



Les déterminants individuels de la durée de chômage : conséquence de l'oubli du biais d'échantillonnage

Thierry Teste

► To cite this version:

Thierry Teste. Les déterminants individuels de la durée de chômage : conséquence de l'oubli du biais d'échantillonnage. [Rapport de recherche] Laboratoire d'analyse et de techniques économiques(LATEC). 1997, 20 p., Table, ref. bib.: 1 p. 3/4. hal-01526842

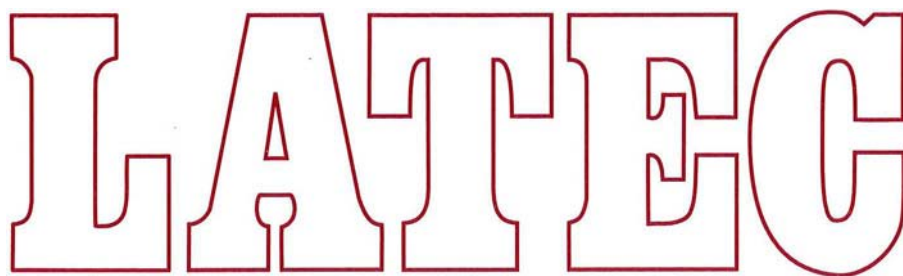
HAL Id: hal-01526842

<https://hal.science/hal-01526842>

Submitted on 23 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**LABORATOIRE D'ANALYSE
ET DE TECHNIQUES ÉCONOMIQUES**

UMR 5601 CNRS

DOCUMENT DE TRAVAIL



Pôle d'Economie et de Gestion

2, boulevard Gabriel - 21000 DIJON - Tél. 03 80 39 54 30 - Fax 03 80 39 54 43

ISSN : 1260-8556

n° 9717

**Les déterminants individuels de la durée du chômage :
Conséquence de l'oubli du biais d'échantillonnage**

Thierry TESTE*

Octobre 1997

Enseignant contractuel à l'IUT de Dijon

LES DETERMINANTS INDIVIDUELS DE LA DUREE DU CHOMAGE : CONSEQUENCE DE L'OUBLI DU BIAIS D'ECHANTILLONNAGE.

Thierry TESTE¹

Résumé:

Si, aujourd'hui, les études sur les comportements des demandeurs d'emploi sur le marché du travail sont de plus en plus importantes, la modélisation de ces comportements comporte encore un bon nombre d'interrogations. Parmi celles-ci figure le problème du biais d'échantillonnage présent dans la plupart des données françaises sur les durées de chômage. Dans un précédent travail, nous avons montré l'importance et les conséquences de l'oubli du biais d'échantillonnage sur les probabilités conditionnelles de sortie du chômage. Dans le travail présenté ici et réalisé à partir de l'enquête « suivi des chômeurs », notre intérêt porte sur les conséquences de l'oubli de ce biais sur les déterminants individuels de la durée du chômage. Pour cela, nous comparons les effets de certaines variables sur les taux de sortie du chômage en reprenant des techniques éprouvées et maintenant bien connues de la théorie des modèles de durée. A partir de deux modèles bien distincts, le modèle de Cox et un modèle de Weibull, nous montrons que le biais d'échantillonnage influence peu l'effet des variables individuelles sur la probabilité conditionnelle de sortie du chômage. C'est à dire qu'un effet positif dans un modèle est le plus souvent positif dans l'autre modèle et réciproquement pour un effet négatif. En revanche, l'oubli du biais d'échantillonnage revient toujours à surestimer l'effet de ces variables sur les taux de sortie du chômage.

Mots clés:

Taux de hasard, modèle de Cox, modèle de Weibull, biais de sélection, vraisemblance conditionnelle

¹L'auteur remercie P. Werquin pour les remarques et suggestions qui ont suscité la révision d'une première version de ce travail.

1. INTRODUCTION.

Si, aujourd'hui, les études sur les comportements des demandeurs d'emploi sur le marché du travail sont de plus en plus importantes, la modélisation de ces comportements comporte encore un bon nombre d'interrogations. Parmi celles-ci figure le problème du biais d'échantillonnage présent dans la plupart des données françaises sur les durées de chômage. En effet, l'échantillonnage dans les enquêtes sur les chômeurs constitue, le plus souvent, une difficulté dans l'estimation des modèles, parce qu'il existe une liaison entre le mode de sélection et la variable de durée observée. Pour que les méthodes traditionnelles de modélisation soient applicables, il faut disposer d'un échantillon composé de périodes de chômage, et non de chômeurs, comme l'ont fait BONNAL et FOUGERE (1990) avec le flux des radiés² d'une agence locale de l'ANPE.

Plusieurs solutions ont été proposées par les économètres pour corriger ce biais. L'une d'elle, présentée par HECKMAN et BORJAS (1980), préconise l'introduction dans le modèle *« d'un effet individuel rendant compte de la récurrence du chômage chez un même individu, et des conséquences qu'elle peut induire sur les probabilités de sortie »*. En effet, seule la première période de chômage est soumise à la sélection endogène. CASES et LOLLIVIER (1993) donnent une solution plus simple en introduisant une variable indicatrice de rang pour la période de chômage observée, qui ne prend en compte que le numéro d'ordre de la période et non la durée des périodes précédentes. *« Toutefois, sa mise en oeuvre se heurte à un problème d'identification, dans la mesure où les durées brèves proviennent presque exclusivement des périodes de chômage répétitives »*. LANCASTER (1990) a également proposé sa solution pour corriger le biais résultant de la sélection endogène. Il ne retient, pour l'étude, que les périodes autres que la première, et qui ne sont donc pas soumises à cette sélection. Mais, dans le cas présent, cela revient à laisser de côté près des 2/3 des périodes observées, ce qui n'est guère envisageable. Enfin, une dernière méthode, proposée par GOURIEROUX et MONFORT (1991), consiste à maximiser une vraisemblance conditionnelle à la date d'entrée au chômage, ce qui entraîne la correction automatique de *« la loi de la variable en fonction de l'ancienneté dans le chômage à la date d'entrée dans l'échantillon. Les estimateurs ainsi obtenus possèdent les propriétés classiques des estimateurs du maximum de vraisemblance, mais sont corrigés des biais liés à la sélection endogène de*

²C'est à dire de chômeurs sortant sortants de l'Anpe.

l'échantillon ». C'est cette dernière solution que nous avons retenue et que nous allons étudier.

Dans un précédent travail et avec la même méthode, nous avons montré l'importance et les conséquences de l'oubli du biais d'échantillonnage sur les probabilités conditionnelles de sortie du chômage [TESTE, (1995)]. Rappelons simplement que ne pas considérer le biais de sélection revient à sous estimer de deux à trois fois les taux de sortie du chômage. Dans le travail présenté ici et réalisé à partir de l'enquête « suivi des chômeurs », notre intérêt porte sur les conséquences de l'oubli de ce biais sur les déterminants individuels de la durée du chômage. Pour cela, nous comparons les effets de certaines variables sur les taux de sortie du chômage en reprenant des techniques éprouvées et maintenant bien connues de la théorie des modèles de durée³. Il est également permis de distinguer plusieurs destinations possibles à la suite d'une période de chômage. Les derniers travaux empiriques réalisés sur données françaises introduisaient généralement quatre états pour modéliser le marché du travail: le chômage, l'emploi stable, l'emploi précaire et l'inactivité. Nous avons reconduit ici ces hypothèses de la façon qui sera précisée plus loin.

