

Validité de l'analyse markovienne de structure latente pour l'estimation de l'erreur de classification des données sur la population active

Paul P. Biemer et John M. Bushery¹

Résumé

Les auteurs ont surtout cherché à vérifier la validité des estimations MLCA (analyse markovienne de structure latente) de l'erreur de classification de la population active et à évaluer la possibilité que ce type d'analyse remplace les méthodes traditionnelles d'évaluation de la qualité des données. Les auteurs ont analysé des données d'interview de la Current Population Survey (CPS, enquête sur la population active) pour les trois premiers mois de chacune de trois années (1993, 1995, 1996) et ils ont mené une analyse supplémentaire des données de réinterview non rapprochées de la CPS pour à peu près les mêmes périodes de temps. Les données de réinterview représentent une autre stratégie d'estimation de l'erreur de classification de la CPS qui, si on la compare aux estimations de type MLCA, aide à examiner la validité de la stratégie MLCA. Les auteurs abordent cinq dimensions de la validité de la stratégie MLCA : a) diagnostics modélisés, b) validité de l'ajustement du modèle pour trois années de la CPS, c) accord entre les estimations de type modèle et de type réinterview test-retest des probabilités de réponse, d) accord entre les estimations de type modèle et de type réinterview test-retest du manque de cohérence et e) plausibilité des profils d'erreur de classification. De plus, les auteurs examinent la robustesse des estimations MLCA à l'égard d'infractions à l'hypothèse markovienne. La présente analyse ne fournit aucune raison de mettre en doute la validité de la stratégie MLCA. La méthode a donné de bons résultats pour les cinq épreuves de validité.

Mots clés : Enquêtes par panel; erreur non due à l'échantillonnage; chômage; qualité des données.

1. Introduction

La Current Population Survey (CPS, enquête sur la population active) est une enquête sur échantillon menée mensuellement auprès des ménages par le Bureau of the Census des États-Unis en vue de la préparation d'estimations de l'emploi, du chômage et d'autres caractéristiques de la population active des États-Unis. Des estimations nationales tirées de la CPS pour ce qui est de la taille, de la composition et de l'évolution de la composition de la population active sont publiées chaque mois par le Bureau of Labor Statistics in Employment and Earnings des États-Unis. Les estimations de la population active tirées de la CPS sont l'un des indicateurs économiques clés du pays; depuis 1942, l'administration fédérale fait appel aux séries de données de la CPS pour suivre l'évolution mensuelle et annuelle de la population active.

Compte tenu de l'importance des séries de données de la CPS pour la politique gouvernementale, l'exactitude des données a été évaluée à plusieurs reprises. Ainsi, depuis le début des années 1950, le Bureau of the Census a dirigé le programme de réinterview de la CPS afin d'évaluer la qualité des données sur la population active. Dans le cadre de ce programme, on tire un petit sous-échantillon (moins de 5 %) des répondants de la CPS et on pose de nouveau certaines questions de l'interview originale, notamment des questions sur la population active. Jusqu'en 1994, on a soumis un quart environ de l'échantillon à une réinterview non

rapprochée et les trois quarts restants à une réinterview rapprochée. Le volet réinterview rapprochée, utilisé surtout en vue du contrôle de la qualité des interviews, a été éliminé en 1994, à cause de préoccupations au sujet de la qualité des données. Toutefois, la réinterview non rapprochée se poursuit encore aujourd'hui et sert à l'estimation de la stabilité (ou de la cohérence des réponses). On trouvera dans Forsman et Schreiner (1991) une description détaillée du programme de réinterview de la CPS.

Plusieurs documents préparés par des chercheurs à l'extérieur du Bureau of the Census ont analysé les données du programme de réinterview de la CPS afin d'estimer l'erreur de classification (voir Sinclair et Gastwirth 1996, 1998; Biemer et Forsman 1992; Chua et Fuller 1987; Poterba et Summers 1986; Abowd et Zellner 1985). Récemment, Poterba et Summers (1995) ont utilisé des données du programme de réinterview de la CPS afin d'estimer les taux d'erreur de classification de la CPS et d'évaluer l'effet de l'erreur de classification sur les taux de changement de situation vis-à-vis de l'activité. Comme dans le document de 1986, l'analyse plus récente de ces auteurs se fonde sur l'hypothèse que le processus de rapprochement des réinterviews de la CPS fournit des données que l'on peut considérer comme véridiques. Abowd et Zellner (1985) ont adopté une stratégie semblable.

Plusieurs auteurs (par exemple Sinclair et Gastwirth 1996, 1998; Biemer et Forsman 1992; Forsman et Schreiner 1991; Schreiner 1980) ont mis en doute l'hypothèse selon

1. Paul P. Biemer, Research Triangle Institute, Research Triangle Park, NC 27709 et John M. Bushery, Bureau of Transportation Statistics, Washington, DC 20590, États-Unis.

laquelle les réinterviews rapprochées fournissent des valeurs véridiques. Ces auteurs présentent d'importantes indications que les données de réinterview se prêtent à des erreurs de classification appréciables. En réalité, cette prise de conscience a été responsable de l'élimination par le Bureau of the Census du volet réinterview rapproché du programme de réinterview de la CPS en 1994.

Comme substitut de l'hypothèse d'infailibilité, Chua et Fuller (1987) et Fuller et Chua (1985) ont appliqué un type de modèle de structure latente aux données de réinterview rapprochée de la CPS afin d'estimer les probabilités de réponse à la CPS. Par souci d'identifiabilité du modèle, ces auteurs imposent de rigoureuses restrictions aux probabilités de réponse, forçant le biais dû à l'erreur de classification à être nul tant pour l'interview que pour la réinterview. De plus, ils supposent des erreurs de classification indépendantes pour l'interview et la réinterview (on parle d'hypothèse ICE dans la documentation) et d'un mois à l'autre dans l'échantillon. L'hypothèse ICE représente une limite de leur analyse puisque certains documents indiquent que l'hypothèse n'est peut-être pas valable pour la CPS (voir par exemple O'Muircheartaigh 1991, Singh et Rao 1995). Par conséquent, il se peut que les probabilités de réponse estimées à l'aide de la stratégie de Chua et Fuller soient biaisées.

Sinclair et Gastwirth (1996) et Sinclair et Gastwirth (1998) ont appliqué une stratégie de modélisation de classe latente aux données d'interview-réinterview de la CPS en fonction de restrictions de modèle proposées originalement par Hui et Walter (1980) pour des épreuves diagnostiques médicales. À l'aide de données d'interview-réinterview recoupées selon le sexe, Sinclair et Gastwirth ont supposé que les probabilités d'erreur de classification sont égales pour les hommes et les femmes, tandis que les taux d'activité de la population active sont différents pour ces groupes. Puisque les paramètres de modèle absorbent tous les degrés de liberté disponibles pour l'estimation des paramètres, aucun degré de liberté résiduel n'est disponible pour la vérification du manque d'ajustement du modèle. Par conséquent, leur analyse n'indique pas directement si ces hypothèses au sujet du modèle sont valables pour les données de la CPS.

