décroît jusqu'à 6 mois, qu'il est constant entre 6 et 12 mois pour ensuite décroître de nouveau: au niveau agrégé le taux de sortie du chômage a tendance à décroître en fonction de la durée. Cela dit, on ne peut rien en conclure sur la forme de la fonction de hasard au niveau individuel: celui-ci peut être constant, alors même que le hasard agrégé présente une dépendance temporelle négative. C'est le phénomène bien connu dit « mover-stayer ».

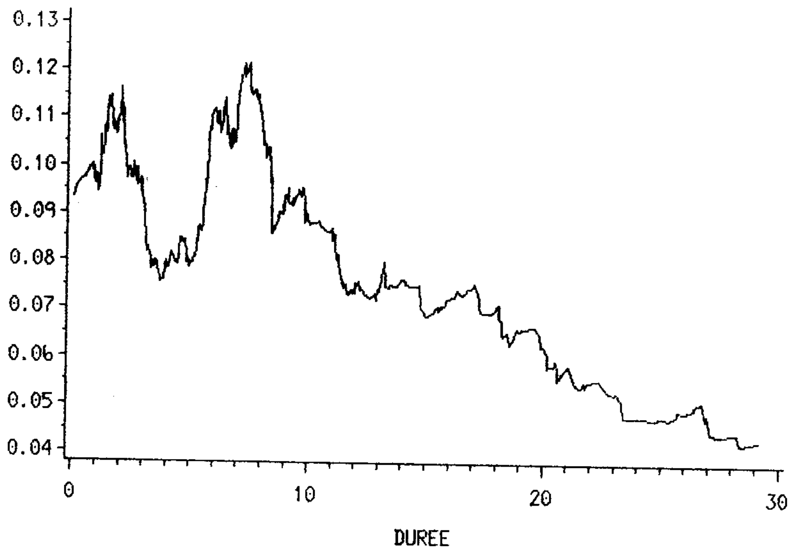

FIGURE 2

3 Spécification et estimation

3.1. Un modèle structurel

Le modèle que nous estimons n'est pas nouveau. En fait, il s'agit du modèle de base de recherche d'emploi (voir par exemple LANCASTER et CHESHER [1983]). Par conséquent, nous rappelerons uniquement les aspects essentiels de ce modèle.

On suppose qu'un chômeur maximise la valeur anticipée de la somme de ses revenus futurs en horizon infini. Les offres d'emploi arrivent en suivant un processus de Poisson d'intensité μ. Ainsi, la durée entre deux offres suit une loi exponentielle de paramètre μ. (La durée moyenne entre deux offres est donc égale à $1/\mu$). A chacune de ces offres est associé un salaire. On suppose que ce salaire est tiré indépendamment de la date d'arrivée des offres dans une loi de fonction de répartition G. Enfin, les revenus futurs sont actualisés au taux ρ.

La stratégie suivie par un individu se caractérise par une fonction $\psi(w)$, $\psi(w)=1$ si l'individu accepte l'emploi avec le salaire w, $\psi(w)=0$ sinon. Dans ce cas, l'individu reçoit des allocations b.

Les hypothèses principales du modèle sont :

- l'environnement est stationnaire,
- la distribution des salaires est connue et non-dégénérée,
- la recherche est séquentielle : le chômeur ne peut choisir entre deux emplois au même instant,
- l'intensité de recherche est exogène.

Sous ces hypothèses, la solution du problème de la recherche optimale est la suivante. En notant I* et ψ^* les solutions du problème de maximisation. on a :

$$(1) \quad I^* = \int_0^\infty \psi^*(w)\frac{w}{\rho}dG(w) + \int_0^\infty (1-\psi^*(w))\,dG(w) \times \left[E_t\left(\int_0^t be^{-\rho s}ds + I^* e^{-\rho t}\right)\right],$$

et :

$$\psi^*(w) = \begin{bmatrix} 1 \text{ si } w \geqslant \xi \\ 0 \text{ sinon} \end{bmatrix}$$

ξ est le salaire de réserve et il est défini par :

$$(2) \qquad \xi = b + \frac{\mu}{\rho}\int_\xi^\infty (1-G(w))\,dw.$$

La durée passée au chômage suit une loi exponentielle : le hasard $\lambda(t)$ est constant :

$$(3) \qquad \lambda(t) = \mu(1-G(\xi)).$$

La fonction de survie est donc :

$$(4) \qquad S(t) = \exp(-\mu(1-G(\xi))\,t).$$

Avec des données sur les salaires de réserve et sur les salaires espérés, Lancaster et Chesher arrivent à déduire (non-paramétriquement) des élasticités et d'autres paramètres d'intérêt. A défaut d'avoir ce type de données, nous sommes obligés d'estimer le modèle.