Ainsi, deux types de modèles ont été retenus dans l'étude pour formaliser les probabilités conditionnelles de sortie du chômage. Le premier est le modèle semi-paramétrique de Cox dont le principal intérêt est d'éviter de choisir *a priori* une formalisation statistique de la loi de la durée. Deux raisons ont motivé le choix de ce modèle: d'une part, la durée de chômage n'intervient pas dans l'estimation et donc le biais d'échantillonnage ne peut être pris en compte dans ce modèle; d'autre part, il apparaît des études récentes que ce modèle donne des estimateurs relativement robustes [voir RIDDER et VERBAKEL (1984), JOUTARD et WERQUIN (1992)]. Le second modèle est un modèle de hasard proportionnel de type Weibull. Ce modèle dans sa formalisation générale est le même que celui de Cox, à la seule différence que la variable de durée est supposée suivre une loi de Weibull. L'intérêt de cette modélisation, outre le fait de proposer une estimation simple des effets des caractéristiques individuelles sur la durée du chômage, est de permettre la prise en compte du biais de sélection issu de l'enquête utilisée.

³Un bref rappel sur les notions élémentaires de l'analyse économétrique des durées se trouve en annexe.

2. PRESENTATION DES DONNEES ET DU BIAIS DE SELECTION.

2.1 L'échantillon retenu.

L'enquête « suivi des chômeurs » a été réalisée par la division emploi de l'INSEE de 1986 à 1988. L'échantillon, constitué de 8238 demandeurs d'emploi, a été tiré de manière aléatoire dans le fichier des chômeurs inscrits à l'ANPE au 31 août 1986.

Les personnes tirées au sort fin août 1986 (date E_0), sont ensuite interrogées à quatre dates successives: novembre 1986 (E_1), mai 1987 (E_2), novembre 1987 (E_3) et mai 1988 (E_4). L'enquête définitive couvre donc une période totale de 21 mois sur laquelle les mêmes personnes ont été interrogées, qu'elles aient ou non déménagé. En E_1 , les individus sont interrogés sur leur situation en E_0 ainsi que sur la durée de la période de chômage en cours. Entre les dates E_2 et E_4 , ils doivent reconstituer à l'aide du calendrier mensuel du questionnaire complémentaire leur situation d'emploi ou de chômage depuis l'enquête précédente. On obtient donc, pour chaque individu, un historique sur sa situation commençant à la date de début de la période de chômage en cours en E_0 et se terminant en E_4 . Cet historique peut éventuellement contenir plusieurs épisodes de chômage et si l'un d'eux est toujours en cours à la date E_4 , la durée de chômage sera considérée comme censurée à droite.

Au total 9396 périodes de chômage sont observées entre août 1986 et mai 1988 dont 6900 en cours en novembre 1986, 2664 censurées à droite, 4328 détenues par des hommes et 5068 par des femmes. A l'issue du chômage, trois types de sorties sont considérés: l'emploi « *stable* », l'emploi « *précaire* », et l'*inactivité* selon la terminologie en vigueur⁴. L'emploi « *stable* » regroupe les contrats à durée indéterminée, à temps partiel ou complet, ainsi que les personnes travaillant à leur compte ou aidant un membre de leur famille. L'emploi « *précaire* » représente l'ensemble des contrats à durée déterminée, les contrats d'intérim, le travail saisonnier, les contrats d'adaptation ou de qualification, les stages d'initiation à la vie professionnelle (SIVP), les travaux d'utilité collective (TUC), les apprentis et les stages ou formations rémunérés. L'*inactivité* désigne les personnes poursuivant des études ou effectuant des stages non rémunérés, les militaires du contingent, les retraités, préretraités ou les personnes dispensées de recherche d'emploi, les femmes au foyer ainsi que toute autre inactif.

⁴Cette terminologie est due à Atkinson et Micklewright (1990). Elle permet de différencier les emplois selon la durée de leur contrat et de schématiser les différentes situations rencontrées sur le marché du travail.

Les variables retenues sont les suivantes: le sexe, l'état matrimonial, l'âge classé en cinq tranches, la nationalité, le niveau d'étude atteint, la tranche d'unité urbaine, la circonstance d'entrée au chômage, le nombre de périodes déjà vécues par le demandeur d'emploi, le taux de chômage⁵. Notons que l'état matrimonial n'a été introduit dans les modèles que pour l'échantillon stratifié homme/femme. Excepté pour le taux de chômage, toutes les variables sont qualitatives et se scindent en différentes modalités. Leur traitement dans les modèles nécessite donc la création d'autant de variables binaires moins une afin d'éviter des problèmes de colinéarité.

2.2 Problème lié à la sélection endogène.

Le mode de sélection de l'enquête constitue une difficulté dans l'estimation des modèles. En effet, du fait de la liaison existant entre le mode de sélection et la variable aléatoire de durée, GOURIEROUX et MONFORT (1991) ont montré que les modèles traditionnels n'offrent pas une estimation robuste des taux de sortie du chômage. Par exemple, si l'on considère d'une part l'origine du temps comme la date de tirage de l'échantillon, E_0 , et d'autre part les individus entrés au chômage à la date $-e$, alors ne subsisteront dans l'échantillon que les individus dont la période de chômage sera $T > e$. *« Les périodes les plus courtes sont ainsi exclues de la base du sondage pour chaque cohorte d'entrants, ce qui va naturellement biaiser l'estimation de la loi des durées ».*

Notons par t , la valeur de la durée observée et par $-e$, la valeur de la date d'entrée dans la période qui nous intéresse. On considère la durée t indépendante de la date d'entrée $-e$. De ce fait, on peut affirmer que, pour un échantillon de taille N , les observations (t_i, e_i) , $i=\{1, \dots, N\}$, sont indépendantes, de même loi qui se décompose en $g_0(e)f_0(t; \theta_0)$, θ_0 représentant la vraie valeur du paramètre.