Dans une étude des déterminants du biais dû aux groupes de renouvellement, Shockey (1988) a également appliqué une analyse de structure latente à la CPS. Son analyse a indiqué que le biais dû aux groupes de renouvellement signalé d'abord par Bailar (1975) est peut-être causé par l'erreur de réponse suscitée dans le cadre de l'interview. Shockey n'a pas utilisé de données de réinterview, se fondant plutôt sur des méthodes confirmatoires d'analyse par facteurs pour appuyer ses affirmations. Ses taux d'erreur étaient beaucoup plus élevés que les taux signalés par d'autres auteurs, ce qui indique peut-être un biais dû au modèle. Malheureusement, comme pour Sinclair et Gastwirth, les données de Shockey ne permettent pas une vérification complète des hypothèses du modèle utilisé.

La méthode de l'analyse markovienne de structure latente, une stratégie prometteuse d'estimation de l'erreur de classification des données d'enquête par panel, n'avait pas encore été appliquée à la CPS. Cette méthode tire avantage de la nature répétitive des enquêtes par panel afin d'extraire de l'information sur l'erreur de classification directement des données d'interview. Le modèle MLCA est en réalité une combinaison de deux modèles : un modèle de chaîne markovienne latente qui représente les changements d'un mois à l'autre de la véritable situation vis-à-vis de l'activité, et un modèle d'erreur de classification qui représente les écarts par rapport à la situation véritable et observée vis-à-vis de l'activité.

Puisque la stratégie MLCA tire avantage de la nature répétitive des enquêtes par panel en vue de l'extraction d'informations sur l'erreur de classification directement des données d'interview, elle n'exige aucune mesure infailible externe ni aucune mesure obtenue à l'aide de méthodes de réinterview. À cet égard, la méthode offre certains avantages par rapport aux méthodes traditionnelles du Bureau of the Census et aux méthodes de Chua et Fuller, Abowd et Zellner, Poterba et Summers, et Sinclair et Gastwirth pour ce qui est de l'évaluation de la qualité des données d'enquête. Dans de nombreuses enquêtes par panel, les réinterviews ne sont pas possibles à cause de restrictions budgétaires, de la complexité des travaux sur place et du fardeau de réponse. La stratégie MLCA est peut-être la seule façon d'évaluer l'erreur de mesure de ces enquêtes. Pour ce qui est des enquêtes par panel comme la CPS, pour lesquelles des données de réinterview sont disponibles, les méthodes de réinterview et MLCA offrent une autre stratégie analytique d'évaluation de l'erreur de classification. Ainsi, comme dans la présente analyse, la méthode MLCA peut servir à modéliser et à vérifier les hypothèses traditionnelles d'analyse des réinterviews. De plus, la stratégie MLCA fournit un schéma statistique pour la combinaison des données recueillies au moyen d'un panel et des données de réinterview en vue de l'obtention d'une information encore plus riche au sujet de l'erreur de classification (van de Pol et Langeheine 1997).

Un autre avantage de la stratégie MLCA est la possibilité d'intégrer l'ensemble des données recueillies au moyen d'un panel aux estimations de l'erreur de classification au lieu de se limiter à l'échantillon relativement petit retenu pour la réinterview. Par conséquent, certains aspects de la qualité des données des enquêtes par panel qu'il n'était pas possible d'examiner jusqu'à présent à cause d'un manque de données sont peut-être abordables maintenant.

Le présent exposé décrit nos résultats pour ce qui est de l'utilité de la stratégie de modélisation MLCA pour l'évaluation de l'erreur de classification de la population active dans le cadre de la CPS. Les logiciels servant à ajuster tout un choix de modèles de structure latente de type MLCA ou autre sont accessibles de différentes sources. Le logiciel utilisé dans notre analyse est *ℓ* EM (Vermunt 1997),

et il permet d'ajuster une grande classe de modèles linéaires logarithmiques en présence ou en l'absence de variables latentes. Grâce à sa souplesse et à sa généralité, ce logiciel permet à l'analyste de vérifier une gamme considérable de modèles d'erreur de classification et d'examiner les hypothèses au sujet de la cause et des corrélations de l'erreur de classification.

À la section suivante, nous décrivons le modèle MLCA ainsi que la méthode d'estimation et ses fondements théoriques. À la section 3, il est question de la méthode MLCA appliquée à la CPS, et nous ajustons une série de modèles à la CPS tout en examinant l'ajustement de ces modèles. Dans la présente section, nous préparons également des estimations de l'erreur de classification en fonction du meilleur modèle MLCA. À la section 4, nous vérifions de différentes façons la validité des estimations MLCA, notamment en comparant celles-ci aux estimations d'une nouvelle analyse interview-réinterview. Enfin, à la section 5, nous résumons nos résultats et formulons des recommandations au sujet de l'utilité de la méthode MLCA pour de futures évaluations de l'erreur de classification de la situation vis-à-vis de l'activité.

2. Analyse markovienne de structure latente pour trois périodes de temps

Les modèles markoviens de structure latente ont d'abord été proposés par Wiggins (1973) et perfectionnés par Poulsen (1982). Van de Pol et de Leeuw (1986) ont établi les conditions dans lesquelles le modèle est identifiable, et ils ont décrit d'autres conditions de l'estimabilité des paramètres de modèle. Dans la présente section, nous élaborons le modèle MLCA dans le contexte de la CPS, et nous en suggérons d'autres applications de même que la généralisation.

Nous divisons la population cible de la CPS en groupes L (âge, race, sexe, par exemple). Soit G_i , une variable qui désigne la composition du groupe. Ainsi, $G_i = 1$ si le $i^{\text{ème}}$ membre de la population se trouve dans le groupe 1, $G_i = 2$ pour le groupe 2 et ainsi de suite. Soit X_{gi} , Y_{gi} , et Z_{gi} , les véritables classifications de la situation vis-à-vis de l'activité pour la $i^{\text{ème}}$ personne du groupe $G = g$ (pour $g = 1, \dots, L$ et $i = 1, \dots, n_g$) où X_{gi} se définit comme

$$X_{gi} = \begin{cases} 1 & \text{si la personne } (g, i) \text{ est employée} \\ & \text{au cours de la période 1} \\ 2 & \text{si la personne } (g, i) \text{ est sans emploi} \\ & \text{au cours de la période 1} \\ 3 & \text{si la personne } (g, i) \text{ est inactive} \\ & \text{au cours de la période 1} \end{cases}$$

avec des définitions semblables pour Y_{gi} et Z_{gi} au cours des périodes 2 et 3 respectivement. Soit $\pi_{x,y,z|g}$, $\Pr(X = x, Y = y, Z = z | G = g)$, $\pi_{y|g,x}$, $\Pr(Y = y | X = x, G = g)$