Les allocations b ne sont pas observées pour les observations non-censurées. Autrement dit, on ne peut pas résoudre dans ce cas l'équation du salaire de réserve (2). Pour les observations censurées, un problème assez similaire se pose. Beaucoup de jeunes vivent toujours avec leurs parents; ils ne prennent pas ceci en considération lorsqu'ils déclarent le montant de leur revenu pendant la période de recherche d'emploi. Le salaire de réserve donné par l'équation (2) sous-estime donc très vraisemblablement le vrai salaire de réserve. Une solution que nous proposons consiste à considérer le salaire de réserve comme une variable aléatoire de support $[\xi_0, \infty)$ et de fonction de répartition $F(\xi)$. Comme b est positif, on trouve la borne inférieure :[2]

$$b=\xi-\frac{\mu}{\rho}\int_{\xi}^{\infty}(1-G(w))\,dw\geqslant 0\Rightarrow\xi\geqslant\xi_0.$$

Il y a deux sortes de contributions à la vraisemblance L.

1. *Observations censurées.*

La probabilité conditionnelle que la durée du chômage soit supérieure à t est :

$$S(t\mid\xi)=\exp(-\mu(1-G(\xi))\,t).$$

Une des questions dans l'enquête porte sur le salaire recherché par le chômeur : $\bar{w}$. La contribution à la vraisemblance est donc la probabilité non-conditionnelle :

$$\text{(5)}\qquad L_A=\int_{\xi_0^1}^{\bar{w}}\exp(-\mu(1-G(\xi))\,t)\,dF(\xi),$$

où ξ_0^1 est la solution implicite de :

$$\xi_0^1=b+\frac{\mu}{\rho}\int_{\xi_0^1}^{\infty}(1-G(w))\,dw.$$

On voit que ξ_0 et ξ_0^1 coïncident si $b=0$.

2. *Observations non-censurées.*

L'unité d'observation pour les durées est le mois. La probabilité conditionnelle d'observer un salaire w et une durée dans l'intervalle $[t-1,t)$ est :

$$(S(t-1\mid\xi)-S(t\mid\xi))\frac{dG(w)}{1-G(\xi)}.$$

Si on prend en compte le fait que les salaires sont donnés par intervalles $[w_1,w_2]$, on trouve la contribution à la vraisemblance :

$$\text{(6)}\qquad L_B=\int_{w_1}^{w_2}\int_{\xi_0}^{w}(S(t-1\mid\xi)-S(t\mid\xi))\frac{dF(\xi)}{1-G(\xi)}\,dG(w).$$

La fonction de vraisemblance s'écrit :

$$\text{(7)}\qquad L=\prod_{i\in A}L_A\prod_{i\in B}L_B.$$

2. Cette remarque est due à Stéphane Gregoir.

3.2. Résultats des estimations

Pour identifier les paramètres structurels du modèle, nous devons spécifier les distributions du salaire et du salaire de réserve. Nous supposons que les salaires suivent une loi log-normale de moyenne m et de variance σ^2. Les salaires de réserve suivent une loi exponentielle translatée:

$$F(\xi) = 1 - \exp(-(\xi - \xi_0)/a), \tag{8}$$

de moyenne $(a + \xi_0)$ et de variance a^2.

En principe nous aurions pu estimer les paramètres μ, ρ, m, σ^2, a et ξ_0 comme des fonctions de variables explicatives. Mais il aurait fallu alors spécifier une certaine forme de dépendance entre les paramètres et les variables explicatives. Nous avons préféré procéder de la manière suivante: l'échantillon a été partagé en 12 classes, selon le sexe et les catégories de diplôme. Nous supposons que les paramètres dépendent uniquement des classes définies par les deux variables sexe et diplôme: dans chaque classe, les individus font face à la même distribution des salaires, le même taux d'arrivée d'offres et ils actualisent au même taux ρ. Le salaire de réserve est néanmoins hétérogène à cause de l'hétérogénéité sur b.

TABLEAU 2

La distribution des salaires

Classe	m	σ^2
Femme et éducation 1 (55).	3.338 (0.177)**	0.874 (0.120)**
Femme et éducation 2 (32).	2.985 (0.390)**	1.835 (0.313)**
Femme et éducation 3 (69).	3.823 (0.286)**	1.621 (0.130)**
Femme et éducation 4 (170).	4.003 (0.201)**	1.697 (0.089)**
Femme et éducation 5 (79).	4.588 (0.172)**	1.536 (0.087)**
Femme et éducation 6 (34).	5.167 (0.789)**	7.422 (0.706)**
Homme et éducation 1 (71).	3.765 (0.272)**	2.340 (0.198)**
Homme et éducation 2 (38).	3.230 (1.194)**	2.930 (1.618)**
Homme et éducation 3 (93).	3.488 (0.434)**	3.333 (0.503)**
Homme et éducation 4 (105).	3.764 (0.252)**	2.261 (0.278)
Homme et éducation 5 (33).	4.893 (1.493)**	5.801 (1.560)**
Homme et éducation 6 (23).	2.996 (7.833)	3.965 (0.574)

(Écart-type): * = significative à 5 %; ** = significative à 1 %.