Si les observations sont toujours indépendantes d'un individu à l'autre et de même loi, cette loi sur l'échantillon n'est pas la même que la loi initiale relative à la population. Pour tout dire, les observations sont menées conditionnellement au fait que $T > e$. La loi conditionnelle à la date d'entrée admet donc la densité:

$$f_0(t|e; \theta_0) = \frac{g_0(e)f(t; \theta_0)}{\int_{t>e} g_0(e)f(t; \theta_0)dt} = P[t|T > e],$$

⁵Cette variable est mesurée tous les ans dans chaque département.

et en introduisant la fonction de survie associée à T , $S(t) = \int_t^\infty f(u) du$, nous obtenons:

$$f_0(t|e; \theta_0) = \frac{f(t; \theta_0)}{S(e; \theta_0)}.$$

GOURIEROUX et MONFORT font remarquer que cette loi conditionnelle ne dépend des vraies lois inconnues f_0 et g_0 , que par l'intermédiaire du paramètre θ_0 , et que dès lors elle peut être utilisée comme solution du problème:

$$\text{Max}_{\theta} \sum_{i=1}^N \text{Log} \frac{f(t_i; \theta)}{S(e_i; \theta)}. \quad (2.1)$$

L'estimateur du maximum de vraisemblance conditionnelle est convergent, asymptotiquement normal, et sa matrice de variance-covariance asymptotique s'écrit:

$$\Omega = E \left[- \frac{\partial^2}{\partial \theta \partial \theta'} \text{Log} \frac{f(T; \theta_0)}{S(e; \theta_0)} \right]^{-1}.$$

Dans le cas de notre échantillon, on observe deux types de durées de chômage soumises à la sélection endogène:

- celles parfaitement connues pour lesquelles la vraisemblance s'écrit:

$$L_1(t|y; \theta) = \frac{h(t; \theta) S(t; \theta)}{S(e; \theta)};$$

- celles censurées à droite pour lesquelles la vraisemblance s'écrit:

$$L_2(t|y; \theta) = \frac{S(t; \theta)}{S(e; \theta)}.$$

Par conséquent, la vraisemblance conditionnelle de l'échantillon s'écrit:

$$L_c(t|e; \theta) = \prod_{i=1}^N \left[\frac{f(t_i; \theta)}{S(e_i; \theta)} \right]^{d_i} \left[\frac{S(t_i; \theta)}{S(e_i; \theta)} \right]^{1-d_i} = \prod_{i=1}^N \left[\frac{h^{d_i}(t_i; \theta) S(t_i; \theta)}{S(e_i; \theta)} \right]$$

où d_i représente la variable indicatrice de censure de l'individu i .

La log-vraisemblance conditionnelle devient donc dans ce cas [CASES et LOLLIVIER (1993)]:

$$\ell_c(\theta) = \sum_{i=1}^N \left[d_i \log h(t_i; \theta) - \int_{e_i}^{t_i} h(u; \theta) du \right] \quad (2.2)$$

- avec
- t_i , la durée observée de l'individu i ;
 - e_i la durée passée au chômage avant la date E_0 ;⁶
 - d_i , la variable indicatrice de censure;
 - θ , le paramètre inconnu à estimer.

3. LES MODELES ESTIMES.

Les modèles choisis pour ce travail l'ont été dans la famille des modèles dits à hasard proportionnel introduits par COX (1972). Ces modèles, les plus utilisés en économétrie, ont fait l'objet de nombreuses études théoriques⁷ et empiriques⁸.

La fonction de hasard, conditionnelle au vecteur x constant, qui caractérise la probabilité conditionnelle de sortie du chômage, se décompose en deux fonctions:

$$h(t|x) = \varphi(x) \cdot h_0(t) \quad (3.1)$$

où $h_0(\cdot)$ est appelé hasard de base (c'est à dire pour $x=0$) et $\varphi(x)$ représente le lien entre le hasard conditionnel et les variables explicatives observées et désignées par x .

D'une autre manière, $h_0(\cdot)$ représente le lien entre le hasard et la durée passée au chômage. Dans ce modèle, $\varphi(x)$ correspond à une constante de proportionnalité entre les hasards $h(t|x)$ et $h_0(t)$. On retient que pour $x=0$, $\varphi(0)=1$, par conséquent, une forme naturelle pour φ est donnée par l'expression:

$$\varphi(x;\beta) = \exp(x' \beta)$$

- où
- x' est un vecteur ligne de p variables exogènes, $x'=(x_1, x_2, \dots, x_p)$;
 - β est un vecteur ($p \times 1$) de paramètres inconnus, $\beta'=(\beta_1, \beta_2, \dots, \beta_p)$.

Considérons donc la fonction de hasard suivante pour modéliser les probabilités conditionnelles de sortie du chômage:

$$h(t|x) = \exp(x' \beta) \cdot h_0(t), \quad (3.2)$$

- où
- $h_0(t)$ est appelé hasard de base (pour $x=0$),
 - x est le vecteur des p variables exogènes, $x'=(x_1, x_2, \dots, x_p)$,
 - β est le vecteur des p paramètres inconnus à estimer, $\beta'=(\beta_1, \beta_2, \dots, \beta_p)$.

⁶Ei = 0 si la période n'est pas soumise à la sélection endogène.

⁷Voir ELBERS et RIDDER (1982), HECKMAN et SINGER (1982, 1984), KIEFER (1984,1985), RIDDER (1990).

⁸Voir LANCASTER (1979,1983 et 1985), LANCASTER et NICKELL (1980), HECKMAN et BORJAS (1980), MOFFIT (1985), FLINN et HECKMAN (1982), DYNARSKI et SHEFFRIN (1990), KATZ et MEYER (1990), FALICK (1991), VISSER (1992), MEULDERS et PLASMAN (1993) et le numéro spécial du *journal of econometrics* (1985).

Cette forme est très souvent utilisée par les économètres car elle permet une interprétation simple des paramètres associés aux variables explicatives. En effet, nous avons:

$$\beta_p = \frac{\partial \log h(t|x;\beta)}{\partial x_p},$$

ce qui permet d'interpréter l'effet de la variable x_p sur le taux de hasard. Un paramètre positif entraîne une hausse du hasard et, inversement, un paramètre négatif entraîne une baisse de ce taux.

A partir du modèle (3.2), deux façons existent de traiter le problème avec des données censurées.

La première consiste à réaliser une estimation semi-paramétrique qui ne porte que sur les seules variables explicatives. Cette version du modèle, encore appelé modèle de Cox, présente dans notre cas l'intérêt d'éviter la prise en compte du biais d'échantillonnage avec la méthode explicitée en 2.2.

La seconde façon de procéder est de définir complètement le modèle (3.2), en particulier, en formalisant la loi de la durée. Ainsi, le modèle étant parfaitement défini, une estimation paramétrique est possible et avec elle, la correction du biais d'échantillonnage. Le plus souvent, la loi statistique décrivant les durées est choisie dans la famille des lois de Weibull. Outre le fait d'être facile à utiliser, l'avantage de cette loi est de donner des taux de hasard strictement croissants ou décroissants et donc de permettre une analyse plus facile de l'évolution de ces taux au cours du temps.

3.1 Le modèle semi-paramétrique de Cox.

On dit du modèle défini en (3.2), qu'il est semi-paramétrique si le hasard de base $h_0(t)$ n'est pas contraint par une loi statistique. Autrement dit, le hasard de base est laissé libre ce qui entraîne plus de souplesse dans le modèle. Cette approche conduit alors à estimer les seuls paramètres contenus dans β par maximisation d'une vraisemblance partielle⁹ dont la construction repose sur l'idée suivante: En ordonnant l'échantillon selon les valeurs croissantes des durées observées t_i , on décompose l'ensemble des données en

⁹Voir LANCASTER (1990, pp 234-238).

deux informations séparées dépendantes. L'une est relative au « label », noté $\mathcal{J}_j$, des individus classés, l'autre contient les seules durées t_j ordonnées telles que $t_1 < t_2 < \dots < t_n$ ¹⁰.