et $\pi_{z|g,y,x}$, $\Pr(Z = z, Y = y, X = x | G = g)$. La probabilité qu'un membre du groupe g se trouve dans une situation d'activité x au cours de la période 1, y au cours de la période 2 et z au cours de la période 3 est

$$\pi_{x,y,z|g} = \pi_{x|g} \pi_{y|g,x} \pi_{z|g,x,y}. \quad (1)$$

Enfin, en vertu de l'hypothèse markovienne d'ordre un, une condition nécessaire de l'identifiabilité du modèle (voir van de Pol et de Leeuw 1986), nous supposons

$$\pi_{z|g,x,y} = \pi_{z|g,y} \quad (2)$$

c'est-à-dire qu'au cours de la période 3, la véritable situation d'une personne ne dépend pas de la situation au cours de la période 1, une fois la situation au cours de la période 2 connue. Une autre interprétation est que la situation courante, compte tenu de la situation au cours de la période précédente, ne dépend pas du changement de la période précédente.

On peut concevoir un certain nombre de scénarios dans lesquels l'hypothèse markovienne n'est peut-être pas valable pour la situation d'activité mensuelle. L'hypothèse serait enfreinte, par exemple, si des personnes sans emploi au cours de la période 2 avaient plus de chances d'être sans emploi au cours de la période 3, étant donné qu'elles étaient également sans emploi au cours de la période 1. Le groupe de personnes sans emploi au cours de la période 2 et de la période 1 englobe probablement une plus forte proportion de chômeurs chroniques que le groupe de personnes sans emploi au cours de la période 2 mais non au cours de la période 1. Ce groupe (sans emploi au cours de la période 2 mais non pas de la période 1) compte probablement une plus forte proportion de chômeurs occasionnels en train de changer d'emploi.

Toutefois, la validité de cette hypothèse ne peut pas être examinée convenablement à l'aide des données observées parce que les données comportent une distorsion dont l'ampleur est inconnue à cause de la présence d'erreurs de classification. Il existe au moins deux méthodes d'évaluation de la validité de l'hypothèse markovienne pour des données recueillies au moyen d'un panel. Van de Pol et de Leeuw (1986) ont suggéré une méthode fondée sur quatre cycles de données recueillies au moyen d'un panel qui substitue une restriction markovienne d'ordre 2 à la restriction d'ordre un en (2). Une autre méthode, suggérée par van de Pol et Langeheine (1997), fait appel à une combinaison de données de type panel sur la situation d'activité et de données de réinterview pour chaque période de temps. Ni l'une ni l'autre de ces méthodes n'a été utilisée dans le présent exposé pour vérifier directement l'hypothèse de la MLCA. Nous avons plutôt évalué la validité globale des estimations MLCA à l'aide des méthodes décrites à la section 3.2 ci-dessous. À la section 3.6, nous présentons des résultats d'une étude de simulation afin d'illustrer la robustesse des estimations MLCA de l'erreur de classification vis-à-vis des infractions à l'hypothèse markovienne.

Nous considérons maintenant les classifications observées de la situation vis-à-vis de l'activité relativement à la CPS notées A_{gi} , B_{gi} et C_{gi} pour les périodes 1, 2 et 3, respectivement, où

$$A_{gi} = \begin{cases} 1 & \text{si la personne } (g, i) \text{ est classée comme} \\ & \text{employée au cours de la période 1} \\ 2 & \text{si la personne } (g, i) \text{ est classée comme} \\ & \text{sans emploi au cours de la période 1} \\ 3 & \text{si la personne } (g, i) \text{ est classée comme} \\ & \text{inactive au cours de la période 1} \end{cases}$$

avec des définitions semblables pour les indicateurs de réponse B_{gi} et C_{gi} pour les périodes 2 et 3, respectivement. À l'aide d'une extension de la notation établie ci-dessus, nous désignons les probabilités de réponse pour chacune de ces classifications par $\pi_{a|g,x} = \Pr(A = a | G = g, X = x)$, avec des définitions semblables pour $\pi_{b|g,y}$ et $\pi_{c|g,z}$. Ainsi, $\pi_{a=1|g,x=2}$ est la probabilité qu'une personne dans le groupe g soit classée par la CPS comme employée ($A = 1$) lorsqu'elle est en réalité sans emploi ($X = 2$). De même, $\pi_{a=2|g,x=2}$ est la probabilité qu'une personne dans le groupe g soit classée correctement comme sans emploi.

Enfin, nous supposons

$$\pi_{a,b,c|g,x,y,z} = \pi_{a|g,x} \pi_{b|g,y} \pi_{c|g,z} \quad (3)$$

ou que l'erreur de classification de la situation observée vis-à-vis de l'activité est indépendante de l'un à l'autre des trois mois. Cette hypothèse, appelée l'hypothèse de l'indépendance locale, a été examinée relativement à la CPS par Meyers (1988) dans son étude de la stratégie d'estimation utilisée par Abowd et Zellner (1985). Meyers a conclu que l'hypothèse semble représenter une approximation raisonnable. Singh et Rao (1995), qui ont étudié la robustesse de l'hypothèse en fonction d'un certain nombre de scénarios de la population active, ont tiré une conclusion semblable. Van de Pol et Langeheine (1997) ont modélisé la distribution combinée des données recueillies au moyen d'un panel et des données de type réinterview à l'aide de modèles de structure latente afin de vérifier l'indépendance locale de divers types de changements de situation vis-à-vis de l'activité. Ils ont trouvé certaines indications que les personnes qui changent de situation vis-à-vis de l'activité manifestent une fiabilité inférieure à celle des personnes qui ne changent pas de situation, l'effet étant toutefois assez faible. Par conséquent, nous allons également supposer (3) sans tenter d'en approfondir davantage la validité dans le présent exposé.

Les classifications de la CPS pour ce qui est de la situation vis-à-vis de l'activité au cours de chaque mois du premier trimestre de l'année sont les variables résultantes de notre analyse. Soit A, B , et C , les classifications observées, et X, Y et Z , les véritables classifications (non observées) pour janvier, février et mars, respectivement. Soit G , une variable de groupement (ou stratification) quelconque à définir ultérieurement au cours de l'analyse. Compte tenu de ces hypothèses, nous pouvons formuler la

probabilité qu'un membre de l'échantillon de la CPS soit classé dans la cellule (g, a, b, c) du tableau $GABC$ comme suit :

$$\pi_{g,a,b,c} = \sum_{x,y,z} \pi_g \pi_{x|g} \pi_{a|g,x} \pi_{y|x,g} \pi_{b|y,g} \pi_{z|g,y} \pi_{c|g,z}. \quad (4)$$

Les extensions à plus d'une variable de groupement se font directement.