La table 2 contient les estimations de m et σ^2. Les résultats sont en milliers de francs 1986. Nous donnons également le nombre d'invidus dans chaque classe.

Comme prévu, les salaires offerts sont plus bas pour les diplômes les moins élevés. A un niveau de diplôme donné, il n'existe pas de grandes différences entre la moyenne des salaires offerts aux femmes et hommes; pourtant, la dispersion (mesurée par σ^2) est plus grande chez les hommes.

Nous avons eu quelques problèmes pour estimer les autres paramètres du modèle pour les sept premières classes. En examinant la vraisemblance pour ces classes, on voit qu'elle est presque plate pour de grandes valeurs de ρ. De fait, le maximum de la vraisemblance est obtenu pour $\rho = \infty$.

Pour contrôler s'il ne s'agissait pas d'un problème purement numérique nous avons reparamétrisé la vraisemblance:[3] si ρ est très grand, ξ_0^1 est proche de b; au lieu de résoudre l'équation (2) à chaque itération, on considère que ξ_0^1 est égale à $b+\int_b^\infty (1-G(w))\,dw$. (On obtient facilement cette expression de l'équation 2 en prenant $\xi \cong b$). Quand $b=0$, on obtient $\xi_0=\frac{\mu}{\rho}m$. Cette reparamétrisation n'a pas modifié les résultats.

Finalement nous avons décidé d'estimer le modèle sur ces classes en fixant $\xi_0=0$ et $\xi_0^1=b$. Cela signifie que l'individu accepte la première offre qui est supérieure à son revenu. L'annexe 4 présente les estimations de μ, ρ, a et les deux premiers moments de la distribution du salaire de réserve.

Regardons maintenant les paramètres d'intérêt de notre modèle: la durée moyenne entre deux offres (μ^{-1}), la probabilité d'accepter un emploi $(1-G(\xi))$ et la durée moyenne du chômage $(\mu^{-1}(1-G(\xi))^{-1})$. En ce qui concerne les deux dernières expressions, on a calculé $\mathbb{E}(1-G(\xi))$ et $\mathbb{E}(\mu^{-1}(1-G(\xi))^{-1})$, où l'espérance est calculée par rapport à la distribution

Tableau 3

Implications des estimations (b = 0)

Classe	Probabilité d'acceptation	Durée moyenne entre offres (mois)	Durée moyenne du chômage (mois)	
			modèle	non para-métrique
Femme et éducation 1 (55)	98	43.4 (7.2)**	66.6	30.3
Femme et éducation 2 (32)	91	20.1 (5.5)**	30.2	17.0
Femme et éducation 3 (69)	95	15.8 (1.5)**	22.3	14.7
Femme et éducation 4 (170)	95	15.0 (1.4)**	22.5	15.0
Femme et éducation 5 (79)	97	9.7 (1.1)**	13.4	9.3
Femme et éducation 6 (34)	88	5.6 (2.1)**	7.5	4.8
Homme et éducation 1 (71)	98	24.8 (3.5)**	26.0	23.5
Homme et éducation 2 (38)	84	11.4 (10.9)	203	15.5
Homme et éducation 3 (93)	78	4.5 (1.6)**	10.2	8.0
Homme et éducation 4 (105)	89	9.9 (1.8)**	17.1	11.1
Homme et éducation 5 (33)	98	8.3 (1.5)**	8.8	9.0
Homme et éducation 6 (23)	15	1.0 (6.1)	9.3	5.9

(Écarts-type); * = significative à 5 %; ** = significative à 1 %.

de ξ. La dernière colonne de la table 3 contient la moyenne empirique (calculée avec l'estimateur de Kaplan-Meier) de la durée du chômage.

Les résultats doivent être interprétés avec prudence. Certaines classes contiennent trop peu d'observations; les paramètres associés sont mal estimés. Voici quand même quelques conclusions: les probabilités d'acceptation

3. Ceci nous a été suggéré par Christian Gouriéroux.

sont élevées (autour de 90 %) et plus élevées pour les femmes que pour les hommes. La classe des hommes ayant le diplôme le plus élevé constitue un cas à part : ils reçoivent nettement plus d'offres et paraissent plus exigeants que les autres chômeurs. Les variations dans les durées moyennes de chômage sont dues aux variations de la durée entre les offres. Les durées qui séparent les offres sont inférieures pour les hommes et pour les plus diplômés. Plus précisément, les jeunes hommes avec un diplôme professionnel (Education 3) reçoivent plus d'offres que les jeunes hommes qui ont suivi une éducation plus générale mais sans diplôme supérieur au bac. En ce qui concerne l'adéquation du modèle aux données, il semble qu'elle soit meilleure pour les durées courtes et pour les hommes. Ceci peut être une indication de non-stationnarité.