La vraisemblance partielle du modèle de Cox se note alors de la manière suivante:

$$L_p(\beta) = \prod_{j=1}^n \frac{\exp(x'_j \beta)}{\sum_{k \in \mathcal{R}_j} \exp(x'_k \beta)}. \quad (3.3)$$

Les procédures du maximum de vraisemblance¹¹ s'appliquent alors sur la fonction (3.3) car elle possède l'essentiel des propriétés de la fonction de vraisemblance traditionnelle. En particulier, les estimateurs du maximum de vraisemblance partielle (MVP) possèdent les propriétés asymptotiques habituelles, et un estimateur convergent de leur variance est donné avec l'inverse de la matrice d'information évaluée avec les estimateurs MVP. De plus, la pratique montre que les estimateurs obtenus avec ce modèle ont une bonne fiabilité et sont suffisamment robustes.

3.2 Le modèle de Weibull.

Avec les spécifications d'une loi de Weibull pour décrire la loi de la durée, le modèle de hasard proportionnel (3.1) devient:

$$h(t|x) = \alpha \lambda \exp(x' \beta) t^{\alpha-1} \quad (3.4)$$

où α et λ sont les paramètres à estimer d'une loi de Weibull.

Précisons que si $\alpha > 1$ alors le taux de hasard est monotone croissant et, inversement, si $\alpha < 1$ alors le taux de hasard est monotone décroissant.

En utilisant la vraisemblance conditionnelle à la date d'entrée au chômage explicitée en 2.2 et en tenant compte des différents schémas de censure, la log-vraisemblance conditionnelle à la date d'entrée au chômage devient donc:

$$\ell_c = \sum_{i=1}^N \left[d_i \left(x'_i \beta + \log h_0(t_i) \right) - \exp(x'_i \beta) \int_{e_i}^{t_i} h_0(u) du \right]$$

et avec une spécification de Weibull pour le hasard de base:

¹⁰Voir COX et OAKES (1984) pour plus de détails.

¹¹La procédure PHREG incluse dans SAS v.6.07 considère le modèle semi-paramétrique de Cox et calcule un estimateur non paramétrique du hasard de base et des estimateurs paramétriques de β des coefficients associés aux covariables affectant le hasard de base sous la forme $\exp(x' \beta)$.

$$\ell_c(\alpha, \beta, \lambda) = \sum_{i=1}^N \left[d_i \left(x_i' \beta + \log \alpha + \log \lambda + (\alpha - 1) \log t_i \right) - \lambda \exp(x_i' \beta) (t_i^\alpha - e_i^\alpha) \right]$$

avec

- t_i , la durée observée de l'individu i ;
- e_i , la durée passée au chômage avant la date E_0 ;
- d_i , la variable indicatrice de censure;
- x_i , le vecteur des variables individuelles caractéristiques;
- β , le vecteur des paramètres inconnus à estimer
- α et λ , les paramètres à estimer d'une loi de Weibull.

3.3 Modèles de transitions multiples.

Nous avons également tenu compte dans cette étude de la possibilité de distinguer entre plusieurs destinations. Si cette hypothèse ne change en rien la nature des deux modèles explicités dans les deux sous-sections précédentes, il faut cependant développer quelques notions propres aux modèles de transitions multiples.

Rappelons que ces modèles considèrent les *intensités de transitions vers la destination k* , $k=\{1,2,3\}$, et que la fonction de hasard est la somme de ces intensités

$$h(t) = \sum_{k=1}^3 h_k(t)$$

Les modèles de transitions multiples définissent également la probabilité marginale de chacune des destinations comme la probabilité, lorsqu'un changement intervient, que sa destination soit k . Notée π_k , elle s'obtient par l'intégrale suivante:

$$\pi_k = \int_0^{\infty} S(u) h_k(u) du,$$

et il est facile de montrer que $\sum_{k=1}^3 \pi_k = 1$.

La fonction de survie à la date t s'écrit:

$$S(t) = \sum_{k=1}^K \pi_k S_k(t),$$

et correspond à une distribution mixée [LANCASTER (1990)]. $S_k(t)$ est la probabilité de rester dans le même état sachant que si une transition survient, elle sera vers k .

Ainsi, avec des observations portant sur un échantillon de taille N , la fonction log-vraisemblance conditionnelle à la date d'entrée au chômage, $-e_i$, s'écrit:

$$\ell_c = \log L = \sum_{i=1}^N \sum_{k=1}^3 \left[d_{ik} \log h_k(t_i) - \int_{e_i}^{t_i} h_k(u) du \right],$$

où d_{ik} est une variable qui indique une transition vers la destination k .

4. RESULTATS.

Dans cette section, le travail qui est présenté est le résultat de deux séries d'estimations, l'une réalisée à partir d'un modèle semi-paramétrique de Cox, l'autre avec un modèle de Weibull. L'intérêt est donc, ici, d'évaluer les effets de certaines caractéristiques individuelles sur la durée du chômage, avec et sans prise en compte du biais de sélection. Ainsi, par l'intermédiaire de ces deux modèles, nous pourrions mesurer les conséquences de l'oubli du biais d'échantillonnage sur ces déterminants. L'ensemble des résultats est présenté dans les tableaux situés en annexe II et dans lesquels figurent les valeurs des paramètres associés aux variables retenues. Les chiffres en gras donnent les valeurs significatives des paramètres par rapport à la statistique de Student pour un seuil de 10%¹². Précisons, également, que les estimations ont d'abord été effectuées sur l'ensemble de l'échantillon puis sur deux sous-groupes, celui des hommes et celui des femmes. Dès lors, il apparaît à la lecture des résultats que:

- Le fait *d'être une femme* diminue, pour les deux modèles, la probabilité de sortir du chômage avec, cependant, un effet inverse pour une transition vers l'inactivité. Ceci montre que les femmes sont plus concernées que les hommes par les sorties du marché du travail. De plus, l'emploi stable est l'issue vers laquelle les femmes ont une probabilité plus faible. La comparaison de l'ensemble des résultats des deux modèles fait apparaître que l'effet de cette variable est toujours plus marqué dans le modèle de Cox que dans le modèle de Weibull. Ce qui signifie que, dans notre étude, l'introduction dans les estimations de la méthode corrigeant le biais d'échantillonnage atténue cet effet.

- Les variables concernant *l'état matrimonial* ont été introduites uniquement au niveau des sous-échantillons hommes/femmes. Leurs effets montrent des résultats ambigus et difficilement interprétables. Toutefois, le modèle de Weibull donnent, pour ces variables, des paramètres beaucoup plus significatifs et marqués que l'autre.

¹²L'immense majorité des paramètres significatifs le sont pour un seuil d'acceptation bien inférieur (5% voire 1%).

- *L'effet de l'âge* est très significatif sur les probabilités conditionnelles de sortie du chômage. Plus exactement, il apparaît que la durée du chômage augmente avec l'âge. Etre âgé de moins de 25 ans augmente le taux instantané de sortie, surtout vers l'emploi précaire, alors qu'un individu de plus de 45 ans a de moins en moins de chance de retrouver un emploi, quelle que soit la nature de celui-ci. Notons que le signe des paramètres reste le même pour les hommes et les femmes. D'une façon générale, on constate qu'un jeune a d'abord plus de chance de trouver un emploi précaire, puis de basculer dans l'inactivité, et enfin de trouver un emploi stable. Ce résultat est en fait conforme à leur inexpérience sur le marché du travail¹³. Encore une fois, la comparaison des deux modèles montre que les paramètres du modèle de Cox sont beaucoup plus marqués que ceux du modèle de Weibull. Par conséquent, tenir compte du biais de sélection dans nos estimations modère les effets de l'âge sur la durée du chômage.