Dans le cadre d'un échantillonnage multinomial, la fonction de vraisemblance pour le tableau $GABC$ est

$$\Pr(GABC) = k \prod_{g,a,b,c} \pi_{g,a,b,c}^{n_{gabc}} \quad (5)$$

où k est la constante multinomiale et \prod désigne le produit des termes pour les indices g, a, b et c . En vertu des hypothèses énoncées antérieurement, les paramètres de modèle peuvent être estimés à l'aide de méthodes d'estimation du maximum de vraisemblance. Van de Pol et de Leeuw (1986) ont fourni la formule permettant d'appliquer l'algorithme E-M afin d'estimer les paramètres de ce modèle, et ils ont décrit les conditions de leur estimabilité. Le logiciel ℓ EM, appliqué aux ensembles de données de la CPS à la section suivante, assure la mise en œuvre de ces méthodes.

3. Application à la CPS

3.1 Notation

Une partie de notre évaluation de la stratégie MLCA est une comparaison des estimations MLCA de l'erreur de classification aux estimations tirées de l'analyse des données d'interview-réinterview. Nous utilisons la notation décrite à la section précédente; soit A et A' , la classification de la situation vis-à-vis de l'activité pour l'original et pour la réinterview, respectivement; nous définissons $\pi_a = \Pr(A = a)$ et $\pi_{a'} = \Pr(A' = a')$. Soit AA' , le tableau de tri croisé des interviews-réinterviews observées $K \times K$; soit $\pi^{AA'|X}$, la matrice $K \times K$ des probabilités de cellule, $\Pr(A = a, A' = a' | X = x)$. Si nous supposons que $\pi_{aa'} = \Pr(A = a, A' = a' | X = x) = \pi_{a|x}^2$, appelée dans la documentation l'hypothèse des mesures parallèles (Bohrnstedt 1983), nous avons

$$\pi^{AA'|X} = \pi^{A|X} (\pi^{A|X})^T \quad (6)$$

où $(\pi^{A|X})^T$ désigne la transposée d'un vecteur de probabilités conditionnelles, $\pi^{A|X}$.

Soit π^X , le vecteur K des véritables probabilités de classification. Nous avons alors

$$\pi^{AA'} = \pi^{AA'|X} \pi^X \quad (7)$$

c'est-à-dire que la probabilité du tableau de classification des interviews-réinterviews observées, $\pi^{AA'}$, est égale au produit de la matrice des probabilités de réponse conditionnelles, $\pi^{AA'|X}$, et du vecteur des véritables probabilités de classification, π^X .

Comme nous l'avons expliqué à la section précédente, la MLCA des données longitudinales de la CPS fournira des estimations du maximum de vraisemblance de $\pi^{A|X}$ et de π^X , permettant l'estimation de $\pi^{AA'}$ à l'aide de (6) et de (7). Nous pouvons estimer la fiabilité test-retest, R , pour toute situation vis-à-vis de l'activité en appliquant les méthodes d'estimation habituelles (voir par exemple Bohrnstedt 1983) à cette estimation de $\pi^{AA'}$. Pour notre analyse, nous calculons l'indice d'incohérence, $I = 1 - R$, qui représente la mesure traditionnelle de la fiabilité des données sur la population active de la CPS (voir U.S. Bureau of the Census 1985). Soit I_a , l'indice d'incohérence pour la catégorie $A = a$. Un estimateur de I_a est

$$\frac{gdr}{2\hat{\pi}_a(1 - \hat{\pi}_a)} \quad (8)$$

où gdr représente le taux de différence brut défini par

$$gdr_a = 2 \sum_{a \neq a'} \hat{\pi}_{a, a'} \quad (9)$$

et où $\hat{\pi}_a$ et $\hat{\pi}_{a, a'}$ désignent des estimations de structure latente de π_a et de $\pi_{a, a'}$, respectivement.

Le Bureau of the Census des États-Unis (1985, 88-91) a fourni les formules des erreurs-types de même qu'une mesure globale de l'incohérence pour toutes les catégories K combinées, appelée l'indice global d'incohérence, I_{AG} . L'indice global est une mesure de la non-fiabilité au niveau des questions égale à $1 - \kappa$ (Hess, Singer et Bushery 2000) où κ est la mesure de la fiabilité kappa de Cohen (Cohen 1960) et une moyenne pondérée des indices au niveau des catégories.

Enfin, pour une estimation de π^X donnée, nous pouvons estimer le vecteur K des biais de mesure, notés β_A , associés aux catégories K de A à l'aide de l'identité

$$\beta_A = \pi^A - \pi^X. \quad (10)$$

3.2 Évaluation de la validité de la méthode MLCA

Le présent exposé a comme principal objectif d'évaluer la validité de la stratégie MLCA. Les études précédentes de la mesure de l'erreur de classification de la CPS n'ont pas complètement abordé la validité des stratégies d'estimation utilisées (Meyers 1988). Nous comptons déterminer si la stratégie MLCA est informative et utile pour l'étude de l'erreur de classification de la CPS. Plus précisément, nous espérons déterminer si les estimations modélisées des probabilités d'erreur, $\pi^{A|X}$, reflètent les véritables niveaux d'erreur des classifications de la situation vis-à-vis de l'activité que l'on trouve dans la CPS. Malheureusement, pour des raisons mentionnées antérieurement, il n'existe aucun étalon généralement accepté pour l'évaluation de l'exactitude de la CPS (voir par exemple Sinclair et Gastwirth 1996, 1998, Biemer et Forsman 1992, ainsi que Schreiner 1980). Par conséquent, il n'est pas possible d'estimer le biais des estimations MLCA.

Dans la suite du texte, nous examinons la validité des estimations MLCA de l'erreur de classification de la CPS à l'aide de cinq critères :