Nos résultats rejoignent ceux des études antérieures : FLINN et HECKMAN [1982] ont estimé un modèle exponentiel et ils ont trouvé une probabilité d'acceptation de 60 %. D'après leurs résultats, la durée moyenne entre les offres est égale à 5 mois tandis que la durée moyenne du chômage est estimé à 8.3 mois. Les estimations sont basés sur un échantillon de jeunes (20-24 ans) bacheliers masculins. FOUGÈRE [1989] a estimé un modèle stationnaire de recherche d'emploi (en introduisant plusieurs types d'offres) sur un échantillon de jeunes femmes vivant en France. La probabilité d'acceptation estimée est proche de 1, la durée moyenne entre les offres varie entre 8.8 et 10 mois tandis que la durée moyenne du chômage varie entre 9.2 et 14.5 mois.

Enfin, les élasticités du salaire de réserve et de la durée sont données dans l'annexe 4. Ces élasticités sont en général plus élevées pour les moins diplômés : un changement d'allocation modifie plus le salaire de réserve des individus non-qualifiés que des individus qualifiés. Les élasticités de la durée varient entre -0.07 et -1.34. Cette dernière estimation implique qu'une réduction de 10 % de l'allocation entraîne une baisse de la durée du chômage de 13 %.

3.3. Modèles réduits

Bien que le modèle à hasard proportionnel soit souvent utilisé, les hypothèses sous-jacentes sont rarement testées. Nous appliquons un test proposé récemment par JAYET et MOREAU [1988]. La méthode n'est utilisable que pour des variables exogènes qualitatives. Pour chaque variable, l'hypothèse H.P. est testée (voir annexe 3). L'hypothèse H.P. est acceptée pour les deux variables que nous utilisons, le sexe et le diplôme. Les valeurs du chi-deux (degrés de liberté) sont 42.80 (31) et 58.34 (45).

Nous allons maintenant essayer de tester l'hypothèse de stationnarité du modèle. Soit le modèle à H.P. suivant :

$$\lambda(t)=\lambda_0(t)\,e^{x'\beta}. \tag{9}$$

Grâce à la méthode proposée par Cox [1972, 1975], le paramètre β peut être estimé sans qu'il soit nécessaire de spécifier le hasard de base $\lambda_0(t)$. Supposons qu'il y a n observations (censurées et non-censurées), notées $t_1, \ldots, t_n$ avec $t_1 < t_2 < \ldots < t_m$ (où $m < n$ si les observations sont groupées).

L'estimateur de β est défini comme le maximum de la vraisemblance partielle :

$$(10) \qquad L(\beta)=\prod_{i=1}^{n}\left[\frac{e^{x_i'\beta}}{\sum_{k\in R_i} e^{x_k'\beta}}\right]^{d_i},$$

où $R_i (i=1, \ldots, m)$ représente l'ensemble des individus à risque à l'instant t_i (*i.e.* $R_i=\{k \mid 1 \leqslant k \leqslant n,\ t_k \geqslant t_i\}$) et où $d_i=0$ si le i-ème épisode est censuré, $d_i=1$ sinon. Les résultats de l'estimation sont regroupés dans la table 4.

TABLEAU 4

Estimation semi-paramétrique de (9)

Sexe	−0.35 (0.08)**
Education :	
2	0.45 (0.19)**
3	0.93 (0.14)**
4	0.70 (0.14)**
5	1.09 (0.16)**
6	1.20 (0.19)**
$n=802$	

(Écart-type); * = significative à 5 %; ** = significative à 1 %.

Le signe des coefficients estimés est bien celui auquel on s'attendait. La valeur estimée de β, $\hat{\beta}$, peut servir pour construire un estimateur non-paramétrique du hasard de base λ_0. Nous utilisons un estimateur de type histogramme pour λ_0 basé sur le hasard intégré estimé :

$$(11) \qquad \hat{\lambda}_0(t)=\frac{\hat{\Lambda}_0(t_i)-\hat{\Lambda}_0(t_i-h)}{h},$$

où le hasard intégré est estimé par (BRESLOW [1974]) :

$$(12) \qquad \hat{\Lambda}_0(t)=\sum_{t_i \leqslant t}\frac{d_i}{\sum_{k\in R_i}\exp(x_k'\hat{\beta})}.$$

On trouvera le graphique de (12) sur la figure 3, accompagné de la droite $y=0.045\,t$. La figure indique que le hasard est constant, au moins pour les 25 premiers mois.

Maintenant nous allons estimer (9) avec l'approche de Prentice et Gloeckler. La version en temps discret du hasard s'écrit :

$$(13) \qquad \lambda(t)=1-\exp(-\exp(x'\beta+y(t)));\ t=1,2,\ldots$$

où $y(t)=ln\left(\int_t^{t+1}\lambda_0(u)\,du\right)$.

Les estimations des coefficients des variables explicatives figurent dans la table 5.