- *L'effet de la nationalité* est pour l'ensemble plus accentué avec le modèle de Cox. De plus, il apparaît qu'un travailleur étranger a une probabilité plus faible de retrouver un emploi qu'un travailleur français. Ce constat s'inverse pour un individu membre de la communauté européenne, ce qui s'explique sans doute par le haut degré de qualification qui caractérise cette immigration.

- le *niveau d'étude* atteint a un rôle déterminant sur les probabilités conditionnelles de sortie du chômage. Ces effets, toujours plus marqués avec le modèle de Cox, montrent que moins le diplôme est élevé plus la probabilité conditionnelle de sortie du chômage est faible. Ce constat reste valable quel que soit l'emploi occupé à l'issue du chômage sauf lorsque la transition se fait vers l'inactivité, dans ce cas les paramètres ne sont pas significatifs.

- *Habiter la région parisienne* accroît les chances de sortie du chômage quels que soient le modèle estimé et le sous-échantillon considéré (hommes ou femmes). Toutefois, dans notre cas, oublier le biais d'échantillonnage contribue à renforcer cet effet sur les probabilités conditionnelles de sortie du chômage.

- *Les licenciements* contribuent à diminuer la probabilité pour retrouver un emploi, aussi bien chez les hommes que chez les femmes, même si chez ces dernières, cet effet est légèrement plus marqué. La distinction entre plusieurs sorties possibles permet le constat suivant: une période de chômage causée par une démission augmente la probabilité de basculer dans l'inactivité alors qu'une période causée par la fin d'un

¹³Voir TESTE (1996).

emploi précaire favorise le statu quo, c'est à dire la précarité des emplois retrouvés. Excepté pour les circonstances non déclarées de l'origine du chômage où le signe des paramètres estimés avec les deux modèles sont opposés, il apparaît encore une fois que les effets de cette variable sont plus déterminants dans le modèle de Cox.

- La *récurrence* du chômage montre que plus un individu connaît de périodes sans emploi, plus sa probabilité de sortir du chômage est élevée, ceci étant valable aussi bien pour les hommes que pour les femmes. Ce constat est particulièrement vrai pour une issue vers les emplois précaires où, dans ce cas, les paramètres sont beaucoup plus significatifs. Cela révèle sans doute les difficultés que rencontrent les demandeurs d'emploi pour quitter le marché des emplois précaires et s'insérer sur celui des emplois stables. La comparaison directe des estimations du modèle de Cox et du modèle de Weibull pour cette variable montre un effet plus important dans le premier modèle.

- Le *taux de chômage*, censé mesurer les tensions sur le marché du travail, a des effets très peu significatifs qui n'influencent pas les probabilités de sortie du chômage.

Enfin, pour terminer l'analyse des tableaux de l'annexe II, nous pouvons décrire l'évolution générale des taux instantanés de sortie du chômage. Les estimations du paramètre α du modèle de Weibull nous renvoient, dans la plupart des cas, à des probabilités monotones décroissantes ($\alpha < 1$) par rapport à la durée passée dans le chômage. Ces probabilités deviennent monotones croissantes ($\alpha > 1$) uniquement lorsque l'issue observée est l'inactivité, ce qui signifie que les sorties du marché du travail sont associées à de longues périodes de chômage. Ce constat conforte l'idée selon laquelle on devient inactif par découragement, suite à une recherche d'emploi trop longue et sans succès.

5. CONCLUSION.

Indépendamment des résultats économiques qui montrent l'influence des déterminants individuels sur la durée du chômage, l'analyse des tableaux des estimations fait apparaître, deux aspects majeurs:

1) Les signes des paramètres dans les deux modèles sont identiques, excepté pour les circonstances non déclarées de l'origine du chômage. Autrement dit, que l'on prenne en compte ou non le biais d'échantillonnage, ne change rien à l'effet des variables individuelles sur la probabilité conditionnelle de sortie du chômage. De ce point de vue,

les résultats traduisent encore une fois la bonne robustesse des estimateurs obtenus avec le modèle de Cox. Ainsi les deux modèles font apparaître comme effet pénalisant sur la durée du chômage, le fait d'être une femme, d'être âgé de plus de 35 ans, de ne posséder aucun diplôme ou d'être étranger. Dans les deux modèles, il apparaît également qu'un chômage ayant pour origine un licenciement diminue les chances de retrouver un emploi. Enfin, le nombre de périodes de chômage vécues influence les taux de sortie dans le sens où plus ce nombre est grand, plus la probabilité de sortir du chômage vers le marché du travail précaire est élevée.

2) Si les signes des paramètres sont identiques dans les deux modèles, il apparaît en revanche dans les résultats que les estimateurs obtenus avec le modèle de Cox révèlent toujours des effets plus forts que ceux obtenus avec le modèle de Weibull. Autrement dit, les effets des variables individuelles sur la durée du chômage sont toujours surestimés dans le modèle de Cox comparativement à ceux obtenus dans l'autre modèle¹⁴.

Ce second aspect des résultats peut, en particulier, être mis en évidence avec les graphiques qui figurent en annexe III. Ces graphiques représentent l'évolution des probabilités conditionnelles de sortie du chômage lorsque le biais d'échantillonnage est pris en compte et lorsqu'il ne l'est pas. Dans le premier cas, les taux de hasard sont calculés à partir des estimations du modèle de Weibull c'est à dire des paramètres associés aux variables explicatives et au hasard de base de type Weibull. Dans le second cas, les taux de hasard sont calculés avec les résultats des estimations du modèle de Cox pour ce qui est des variables explicatives, alors que le hasard de base est repris du modèle précédent. De cette sorte, l'écart entre les taux de hasard sur chaque graphique représente uniquement l'effet de surestimation constaté dans les résultats et qui a pour conséquence de sous-évaluer de moitié les probabilités conditionnelles de sortie du chômage.

Le travail présenté dans cet article vise à montrer les conséquences de la non prise en compte du biais d'échantillonnage sur les déterminants individuels de la durée du chômage. A partir de deux modèles bien distincts, le modèle de Cox et un modèle de Weibull, nous avons trouvé que le biais d'échantillonnage influence peu l'effet des variables individuelles sur la probabilité conditionnelle de sortie du chômage. C'est à dire

¹⁴Ce constat entraîne parfois un manque de significativité des paramètres dans l'un ou l'autre des modèles.

qu'un effet positif dans un modèle est le plus souvent positif dans l'autre modèle et réciproquement pour un effet négatif. En revanche, l'oubli du biais d'échantillonnage revient toujours à surestimer l'effet de ces variables sur les taux de sortie du chômage. Par conséquent, même si les effets des exogènes sur les taux de sortie gardent le même signe, quel que soit le modèle estimé, la surestimation de ces effets, indépendamment de la sous-estimation des taux de sortie qu'elle entraîne, montre la nécessité et l'importance des méthodes pour corriger le biais de sélection et en particulier, celle proposée par GOURIEROUX et MONFORT.