1. **Diagnostics modélisés.** Une condition nécessaire de la validité d'un modèle est que celui-ci soit plausible (que les hypothèses soient raisonnables et concordent avec la réalité) et qu'il ajuste les données de façon adéquate. Nous utilisons le critère traditionnel chi-carré de la validité de l'ajustement ainsi que d'autres mesures diagnostiques de l'ajustement du modèle afin d'évaluer le bien-fondé de celui-ci et la mesure dans laquelle les données concordent avec lui.
2. **Validité de l'ajustement du modèle d'une année à l'autre de la CPS.** Une méthode répandue de validation d'un modèle consiste à évaluer l'ajustement du modèle pour des données qui sont indépendantes des données utilisées pour la modélisation (voir par exemple Kleinbaum, Kupper et Muller 1988, 330). Cette méthode permet d'éviter un surparamétrage du modèle et une sélection de modèle guidée par les données (plutôt que par la théorie). Dans la présente étude, l'ajustement du même modèle aux données de chaque année séparément est un aspect de cette technique de vérification indépendante du modèle. L'accord du modèle d'une année à l'autre aurait tendance à confirmer la validité de sa structure. Cette méthode comporte une difficulté pour la présente application. Après 1993, le questionnaire imprimé de la CPS a été remanié en fonction de l'IPAO (interview sur place assistée par ordinateur), de sorte que l'ampleur des erreurs de réponse a pu changer après 1993. Toutefois, si les principales sources d'erreur de réponse de la CPS n'ont pas changé sous l'effet du remaniement, une structure modélisée qui décrit convenablement l'erreur pour 1993 devrait également décrire l'erreur pour 1995 et 1996.
3. **Accord entre les estimations MLCA et les estimations test-retest de Hui-Walter pour les probabilités de réponse.** La méthode de Hui-Walter (H-W) (Hui et Walter 1980) d'estimation des probabilités de réponse de la CPS fait appel à des données de réinterview non rapprochée (Sinclair et Gastwirth 1996; 1998). Même si les méthodes MLCA et H-W font toutes deux appel à des modèles de structure latente, les hypothèses de modélisation sont très différentes. Par exemple, la méthode H-W n'exige pas d'hypothèse markovienne pour l'identifiabilité du modèle. De plus, dans la présente recherche, les entrées de données pour la méthode H-W sont indépendantes de celles utilisées pour la méthode MLCA. Le bon accord entre les deux ensembles d'estimations confirme la validité des deux méthodes, tandis qu'un faible accord indique qu'une des stratégies au moins n'est pas valable. Une forte concordance entre les estimations MLCA et H-W fournit également une certaine assurance que les estimations MLCA des probabilités de réponse sont relativement robustes vis-à-vis d'infractures éventuelles à l'hypothèse markovienne.

4. **Accord entre les estimations modélisées et test-retest de l'indice d'incohérence.** Ce critère est semblable au troisième critère car il compare des estimations tirées de la MLCA à des estimations fondées sur des données de réinterview non rapprochée. Toutefois, cette analyse ne fait pas appel à la validité de la méthode d'estimation de Hui-Walter pour l'évaluation de la validité des estimations MLCA. Nous utilisons plutôt les estimations MLCA de l'erreur de classification pour obtenir des estimations de l'indice d'incohérence à l'aide de (7) à (9). Nous comparons ces estimations de la fiabilité directement aux estimations de la fiabilité tirées du programme de réinterview de la CPS, obtenues de données de réinterviews non rapprochées. Un bon accord entre les estimations de réinterview et de type MLCA confirme la validité des deux méthodes, tandis qu'un faible accord indique qu'une des stratégies au moins n'est pas valable.
5. **Plausibilité des profils d'erreur de classification.** Enfin, la plausibilité (ou la validité apparente) des estimations des probabilités de réponse peut également servir d'épreuve de la validité. Par exemple, il semble peu plausible que des réponses par procuration à des questions sur la population active soient plus exactes que des réponses sans procuration. D'autres profils d'erreur de classification peuvent être examinés et évalués en fonction de la plausibilité. Pour autant que les estimations modélisées semblent plausibles, leur validité apparente est confirmée.

À la section suivante, nous discutons de nos résultats de modélisation MLCA dans le contexte de ces critères de validité. Nous commençons par décrire les ensembles de données de la CPS et les résultats du processus de sélection du modèle.

3.3 Ensembles de données de la CPS

En 1994, dans le cadre de la mise en œuvre de l'IPAO (interview sur place assistée par ordinateur), la CPS a été remaniée de fond en comble, avec restructuration des questions servant à déterminer la situation vis-à-vis de l'activité. Rothgeb (1994) a décrit le remaniement de la CPS. Comme conséquence de ces améliorations, nous nous attendons à observer une différence (plus précisément une réduction) de l'erreur de classification de la CPS après 1994 comparativement à 1993. Bien qu'il ne s'agisse pas d'un objectif principal de la présente recherche, nous avons comparé l'erreur de la CPS avant et après le remaniement. Nous avons vérifié la stratégie MLCA pour trois années de la CPS : 1993, 1995 et 1996, puisque des données de réinterview non rapprochée de la CPS étaient facilement accessibles pour cette période de temps.

Les ménages de la CPS sont interviewés au cours de quatre mois consécutifs, ne font pas partie de l'enquête pendant huit mois, puis sont de nouveau interviewés au cours d'une deuxième série de quatre mois consécutifs. La

MLCA exige au moins trois interviews consécutives pour l'identifiabilité des paramètres de modèle. Nous avons un choix d'ensembles de données comprenant toutes les personnes interviewées au cours de trois ou quatre mois consécutifs de la CPS. Puisque le recours à quatre mois de données réduirait la taille de l'échantillon pour l'analyse de moitié, nous avons choisi de faire porter l'analyse sur trois mois consécutifs (janvier, février et mars) pour les trois années de données. Les cas de non-réponse et les cas de changement du ménage tout entier pour un ou plusieurs des trois mois ont été exclus de l'analyse.

Le modèle MLCA le plus simple précise que les probabilités de réponse $\pi_{a|x}$, $\pi_{b|y}$, et $\pi_{c|z}$, ainsi que les probabilités de changement, $\pi_{y|x}$, $\pi_{z|y}$ sont identiques pour toutes les personnes de la population cible (c'est ce qu'on appelle l'homogénéité). Toutefois, notre analyse préliminaire (Biemer, Bushery et Flanagan 1997) a indiqué que les probabilités de réponse et de changement ne sont pas homogènes. Afin de rendre compte de cette hétérogénéité, nous avons examiné un certain nombre de covariables et de variables de stratification à inclure dans les modèles, y compris : le sexe, la scolarité, le mode d'interview, le type de réponse (avec ou sans procuration) et la race. Parmi celles-ci, c'est la variable tirée de l'indicateur de réponse avec ou sans procuration de la CPS qui a le mieux tenu compte de l'hétérogénéité de la population. Cette variable, notée P , se définit comme suit :

$$P = \begin{cases} 1 & \text{si les trois interviews sont toutes menées sans procuration (SELF)} \\ 2 & \text{si deux des trois interviews sont menées sans procuration (MOSTLY SELF)} \\ 3 & \text{si deux des trois interviews sont toutes menées par procuration (MOSTLY PROXY)} \\ 4 & \text{si les trois interviews sont toutes menées par procuration (PROXY)} \end{cases}$$

À noter que nous utilisons maintenant P pour représenter la variable de groupement, au lieu de G , que nous avons utilisé à la section 2. Compte tenu de recherches antérieures (par exemple O'Muircheartaigh 1991), nous nous attendons à ce que le groupe Self ($P = 1$) comporte moins d'erreurs de classification que le groupe Proxy ($P = 4$). Nous vérifions cette hypothèse dans le cadre du critère de plausibilité des estimations (critère 4 ci-dessus).