HASARD DE BASE INTEGRE

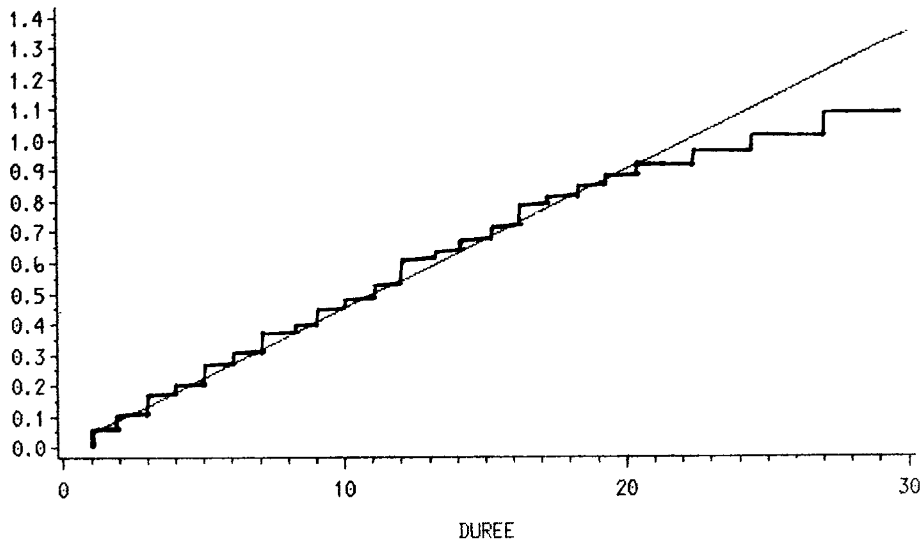

FIGURE 3

TABLEAU 5

Modèle (13)

Sexe.	−0.35 (0.08)**
Education :	
2.	0.55 (0.20)**
3.	0.95 (0.16)**
4.	0.78 (0.15)**
5.	1.16 (0.17)**
6.	1.32 (0.20)**
$n = 802$	

(Écart-type); * = significative à 5 %; ** = significative à 1 %.

Pour estimer (13), nous avons choisi les intervalles $[1, 2), \ldots, [17, 18)$, $[18, 20)$, $[20, 22)$, $[22, 24)$ et $[24, \infty)$. La figure 4 reprend les estimations de :

$$\exp\left(-\int_{t_i}^{t_{i+1}} \lambda_0(u)\,du\right)/(t_{i+1} - t_i).$$

Si le hasard de base était constant sur les intervalles $[t_i, t_{i+1})$, la figure (4) montrerait la forme de $\lambda_0(u)$. L'hypothèse de stationnarité se traduit par $H_0 : y(1) = y(2) = \ldots$

H_0 est accepté jusqu'à 20 mois.

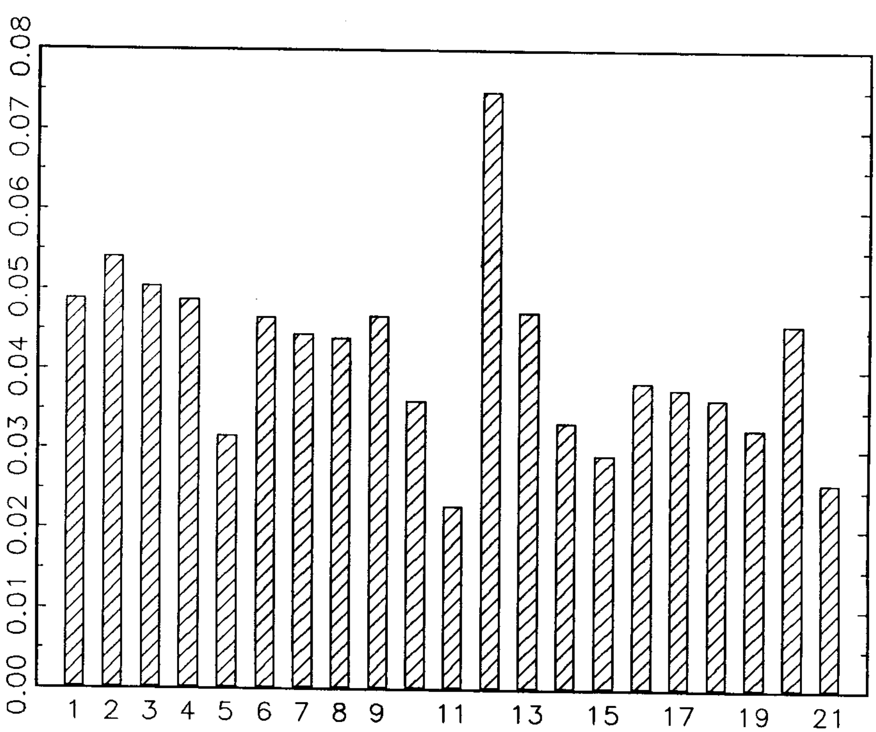

FIGURE 4

4 Conclusion

La décomposition de la fonction de hasard en un taux d'arrivée d'offres et une probabilité d'accepter un emploi est un des principaux attraits des modèles structurels. Notre but, dans cette étude, était de tenter d'expliquer les différences observées entre les durées moyennes de chômage des hommes et des femmes, des plus ou moins diplômés : certains chômeurs sont-ils plus exigeants que d'autres, ou bien des durées de chômage plus longues sont-elles dues à un rythme d'arrivée des offres d'emploi plus faible ? Bien entendu, les résultats obtenus doivent être interprétés avec prudence : le modèle adopté est très simple, les données ne sont pas les mieux adaptées. Mais, avec ces réserves, deux conslusions claires se dégagent :