BIBLIOGRAPHIE.

- Atkinson A. B. et J. Micklewright**, 1991, Unemployment Compensation and Labor Market Transitions: A Critical Review, *Journal of Economic Literature*, vol. XXIX, pp. 1679-1727.
- Bonnal L., Fougère D.**, 1990, « Les déterminants individuels de la durée du chômage », *Economie et Prévision*, 1990-5, 96, p.45-82.
- Bonnal L., Fougère D. et Serandon**, 1994, « Evaluating the impact of french employment policies on individual labour market histories », document de travail du CREST et du département de la recherche n° 9417, INSEE, mars 1994.
- Cases C., Lollivier S.**, 1993, « Estimation de la durée du chômage en France en 1986 », document de travail du CREST et du département de la recherche n° 9309, INSEE, mars 1993.
- Cases C. et S. Lollivier**, 1994, « Estimation d'un modèle de sortie de chômage à destinations multiples », *Economie et Prévision*, n°113-114.
- Cox D. R.**, 1972, « Regression model and life tables », *Journal of Royal Statistical Society, B*, 34, 187-220.
- Cox D. R.**, 1975, « Partial Likelihood », *Biometrika*, vol.62, n°2, p.269.
- Cox D.R. et D. Oakes**, 1984, « Analysis of Survival Data », *Monographs on Statistics and Applied Probability* - Chapman and Hall.
- Dynarski M. et S.M. Sheffrin**, 1990, « The Behavior of Unemployment Durations Over the Cycle », *The Review of Economic and Statistics*, vol.72, n° 2, may.
- Elbers C. et G. Ridder**, 1982, « True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model », *The Review of Economic Studies*, 49, july, pp.403-409.
- Fallick B.C.**, 1991, « Unemployment Insurance and the Rate of Re-employment of Displaced Workers », *The Review of Economic and Statistics*, vol.73, n° 2, may.
- Flinn C.J. et J.J. Heckman**, 1982, « New Methods for Analyzing Structural Models of Labor Force Dynamics », *Journal of Econometrics*, Vol.18 pp. 115-168.
- Gourieroux C., Monfort A.**, 1991, « Modèles de durée et effets de génération », document de travail du CREST et du département de la recherche n° 9125, INSEE, octobre 1991.

- Heckman J.J. et G.J. Borjas**, 1980, « Does Unemployment cause Future Unemployment? Definitions, Questions and Answers from a continuous Time Model of Heterogeneity and State dependence », *Economica*, vol.47, 247-283.
- Heckman J.J. and B. Singer**, 1984a, « Econometric Duration Analysis », *Journal of Econometrics*, 24, 63-132.
- Heckman J.J. et B. Singer**, 1984b, « The Identifiability of the Proportional Hazard Model », *Review of Economic Studies*, vol.51 pp.231-241.
- Joutard X., Werquin P.**, « Les déterminants individuels de la durée du chômage: de l'intérêt de distinguer les emplois stables des emplois précaires », *Economie et Prévision*, n° 102-103, 1992.
- Katz L.F. et B.D. Meyer**, 1990, « The Impact of the Potential Duration of Unemployment Benefits on the Duration of Unemployment », *Journal of Public Economics*, vol.41, p.45-72.
- Kiefer N. M.**, 1988, « Economic Duration Data and Hazard Functions », *Journal of Economic Literature*, vol. XXVI, june, pp.646-679.
- Lancaster T.**, 1979, « Econometric Methods for the Duration of Unemployment », *Econometrica*, vol.47, n°4, july.
- Lancaster T.**, 1985, « Generalised Residuals and Heterogeneous Duration Models with Application to the Weibull Model », *Journal of Econometrics*, Annal I, vol.28, pp. 155-169.
- Lancaster T.**, 1990, « The Econometric Analysis of Transition Data », *Econometric Society Monograph*, Cambridge University Press.
- Lancaster T. et A.D. Chesher**, 1983, « An Econometric Analysis of Reservation Wages », *Econometrica* 51, 6 : 1661-1676.
- Lancaster T. and S. Nickell**, 1980, « The Analysis of Re-Employment Probabilities for the Unemployment », *Journal of the Royal Statistical Society*, vol. 143, n°2, pp. 141-165.
- Meulders D. et R. Plasman**, 1993, « Formation Initiale et Politique de Formation des Chômeurs: Estimation de leur Impact sur la Durée du Chômage », *Communication au congrès de L'AFSE- Dijon mai 1993*.
- Ridder G. and W. Verbakel**, 1983, « On the estimation of the proportionnal hazard model in the presence of heterogeneity », *University of Amsterdam, Faculty of Actuarial Science and Econometrics, A&E Report 22/83*.
- Teste T.**, 1995, « Analyse des durées du chômage en France et biais d'échantillonnage », *XV^e journées de L'Association d'Economie Sociale - L'analyse longitudinale en économie sociale - Nancy, septembre 1995*.
- Teste T.**, 1996, « Les déterminants individuels de la durée du chômage en France avec biais d'échantillonnage », *3^{ème} journées d'études du Céreq-Céretim-Lasmas-Idl, Rennes, mai 1996*, in DEGENNE A., MANSUY M., PODEVIN G. Et Werquin P., *Typologie des marchés du travail: suivi et parcours*, Céreq, document n° 115, Série séminaires, juillet 1996.
- Teste T.**, 1996, « Les modèles de durée: application pour un traitement micro-économique des durées de chômage », *Thèse de doctorat, Université de Bourgogne, novembre 1996*.
- Visser M.**, 1992, « Analysis of Labor Market Histories with Panel Data », *Document de Travail du CREST, INSEE, n° 9209, avril*.

ANNEXE I:

NOTIONS ELEMENTAIRES DE L'ANALYSE ECONOMETRIQUE DES DUREES

On définit T comme la variable aléatoire de durée, réelle et positive, de loi continue sur $\mathfrak{R}^+$ définie par:

i) $f(t)$, la fonction de densité de T (strictement positive);

ii) $F(t) = P[T \leq t] = \int_0^t f(u)du$, sa fonction de répartition.

Ces deux fonctions caractérisent la loi de la variable aléatoire T .

Les fonctions de hasard et de survie.

La fonction de survie en t , notée $S(t)$, correspond à la probabilité pour que la durée d'une période dépasse t , c'est à dire:

$$S(t) = P[T > t] = 1 - F(t)$$

et caractérise également la loi de la variable aléatoire T .

Le hasard pour la variable aléatoire T , noté $h(t)$, correspond à la limite en probabilité pour que la période s'achève dans l'intervalle $[t, t+dt]$ lorsque dt tend vers 0 et sachant que $T > t \geq 0$, c'est à dire:

$$h(t) = \lim_{dt \rightarrow 0} \frac{P[t < T < t + dt | T > t]}{dt} = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \geq 0$$

Il s'interprète comme le **taux instantané de sortie** de l'état considéré. De plus, il est facile de montrer le lien qui existe entre ces deux fonctions et qui est donné par l'expression suivante:

$$S(t) = \exp \left\{ - \int_0^t h(u)du \right\}.$$

Construction de la fonction de vraisemblance dans le cas de données censurées à droite.