La taille des échantillons pour les trois ensembles de données utilisés dans notre analyse est la suivante :

1993 :	45 291 personnes
1995 :	49 347 personnes
1996 :	41 751 personnes

Pour 1993, un tiers environ de l'échantillon se trouve dans le groupe Self, un quart environ dans le groupe Proxy, les autres membres de l'échantillon étant répartis à peu près également entre les groupes Mostly Self et Mostly Proxy. Pour 1995 et 1996, le nombre de membres de l'échantillon

dans le groupe Proxy est légèrement plus élevé (un tiers au lieu d'un quart).

3.4 Ajustement des modèles MLCA

Pour l'ajustement d'un modèle MLCA comportant une seule variable de groupement, P , l'ensemble de données d'entrée était un tableau $4 \times 3 \times 3 \times 3$ à dénombrement de cellules défini par le tri croisé de $P \times A \times B \times C$, où A , B et C représentent la classification de la situation vis-à-vis de l'activité pour janvier, février et mars, respectivement.

Le logiciel ℓ EM et d'autres progiciels d'ajustement de modèle MLCA supposent un échantillonnage aléatoire simple, de sorte que le plan de sondage complexe de la CPS ne peut pas être modélisé exactement. Il est possible de rendre compte de la structure d'échantillonnage à probabilités inégales de la CPS à l'aide de dénombrements de cellules pondérés et remis à l'échelle plutôt que de totaux de cellules bruts (Clogg et Eliason 1985). Toutefois, le recours à des données non pondérées pour la stratégie MLCA offre deux avantages importants. Premièrement, nous pouvons comparer les estimations MLCA aux estimations tirées des études citées antérieurement de l'erreur de classification de la CPS, toutes fondées sur des données non pondérées. Deuxièmement, les données de réinterview de la CPS servant à évaluer les critères 3 et 4 ne sont pas pondérées et aucun poids n'est disponible. Par conséquent, une partie au moins de l'analyse exige des données non pondérées; l'utilisation de données pondérées pour les autres critères risquerait de produire de fausses incohérences dans les résultats.

Afin d'examiner la validité d'inférences pour la population totale fondées sur une analyse non pondérée, nous avons estimé les erreurs de classification pour des données à la fois pondérées et non pondérées et nous avons observé que les estimations d'erreur de classification exprimées sous forme de proportions étaient pratiquement identiques, la différence n'étant manifeste qu'à la troisième décimale. Les résultats présentés ci-dessous pour des dénombrements de cellules non pondérés sont donc appropriés pour une inférence au-delà de l'échantillon de la CPS à la population totale.

Un autre aspect du recours à une analyse non pondérée est l'estimation des erreurs-types. Puisqu'elles sont calculées en fonction d'hypothèses d'échantillonnage aléatoire simple, les estimations ℓ EM de l'erreur-type risquent d'être sous-estimées par manque de connaissance de l'effet de groupement dans l'échantillon de la CPS. Pour tenir compte approximativement de ce phénomène, nous pouvons multiplier les variances ℓ EM par un effet de plan de sondage calculé à partir des estimations de la population active de la CPS. Le Bureau of the Census des États-Unis (2000, 14-9) a indiqué que les effets de plan de sondage pour les estimations de la population active de la CPS ne dépassent pas 1,3; le fait de multiplier les erreurs-types ℓ EM par $(1,3)^{1/2}$ devrait entraîner une inflation suffisante des erreurs-types pour tenir compte du groupement. Une

stratégie équivalente consiste à utiliser un niveau de signification de 3 % plutôt que de 5 % au moment de déclarer la différence entre deux estimations. Cette dernière stratégie sera considérée comme appropriée dans l'analyse à venir. Nous croyons que cela représente un test prudent puisque l'effet du plan de sondage de la CPS reflète l'augmentation de la variance attribuable tant au groupement de l'échantillon qu'à la pondération inégale, tandis que seuls des effets de groupement figurent dans nos estimations non pondérées.

Le tableau 1 indique les résultats de l'ajustement d'une série de modèles MLCA de plus en plus complexes pour chacun des trois ensembles de données. Le modèle de base est le modèle MLCA le plus simple; il précise que les probabilités de changement et les probabilités de réponse sont homogènes (c'est-à-dire qu'elles ne diffèrent pas selon le groupe P) et stationnaires (c'est-à-dire qu'elles sont identiques pour les trois mois). Ce modèle peut s'écrire comme suit

$$\pi_{p,a,b,c} = \sum_{x,y,z} \pi_p \pi_{x|g} \pi_{a|x}^3 \pi_{y|x}^2 \quad (11)$$

que l'on obtient de (4) en imposant les contraintes

$$\pi_{z|yp} = \pi_{y|xp} = \pi_{y|x} \quad (12)$$

et

$$\pi_{a|xp} = \pi_{b|yp} = \pi_{c|zp} = \pi_{a|x} \quad (13)$$

pour tous les p .

Pour le modèle 1 nous réduisons la contrainte (12) à

$$\pi_{z|yp} = \pi_{y|xp} \text{ pour } p = 1, \dots, 4 \quad (14)$$

ce qui permet aux changements survenus de janvier à février et de février à mars de varier selon le groupe Self/Proxy, P . Pour le modèle 2, nous réduisons la contrainte (12) davantage à

$$\pi_{z|xp} = \pi_{y|x} \text{ et } \pi_{z|yp} = \pi_{z|y} \quad (15)$$

pour tous les p . Le modèle 3 réduit les contraintes tant d'homogénéité que de stationnarité pour les probabilités de changement de sorte que $\pi_{y|xp} \neq \pi_{z|yp}$. Ce modèle permet donc aux probabilités de changement de varier selon le groupe et selon le mois. Toutefois, les probabilités de réponse doivent toujours être égales d'un groupe à l'autre et d'un mois à l'autre.

Le modèle 4 est le plus général des modèles identifiables que nous avons considérés. Le modèle 4 permet aux probabilités de changement de janvier à février et de février à mars de varier indépendamment de l'un à l'autre des quatre groupes Proxy/Self. Ce modèle précise également que les probabilités de réponse sont les mêmes pour janvier, février et mars, mais qu'elles peuvent varier de l'un à l'autre des quatre groupes Proxy/Self. Nous avons obtenu ce modèle à partir du modèle 3 en réduisant les contraintes exigeant des