– Les probabilités d'acceptation d'une offre d'emploi sont très semblables d'un groupe à l'autre. Les différences entre durées moyennes de chômage observées sont dûes essentiellement à un rythme d'arrivée des offres plus lent chez les femmes et chez les moins diplômés.

– Un diplôme professionnel est un atout pour un jeune qui se présente sur le marché du travail. Le tableau 3 montre que la durée moyenne entre deux offres d'emploi est plus faible pour les jeunes qui ont un CAP ou un

BEP que pour ceux qui ont atteint un niveau d'enseignement général supérieur, sans diplôme professionnel.

En général, nos résultats sont semblables à ceux d'autres études. Les élasticités de la durée du chômage par rapport aux allocations sont plus petites que les élasticités (non-paramétriques) trouvées par LANCASTER et CHESHER [1983]. Une explication pourrait être que nos estimations sont basées sur un échantillon de jeunes dont beaucoup vivent encore chez leurs parents. Autrement dit, ils ont d'autres sources de revenu que les allocations.

En ce qui concerne l'hypothèse de stationnarité, le modèle structurel et le modèle réduit semblent indiquer la même conclusion : l'hypothèse de stationnarité est valide pour les deux premières années de recherche d'emploi.

Durée (mois)	Fréquences de la durée — Fréquence	Pourcentage	Pourcentage cumulé
1	72	9.0	9.0
2	73	9.1	18.1
3	68	8.5	26.6
4	53	6.6	33.2
5	33	4.1	37.3
6	45	5.6	42.9
7	60	7.5	50.4
8	40	5.0	55.4
9	53	6.6	62.0
10	31	3.9	65.8
11	12	1.5	67.3
12	30	3.7	71.1
13	19	2.4	73.4
14	14	1.8	75.2
15	11	1.4	76.6
16	14	1.7	78.3
17	11	1.4	79.7
18	10	1.2	80.9
19	18	2.2	83.2
20	4	0.5	83.7
21	18	2.2	85.9
22	10	1.2	87.2
23	3	0.4	87.5
24	7	0.9	88.4

Pour estimer le hasard $\lambda(t)$, nous utilisons la méthode proposée par Liu et van Ryzin (1985).

Soit (U_j) $j=1, 2, \ldots, d$ la statistique d'ordre des observations non-censurées. Définissons :

$$\begin{aligned} T_n(x) &= \inf\left(j; x \leqslant \frac{u_j+u_{j+k}}{2}\right) \quad \text{si} \quad u_1 \leqslant x \leqslant u_d \\ &= 0 \quad \text{si} \quad x < u_1 \quad \text{ou} \quad d+1-k<0 \\ &= d+1-k \quad \text{si} \quad x > u_d. \end{aligned}$$

Soit $H_n(x)$ la fonction de répartition des observations et

$$F_n(x) = n^{-1} \sum_{i=1}^{d} \mathbf{1}\{U_i \leqslant x\}$$

En écrivant T pour $T_n(x)$, on définit :

$$\hat{h}_n(x) = [F_n(U_{T+k}) - F_n(U_T)]\,[U_{T+k} - U_T]^{-1}\,[1 - H_n(x)]^{-1}.$$

Pour des valeurs bien choisies de k, $\hat{h}_n(x)$ est un estimateur convergent de $h_n(x)$. Ici nous prenons $k =$ partie entière de $n^{0.7}$

Un désavantage de cette approche est qu'on ne peut pas avoir un estimateur convergent de λ sur tout le domaine des observations. La figure 2 montre l'estimateur uniquement pour les 21 premiers mois de chômage.

Pour tester l'hypothèse H.P., nous utilisons le test proposé par JAYET et MOREAU [1988]. Nous décrivons ici les idées principales.