La difficulté majeure dans l'estimation des modèles de durée provient de l'existence de durées censurées, c'est à dire de durées dont la fin ne peut être observée. On parle alors de censure à droite.

Considérons la variable aléatoire binaire d telle que:

$$\begin{cases} d = 1 & \text{s'il n'y a pas censure} \\ d = 0 & \text{s'il y a censure} \end{cases}$$

Dès lors, deux cas sont envisageables:

- i) lorsqu'il n'y a pas de censure, la durée est complètement observée et la densité s'écrit comme le produit du hasard et de la survie, $f_1(t) = h(t) \cdot S(t)$;
- ii) lorsqu'il y a censure, la durée s'étend au-delà de la période d'observation et la fonction de densité s'identifie à la survie, $f_2(t) = S(t)$.

En raisonnant conditionnellement à T , la densité jointe du couple (t, d) s'écrit:

$$\begin{aligned} f(t, d) &= f_1(t)^d \cdot f_2(t)^{1-d} \\ &= h(t)^d \cdot S(t) \end{aligned}$$

Pour N observations indépendantes, la fonction de vraisemblance complète s'écrit, pour $i = \{1, 2, \dots, N\}$:

$$L(t, d) = \prod_{i=1}^N f(t_i, d_i)$$

et la log-vraisemblance:

$$\ell(t, d; \theta) = \sum_{i=1}^N d_i \log h(t_i; \theta) + \sum_{i=1}^N \log S(t_i; \theta)$$

où θ est un paramètre inconnu à estimer.

ANNEXE II-A: Résultats des estimations pour l'ensemble de l'échantillon.

N=9396 Paramètres	Modèles semi-paramétrique de Cox				Modèles de Weibull			
	Ensemble	Emploi stable	Emploi précaire	Inactivité	Ensemble	Emploi stable	Emploi précaire	Inactivité
Être un homme	<i>réf.</i>				<i>réf.</i>			
Être une femme	-0,306	-0,390	-0,370	0,194	-0,276	-0,362	-0,352	0,275
Moins de 25 ans	0,497	0,222	0,655	0,534	0,414	0,140	0,585	0,381
Entre 25 et 35 ans	<i>réf.</i>				<i>réf.</i>			
Entre 35 et 45 ans	-0,063	0,112	-0,127	-0,510	-0,043	0,127	-0,107	-0,533
Entre 45 et 55 ans	-0,399	-0,174	-0,582	-0,504	-0,349	-0,127	-0,541	-0,469
55 ans et plus	-1,052	-1,811	-2,193	0,530	-1,026	-1,824	-2,183	0,672
Être français	<i>réf.</i>				<i>réf.</i>			
Être européen	0,215	0,200	0,241	0,149	0,135	0,123	0,171	0,018
Être étranger	-0,059	-0,351	0,169	-0,460	-0,087	-0,367	0,148	-0,446
Diplôme non déclaré	0,096	-0,292	0,125	0,884	0,465	0,085	0,514	1,295
Dip. supérieur	-0,089	-0,012	-0,140	-0,514	-0,050	0,047	-0,114	-0,420
Bac +2ans	<i>réf.</i>				<i>réf.</i>			
Bac ou BP	-0,124	-0,369	-0,051	0,285	-0,090	-0,322	-0,018	0,388
CAP ou BEP	-0,258	-0,369	-0,221	-0,150	-0,216	-0,314	-0,181	-0,012
BEPC seul	-0,284	-0,483	-0,235	0,085	-0,224	-0,404	-0,175	0,210
Aucun diplôme	-0,475	-0,690	-0,416	-0,128	-0,381	-0,574	-0,340	0,076
Communes rurales	0,046	0,135	-0,005	0,044	0,003	0,096	-0,049	0,011
Moins de 20000 hbt	0,042	-0,027	0,081	0,040	-0,022	-0,085	0,022	-0,027
De 20000 à 200000 hbt	<i>réf.</i>				<i>réf.</i>			
Plus de 200000 hbt	0,114	0,260	-0,016	0,277	0,042	0,197	-0,079	0,168
Agglo. parisienne	0,267	0,484	0,144	0,087	0,259	0,476	0,141	0,095
Circ. non déclarée	0,198	0,029	0,200	0,459	-0,475	-0,818	-0,383	-0,032
Fin de CDD	0,302	0,069	0,383	0,207	0,234	-0,016	0,324	0,187
Licenciement	<i>réf.</i>				<i>réf.</i>			
Démission ou autre	0,245	0,035	0,265	0,500	0,201	-0,020	0,232	0,485
Une période	-1,316	-1,278	-1,331	-1,316	-0,423	-0,331	-0,496	-0,359
Deux périodes	<i>réf.</i>				<i>réf.</i>			
Trois périodes	0,156	0,113	0,218	-0,540	0,138	0,079	0,203	-0,499
Quatre périodes et plus	0,489	0,224	0,668	-1,351	0,455	0,154	0,634	-1,131
Taux de chômage	-0,004	-0,019	0,003	0,009	0,011	-0,007	0,016	0,039
α					0,946	0,908	0,933	1,165
λ					0,145	0,078	0,087	0,002
log-vraisemblance	-54202,2	-18486,3	-28912,0	-6360,9	-21408,4	-9526,7	-13357,2	-4339,8

Notes: - les chiffres en gras donnent les valeurs significatives par rapport à la statistique de Student des paramètres pour un seuil de 10%
- un paramètre positif augmente la probabilité de sortie du chômage.

ANNEXE II-B: Résultats des estimations pour les hommes.