$$\pi_{a|xp} = \pi_{b|yp} = \pi_{c|zp} \quad (16)$$

Tableau 1
Diagnostics modélisés pour d'autres modèles MLCA selon l'année

Données de 1993	<i>df</i>	npar ¹	<i>L</i> ²	valeur <i>p</i>	BIC	<i>d</i>
Modèle de base : Changements homogènes et stationnaires et probabilités de réponse	90	17	645	0,000	-320	0,048
Modèle 1 : Changements non homogènes	84	23	632	0,000	-269	0,047
Modèle 2 : Changements non stationnaires	66	41	99	0,006	-609	0,007
Modèle 3 : Changements non homogènes et non stationnaires	42	65	64	0,016	-386	0,005
Modèle 4 : Changements non homogènes et non stationnaires et probabilités de réponse non homogènes	24	83	23	0,501	-234	0,002
Données de 1995	<i>df</i>	npar ¹	<i>L</i> ²	valeur <i>p</i>	BIC	<i>d</i>
Modèle de base : Changements homogènes et stationnaires et probabilités de réponse	90	17	697	0,000	-275	0,044
Modèle 1 : Changements non homogènes	84	23	668	0,000	-240	0,043
Modèle 2 : Changements non stationnaires	66	41	146	0,000	-567	0,008
Modèle 3 : Changements non homogènes et non stationnaires	42	65	82	0,000	-372	0,005
Modèle 4 : Changements non homogènes et non stationnaires et probabilités de réponse non homogènes	24	83	25	0,410	-234	0,002
Données de 1996	<i>df</i>	npar ¹	<i>L</i> ²	valeur <i>p</i>	BIC	<i>d</i>
Modèle de base : Changements homogènes et stationnaires et probabilités de réponse	90	17	632	0,000	-325	0,045
Modèle 1 : Changements non homogènes	84	23	585	0,000	-308	0,044
Modèle 2 : Changements non stationnaires	66	41	159	0,000	-543	0,010
Modèle 3 : Changements non homogènes et non stationnaires	42	65	82,6	0,000	-364	0,005
Modèle 4 : Changements non homogènes et non stationnaires et probabilités de réponse non homogènes	24	83	39,3	0,026	-216	0,003

¹ À noter que « npar » représente le nombre de paramètres du modèle.

pour tous les p . En vertu de ces contraintes, (4) peut s'écrire comme suit

$$\pi_{p,a,b,c} = \sum_{x,y,z} \pi_p \pi_{x|p} \pi_{y|xp} \pi_{z|py} (\pi_{a|p,x})^3.$$

Dans le tableau 1, nous indiquons la statistique d'ajustement de base pour l'ensemble des cinq modèles et des trois années. La colonne 4 du tableau indique la statistique chi-carré habituelle du rapport des vraisemblances L^2 (voir Agresti 1990, 48), tandis que la colonne 5 indique la valeur p correspondante. Une valeur p égale ou supérieure à 0,05 est le critère habituel d'un ajustement de modèle adéquat. Toutefois, à cause de la taille élevée des échantillons de notre analyse, le fait d'exiger une valeur p aussi grande pourrait entraîner un surajustement du modèle. Nous considérons une valeur p aussi petite que 0,01 comme étant acceptable. La mesure BIC du tableau se définit comme suit :

$$\text{BIC} = L^2 - (\log N) df$$

où N est la taille de l'échantillon total et df représente les degrés de liberté du modèle. Le BIC résume essentiellement le compromis entre l'ajustement du modèle (L^2) et la parcimonie du modèle (df). Puisque de petites valeurs du BIC sont favorables, nous considérons le modèle comportant le plus faible BIC comme le meilleur relativement à la validité de l'ajustement et à la parcimonie. Liu et Dayton (1997) ont discuté de cette stratégie pour des modèles de structure latente.

Enfin, l'indice de dissimilitude (d) est la proportion des observations qui auraient à changer de cellule pour que le modèle soit parfaitement ajusté. En général, on considère des modèles comportant $d \leq 0,05$ (erreur de modèle de 5 %) comme étant bien ajustés aux données (Vermunt 1997).

Pour chaque année de données, le modèle 4 est le seul à fournir un ajustement acceptable lorsque le critère des valeurs p est considéré. Le modèle 4 est également plausible du point de vue de la théorie des réponses puisqu'il postule que l'erreur de classification varie selon le groupe Self/Proxy. Cela, comme nous l'avons dit, correspond à la documentation publiée sur les méthodes d'enquête (voir par exemple O'Muircheartaigh 1991 et Moore 1988). L'indice de dissimilitude d , pour le modèle est de 0,3 %, ce qui indique un ajustement de modèle très valable. Nous utilisons donc le modèle 4 pour produire les estimations de l'erreur de classification de la population active.

3.5 Estimation de l'erreur de classification

Le tableau 2 montre les estimations des probabilités de réponse tirées du modèle 4 pour les catégories employé, sans emploi et inactif. Pour les personnes réellement employées et celles réellement inactives, la probabilité d'une réponse correcte est assez élevée : au moins 98 % pour les employés et 97 % pour les inactifs. Toutefois, pour les personnes réellement sans emploi, la probabilité d'une réponse correcte varie d'une année à l'autre et d'un groupe à l'autre entre 68 % environ et 86 % environ. Comme prévu,

le groupe Self comporte la plus forte probabilité d'une réponse correcte (statistiquement significative pour 1993 et 1996). Le groupe Mostly Self manifeste également une tendance à une plus forte probabilité de réponse correcte que le groupe Mostly Proxy, mais la différence n'est pas statistiquement significative.

Un résultat surprenant du tableau 2 est la direction que prend la différence d'exactitude de déclaration pour les personnes réellement sans emploi d'une des trois années à l'autre. On se souviendra que le questionnaire d'interview de la CPS a été remanié en 1994 afin d'améliorer l'exactitude de déclaration. Toutefois, ces résultats indiquent que l'exactitude de déclaration est plus élevée pour 1993 (année antérieure au remaniement majeur) que pour 1995 et 1996 (les années suivant le remaniement). En 1993, la probabilité d'une réponse correcte était de 81,8 %, comparativement à 76,1 % et à 74,4 % pour 1995 et 1996, respectivement (valeur $p < 0,001$). Cet effet est peut-être une conséquence du remaniement, ou reflète peut-être des changements réels de la population, ou les deux à la fois. Des recherches se poursuivent actuellement au Bureau of the Census pour comprendre les causes qui sous-tendent ces résultats.

Le tableau indique que la classification erronée des personnes sans emploi comme des personnes inactives est

un problème plus grave que la classification erronée de ces personnes comme des personnes employées. En moyenne, pour les trois années, deux tiers environ de l'erreur de classification des sans-emploi est une classification erronée de ces personnes comme inactives, mais le taux est élevé pour les deux types d'erreur.

Nous comparons ensuite nos estimations des probabilités de classification de la CPS à des estimations semblables tirées de la documentation. Au tableau 3, les estimations MLCA pour chacune des trois années sont comparées aux estimations de Chua et Fuller (1987), de Poterba et Summers (1995) et du programme de réinterviews non rapprochées de la CPS. Encore une fois, les trois derniers ensembles d'estimations se fondent sur des données de réinterview, tandis que les estimations MLCA sont préparées directement à partir des données d'interview de la CPS. En général, la valeur relative des estimations MLCA d'une catégorie d'activité à l'autre concorde avec les estimations antérieures. La plus grande différence est observée pour la population des personnes réellement sans emploi. Pour ce groupe, plus élevées que les estimations MLCA correspondantes pour 1993, la période de temps qui correspond le plus étroitement aux estimations comparatives.