Supposons qu'il y ait deux groupes dans la population. Les hasards intégrés sont respectivement $\Lambda_1(t)$ et $\Lambda_0(t)$. Si l'hypothèse H.P. est vérifiée, $\Lambda_1(t)$ et $\Lambda_0(t)$ satisfont :

$$\Lambda_1(t) = e^b \Lambda_0(t) \Rightarrow \ln(\Lambda_1(t)) - \ln(\Lambda_0(t)) = b \tag{14}$$

Un estimateur non-paramétrique de $\Lambda_i(t)$ facilement calculable est l'estimateur de Nelson-Aalen (voir ANDERSEN et BORGAN [1985]). Une façon simple de vérifier l'hypothèse H.P. est de tester si les estimations non paramétriques $\hat{\Lambda}_i(t)$ satisfont l'équation (14). La distribution asymptotique de $n^{1/2}(\hat{\Lambda}_i(t) - \Lambda_i(t))$ est connue. Supposons que nous estimons $\Lambda_i(t)$ pour m points $t_1, \ldots, t_m$. Alors, pour $i = 1, \ldots, m$:

$$n^{1/2}[\ln(\hat{\Lambda}_1(t_i)) - \ln(\hat{\Lambda}_0(t_i)) - (\ln(\Lambda_1(t_i) - \ln(\Lambda_0(t_i)))] \to N(0, \Omega),$$

où Ω peut être calculé facilement à l'aide de l'estimateur de la matrice de variance-covariance asymptotique de $\hat{\Lambda}_i(t)$. Soit Λ le vecteur constitué par les éléments $\ln(\hat{\Lambda}_1(t_i)) - \ln(\hat{\Lambda}_0(t_i))$. Le minimum de

$$\xi = (\Lambda - b)' \Omega^{-1} (\Lambda - b)$$

donne un estimateur convergent de b et, sous l'hypothèse nulle, $n\xi$ évalué à l'optimum suit asymptotiquement un chi-deux avec $m-1$ degrés de liberté (voir GOURIÉROUX et MONFORT [1989]). Il est très simple d'appliquer l'idée précédente à plus de deux groupes.

ANNEXE 4

Tableau A1

Classe	μ	ρ	a	ξ_0
Femme et éducation 1 (55)......	0.023**	∞	0.779*	0.000
	(0.004)		(0.439)	
Femme et éducation 2 (32)......	0.050**	∞	0.990	0.000
	(0.014)		(0.773)	
Femme et éducation 3 (69)......	0.063**	∞	1.064*	0.000
	(0.006)		(0.534)	
Femme et éducation 4 (170).....	0.067**	∞	1.161**	0.000
	(0.006)		(0.345)	
Femme et éducation 5 (79)......	0.103**	∞	1.197*	0.000
	(0.012)		(0.484)	
Femme et éducation 6 (34)......	0.179	∞	1.878	0.000
	(0.068)		(1.322)	
Homme et éducation 1 (71).....	0.040**	∞	0.735*	0.000
	(0.006)		(0.360)	
Homme et éducation 2 (38).....	0.008	1.054	1.238	0.247
	(0.084)	(37.970)	(4.624))	
Homme et éducation 3 (93).....	0.222**	1.489	1.534	0.451
	(0.078)	(5.956)	(1.155)	
Homme et éducation 4 (105).....	0.101**	0.744	1.227	0.449
	(0.018)	(3.726)	(0.991)	
Homme et éducation 5 (33).....	0.112**	0.412	0.053	1.027
	(0.190)	(2.496)	(0.077)	
Homme et éducation 6 (23).....	1.014	0.124	1.228	3.725
	(6.237)	(0.073)	(1.135)	

(Écart-type); * = significative à 5 %; ** = significative à 1 %.

Tableau A2

La distribution du salaire de réserve

Classe	Moyenne	Variance
Femme et éducation 1 (55)....	0.779	0.607
Femme et éducation 2 (32)....	0.990	0.980
Femme et éducation 3 (69)....	1.064	1.133
Femme et éducation 4 (170)....	1.161	1.348
Femme et éducation 5 (79)....	1.197	1.433
Femme et éducation 6 (34)....	1.878	3.527
Homme et éducation 1 (71)...	0.735	0.540
Homme et éducation 2 (38)...	1.485	1.533
Homme et éducation 3 (93)...	1.985	2.352
Homme et éducation 4 (105)...	1.676	1.506
Homme et éducation 5 (33)...	1.080	0.003
Homme et éducation 6 (23)...	4.953	1.509

Les élasticités du salaire de réserve et du hasard par rapport à l'allocation sont calculées de la façon suivante. Nous savons que :

$$\frac{\partial \ln \xi}{\partial \ln b} = \frac{b}{\xi}\left(1 + \frac{\mu}{\rho}[1 - G(\xi)]\right)^{-1}$$

$$\frac{\partial \ln \lambda(t)}{\partial \ln b} = - \frac{g(\xi)}{1 - G(\xi)} b \left[1 + \frac{\mu}{\rho}(1 - G(\xi))\right]^{-1}$$

Les élasticités calculées sont :

$$e_{\xi/b} = \int_{\xi_0^1}^{\infty} \frac{\partial \ln \xi}{\partial \ln b} \frac{dF(\xi)}{1 - F(\xi_0^1)}$$

$$e_{\lambda/b} = \int_{\xi_0^1}^{\infty} \frac{\partial \ln \lambda(t)}{\partial \ln b} \frac{dF(\xi)}{1 - F(\xi_0^1)}$$

Où ξ_0^1 est la solution de $\xi_0^1 = b + \frac{\mu}{\rho}\int_{\xi_0^1}^{\infty}(1 - G(w))\,dw$.