N=4328 Paramètres	Modèles semi-paramétrique de Cox				Modèles de Weibull			
	Ensemble	Emploi stable	Emploi précaire	Inactivité	Ensemble	Emploi stable	Emploi précaire	Inactivité
Etat matrimonial n.d.	<i>réf.</i>				<i>réf.</i>			
Célibataire	-0,159	-0,998	0,289	0,091	-1,142	-1,821	-0,610	-1,380
Marié	0,160	-0,568	0,545	0,105	-0,866	-1,427	-0,390	-1,452
Veuf	0,021	-0,430	0,004	-0,088	-1,009	-1,224	-0,876	-1,561
Divorcé	-0,055	-1,054	0,465	0,304	-1,040	-1,853	-0,444	-1,146
Moins de 25 ans	0,562	0,210	0,657	1,921	0,488	0,138	0,589	1,794
Entre 25 et 35 ans	<i>réf.</i>				<i>réf.</i>			
Entre 35 et 45 ans	-0,246	-0,104	-0,341	-1,203	-0,190	-0,059	-0,283	-1,113
Entre 45 et 55 ans	-0,624	-0,530	-0,838	0,624	-0,505	-0,434	-0,726	0,860
55 ans et plus	-1,220	-2,143	-2,325	2,142	-1,248	-2,201	-2,344	2,248
Être français	<i>réf.</i>				<i>réf.</i>			
Être européen	0,232	0,314	0,301	-0,903	0,267	0,321	0,329	-0,799
Être étranger	0,033	-0,399	0,355	-0,142	0,035	-0,391	0,366	-0,338
Diplôme non declare	-0,082	-1,121	0,434	0,732	-0,108	-1,066	0,458	0,317
Dip. supérieur	-0,092	-0,094	-0,018	-1,173	-0,128	-0,103	-0,081	-1,261
Bac +2ans	<i>réf.</i>				<i>réf.</i>			
Bac ou BP	-0,236	-0,542	-0,136	0,726	-0,235	-0,536	-0,138	0,683
CAP ou BEP	-0,245	-0,523	0,000	-0,044	-0,211	-0,483	0,034	-0,119
BEPC seul	-0,317	-0,594	-0,121	0,216	-0,229	-0,512	-0,026	0,285
Aucun diplôme	-0,496	-0,878	-0,230	0,123	-0,423	-0,774	-0,171	0,033
Communes rurales	0,132	0,203	0,107	0,133	0,081	0,162	0,054	0,056
Moins de 20000 hbt	0,132	-0,014	0,206	0,222	0,059	-0,081	0,140	0,069
De 20000 à 200000 hbt	<i>réf.</i>				<i>réf.</i>			
Plus de 200000 hbt	0,142	0,369	-0,050	0,317	0,077	0,306	-0,110	0,181
Agglo. parisienne	0,181	0,429	-0,005	0,070	0,174	0,422	-0,017	0,155
Circ. non déclarée	0,103	-0,134	0,169	0,569	-0,731	-1,166	-0,549	-0,008
Fin de CDD	0,295	0,029	0,402	0,342	0,200	-0,097	0,320	0,384
Licenciement	<i>réf.</i>				<i>réf.</i>			
Démission ou autre	0,254	-0,104	0,325	0,950	0,201	-0,139	0,273	0,884
Une période	-1,317	-1,267	-1,404	-0,855	-0,466	-0,378	-0,603	0,051
Deux périodes	<i>réf.</i>				<i>réf.</i>			
Trois périodes	0,120	0,093	0,129	-0,036	0,110	0,063	0,132	-0,041
Quatre périodes et plus	0,521	0,301	0,594	0,026	0,590	0,323	0,708	0,260
Taux de chômage	-0,009	-0,034	0,004	0,030	0,003	-0,025	0,015	0,056
α					0,972	0,933	0,994	1,092
λ					0,471	0,667	0,120	0,002
log-vraisemblance	-23426,2	-8433,4	-12648,1	-1926,8	-9677,8	-4499,3	-6216,8	-1482,2

Notes: les chiffres en gras donnent les valeurs significatives par rapport à la statistique de Student des paramètres pour un seuil de 10%.
un paramètre positif augmente la probabilité de sortie du chômage.
l'hypothèse d'exponentialité a été acceptée sur le paramètre souligné α du modèle de hasards proportionnels de type Weibull.

ANNEXE II-C: Résultats des estimations pour les femmes.

N=5068 Paramètres	Modèles semi-paramétrique de Cox				Modèles de Weibull			
	Ensemble	Emploi stable	Emploi précaire	Inactivité	Ensemble	Emploi stable	Emploi précaire	Inactivité
Etat matrimonial n.d.	<i>réf.</i>				<i>réf.</i>			
Célibataire	-0,431	-1,101	-0,227	-0,070	-1,161	-1,922	-0,936	-0,727
Mariée	-0,686	-1,374	-0,725	0,635	-1,425	-2,202	-1,440	-0,059
Veuve	-0,484	-1,098	-0,512	0,732	-1,273	-1,998	-1,266	-0,053
Divorcée	-0,490	-1,255	-0,227	0,107	-1,249	-2,094	-0,951	-0,627
Moins de 25 ans	0,440	0,267	0,576	0,259	0,329	0,169	0,484	0,101
Entre 25 et 35 ans	<i>réf.</i>				<i>réf.</i>			
Entre 35 et 45 ans	0,062	0,293	0,024	-0,430	0,070	0,305	0,033	-0,440
Entre 45 et 55 ans	-0,294	0,067	-0,445	-0,867	-0,277	0,111	-0,440	-0,880
55 ans et plus	-0,998	-1,659	-2,133	-0,169	-0,843	-1,522	-2,022	0,064
Être française	<i>réf.</i>				<i>réf.</i>			
Être européenne	0,186	0,060	0,142	0,578	0,028	-0,092	-0,002	0,361
Être étrangère	-0,206	-0,283	-0,083	-0,525	-0,288	-0,358	-0,155	-0,597
Diplôme non déclaré	0,041	-0,436	-0,153	1,304	0,377	-0,112	0,191	1,671
Dip. supérieur	-0,103	0,052	-0,255	-0,212	-0,027	0,141	-0,186	-0,133
Bac +2ans	<i>réf.</i>				<i>réf.</i>			
Bac ou BP	-0,034	-0,203	-0,014	0,182	0,007	-0,149	0,033	0,232
CAP ou BEP	-0,272	-0,256	-0,377	0,037	-0,221	-0,196	-0,324	0,097
BEPC seul	-0,243	-0,365	-0,302	0,202	-0,211	-0,321	-0,267	0,234
Aucun diplôme	-0,483	-0,563	-0,580	0,013	-0,383	-0,462	-0,492	0,173
Communes rurales	0,006	0,098	-0,041	-0,096	-0,034	0,062	-0,081	-0,110
Moins de 20000 hbt	-0,015	-0,027	0,020	-0,123	-0,063	-0,074	-0,031	-0,124
De 20000 à 200000 hbt	<i>réf.</i>				<i>réf.</i>			
Plus de 200000 hbt	0,112	0,218	0,029	0,193	0,061	0,169	-0,006	0,106
Agglo. parisienne	0,294	0,504	0,205	0,125	0,281	0,494	0,205	0,085
Circ. non déclaré	0,237	0,175	0,157	0,507	-0,240	-0,374	-0,263	0,017
Fin de CDD	0,292	0,134	0,318	0,221	0,276	0,123	0,305	0,209
Licenciement	<i>réf.</i>				<i>réf.</i>			
Démission ou autre	0,223	0,154	0,175	0,414	0,237	0,156	0,193	0,397
Une période	-1,327	-1,335	-1,254	-1,627	-0,414	-0,361	-0,400	-0,623
Deux périodes	<i>réf.</i>				<i>réf.</i>			
Trois périodes	0,191	0,143	0,312	-0,895	0,170	0,104	0,287	-0,836
Quatre périodes et plus	0,471	0,142	0,754	-15,681	0,327	-0,108	0,607	-
Taux de chômage	-0,006	-0,012	-0,006	0,013	0,010	0,003	0,007	0,041
α					0,944	0,918	0,901	1,232
λ					0,381	0,282	0,249	0,004
log-vraisemblance	-26070,5	-8406,1	-13693,1	-3699,4	-11609,5	-4938,0	-7040,1	-2748,2

Notes: les chiffres en gras donnent les valeurs significatives par rapport à la statistique de Student des paramètres pour un seuil de 10%.
un paramètre positif augmente la probabilité de sortie du chômage.

ANNEXE III: COMPARAISON DES TAUX DE HASARD