TABLEAU A3

Élasticités

Classe	b, ξ_0^1	$e_{\xi/b}$	$e_{\lambda/b}$
Femme et éducation 1 (55).....	1000, 1000	0.64	−0.19
	2000,2000	0.76	−1.19
Femme et éducation 2 (32).....	1000, 1000	0.60	−0.39
	2000, 2000	0.72	−1.34
Femme et éducation 3 (69).....	1000, 1000	0.58	−0.19
	2000, 2000	0.71	−0.89
Femme et éducation 4 (170)....	1000, 1000	0.57	−0.18
	2000, 2000	0.70	−0.82
Femme et éducation 5 (79).....	1000, 1000	0.56	−0.12
	2000, 2000	0.69	−0.54
Femme et éducation 6 (34)....	1000, 1000	0.47	−0.17
	2000, 2000	0.61	−0.52
Homme et éducation 1 (71)....	1000, 1000	0.65	−0.14
	2000, 2000	0.77	−0.78
Homme et éducation 2 (38)....	1000, 1171	0.47	−0.39
	2000, 2104	0.63	−1.09
Homme et éducation 3 (93)....	1000, 1322	0.39	−0.38
	2000, 2208	0.56	−0.99
Homme et éducation 4 (105)...	1000, 1331	0.41	−0.29
	2000, 2215	0.59	−0.97
Homme et éducation 5 (33)...	1000, 1815	0.42	−0.07
	2000, 2614	0.61	−0.34
Homme et éducation 6 (23)...	1000, 4072	0.10	−0.27
	2000, 4466	0.20	−0.59

• Références bibliographiques

Breslow, J. (1974). – « Covariance Analysis of Censored Survival Data », *Biometrics*, 30, pp. 89-99.

Cox, D. R. (1972). – « Regression Models and Life Tables », *Journal of the Royal Statistical Society*, B, 34, pp. 187-220.

Cox, D. R. (1975). – « Partial Likelihood », *Biometrika*, 62, pp. 269-276.

Cox, D. R. et Oakes, D. (1984). – *Analysis of Survival Data*, Chapman and Hall.

Flinn, C. et Heckman, J. (1982). – « New Methods for Analyzing Structural Models of Labor Force Dynamics », *Journal of Econometrics*, 18, pp. 115-168.

Fougère, D. (1989). – « Recherche d'emploi en présence de contrats de courte durée : modélisation et estimation sur données individuelles », *Annales d'Économie et Statistique*, 14, p. 225-257.

Gouriéroux, C. et Monfort, A. (1989). – « A General framework for Testing a Null Hypothesis in a Mixed Form », *Econometric Theory*, 5, pp. 63-82.

INSEE (1986). – L'activité et le chômage en mars 1986. Premiers résulats, n° 72.

INSEE (1987). – Population active, emploi et chômage depuis 30 ans. Les collections de l'INSEE, D123.

Jayet, H. et Moreau, A. (1988). – « Analysis of Survival Data: Estimation and Test using Asymptotic Least Squares », à paraître dans *Journal of Econometrics*.

Kiefer, N. M. et Neumann, G. R. (1979). – « An Empirical Job Search Model, with a Test of the Constant-Reservation Wage Hypothesis », *Journal of Political Economy*, 87, pp. 89-108.

Lancaster, T. (1979). – « Econometric Methods for the Duration of Unemployment », *Econometrica*, 47, pp. 939-956.

Lancaster, T. et Chesher, A. (1983). – « An Econometric Analysis of Reservation Wages », *Econometrica*, 51, pp. 1661-1676.

Liu, R. Y. C. et van Ryzin, J. (1985). – « A Histogram Estimator of the Hazard Rate with Censored Data », *The Annals of Statistics*, 13, pp. 592-605.

Lynch, L. (1985). – « State Dependency in Youth Unemployment: a Lost Generation? », *Journal of Econometrics*, 28, pp. 29-50.

Narendranathan, W. et Nickell, S. (1985). – « Modelling the Process of Job Search », *Journal of Econometrics*, 28, pp. 29-50.

Nickell, S. (1979). – « Estimating the Probability of Leaving Unemployment, *Econometrica*, 47, pp. 57-67.

Prentice, R. et Gloeckler, L. (1978). – « Regression Analysis of Grouped Survival Data with an Application to Breast Cancer Data », *Biometrics*, 34, pp. 57-67.

Ridder, G. et Gorter, K. (1986). – « Unemployment Benefits and Search Behavior: an Empirical Investigation », Working paper, Cornell University.

Thélot, C. (1988). – « La sortie du chômage », Mélanges économiques, *Essais en l'honneur de Edmond Malinvaud*, Economica.