Tableau 2
Probabilités de classification estimatives de la situation vis-à-vis de l'activité selon le groupe et l'année
(Les erreurs-types sont indiquées entre parenthèses)

Observé	Groupe	Réellement employé			Réellement sans emploi			Réellement inactif		
		1993	1995	1996	1993	1995	1996	1993	1995	1996
Employé	Sans procuration	98,90 (0,12)	98,98 (0,12)	99,03 (0,13)	4,24 (0,96)	5,66 (1,28)	5,78 (1,52)	1,05 (0,13)	0,67 (0,10)	0,92 (0,13)
	Sans procuration surtout	98,91 (0,18)	98,61 (0,21)	98,86 (0,22)	8,11 (1,89)	7,66 (2,50)	12,20 (2,96)	1,73 (0,27)	1,59 (0,30)	0,93 (0,24)
	Par procuration surtout	98,62 (0,18)	98,56 (0,22)	98,76 (0,20)	11,19 (1,96)	12,86 (2,97)	10,18 (2,73)	1,68 (0,27)	0,95 (0,25)	1,50 (0,30)
	Par procuration	98,66 (0,16)	98,63 (0,13)	98,66 (0,15)	6,60 (1,43)	7,77 (1,44)	8,30 (1,62)	1,56 (0,20)	1,50 (0,17)	1,30 (0,18)
	Total	98,77 (0,11)	98,73 (0,11)	98,83 (0,11)	7,06 (0,70)	7,86 (0,90)	8,57 (1,00)	1,41 (0,11)	1,11 (0,11)	1,13 (0,11)
Sans emploi	Sans procuration	0,33 (0,08)	0,39 (0,08)	0,41 (0,09)	85,92 (1,62)	80,52 (2,04)	79,66 (2,44)	0,48 (0,11)	0,35 (0,10)	0,40 (0,12)
	Sans procuration surtout	0,28 (0,12)	0,50 (0,13)	0,28 (0,15)	82,68 (2,48)	71,71 (3,40)	73,60 (3,71)	1,03 (0,23)	0,91 (0,21)	0,79 (0,28)
	Par procuration surtout	0,37 (0,11)	0,54 (0,15)	0,37 (0,13)	76,54 (2,55)	68,09 (3,37)	72,74 (3,70)	0,81 (0,22)	0,90 (0,28)	1,53 (0,26)
	Par procuration	0,34 (0,10)	0,55 (0,09)	0,39 (0,09)	80,09 (2,19)	77,12 (2,18)	71,63 (2,49)	0,95 (0,19)	0,94 (0,17)	1,18 (0,20)
	Total	0,34 (0,11)	0,49 (0,11)	0,37 (0,11)	81,81 (0,90)	76,09 (1,21)	74,42 (1,21)	0,75 (0,11)	0,69 (0,11)	0,87 (0,11)
Inactif	Sans procuration	0,77 (0,10)	0,63 (0,08)	0,55 (0,09)	9,84 (1,39)	13,82 (1,76)	14,56 (2,11)	98,47 (0,17)	98,98 (0,14)	98,68 (0,18)
	Sans procuration surtout	0,81 (0,13)	0,89 (0,15)	0,86 (0,16)	9,21 (1,83)	20,63 (2,94)	14,20 (2,94)	97,24 (0,35)	97,50 (0,36)	98,28 (0,37)
	Par procuration surtout	1,01 (0,14)	0,90 (0,15)	0,87 (0,15)	12,27 (1,97)	19,05 (2,58)	17,08 (3,16)	97,52 (0,35)	98,15 (0,38)	96,96 (0,40)
	Par procuration	0,10 (0,13)	0,82 (0,10)	0,95 (0,12)	13,31 (1,83)	15,11 (1,86)	20,07 (2,22)	97,49 (0,27)	97,56 (0,24)	97,52 (0,27)
	Total	0,89 (0,11)	0,78 (0,11)	0,79 (0,11)	11,13 (0,90)	16,04 (1,21)	17,00 (1,21)	97,84 (0,11)	98,20 (0,11)	98,00 (0,11)

Tableau 3
 Comparaison des estimations MLCA aux estimations publiées antérieurement

Classification		MLCA	Chua et Fuller (données de 1982)	Poterba et Summers (données de 1981)	CPS Réinterview rapprochée (1977-1982)
Réelle	Observée				
Employé	Employé	98,77 (1993)	98,66 (1 mois)	97,74	98,78
		98,73 (1995)	98,65 (2 mois)		
		98,73 (1996)			
	Sans emploi	0,34 (1993)	0,32 (1 mois)	0,54	0,19
		0,49 (1995)	0,34 (2 mois)		
		0,37 (1996)			
	Inactif	0,89 (1993)	1,02 (1 mois)	1,72	1,03
		0,78 (1995)	1,01 (2 mois)		
		0,79 (1996)			
Sans emploi	Employé	7,06 (1993)	3,52 (1 mois)	3,78	1,91
		7,86 (1995)	3,51 (2 mois)		
		8,57 (1996)			
	Sans emploi	81,81 (1993)	88,27 (1 mois)	84,76	88,57
		76,09 (1995)	88,23 (2 mois)		
		74,42 (1996)			
	Inactif	11,13 (1993)	8,21 (1 mois)	11,46	9,53
		16,04 (1995)	8,16 (2 mois)		
		17,00 (1996)			
Inactif	Employé	1,41 (1993)	1,60 (1 mois)	1,16	0,5
		1,11 (1995)	1,61 (2 mois)		
		1,13 (1996)			
	Sans emploi	0,75 (1993)	1,19 (1 mois)	0,64	0,29
		0,69 (1995)	1,24 (2 mois)		
		0,87 (1996)			
	Inactif	97,84 (1993)	97,21 (1 mois)	98,2	99,21
		98,20 (1995)	97,15 (2 mois)		
		98,00 (1996)			

Une explication de cette différence est que les estimations comparatives manifestent un biais vers le haut à cause de corrélations entre les erreurs d'interview et de réinterview. Une autre explication est que les estimations MLCA comportent un biais vers le bas à cause du manque de validité de l'hypothèse markovienne. Nous estimons que les deux explications sont valables dans une certaine mesure. Toutefois, nous indiquons à la section suivante que l'échec de l'hypothèse markovienne exerce probablement une faible influence sur les estimations de l'erreur de classification.

3.6 Robustesse de la stratégie MLCA à l'égard des changements non markoviens de la situation vis-à-vis de l'activité

Plusieurs auteurs ont examiné l'effet de la situation actuelle et antérieure vis-à-vis de l'emploi sur la situation future vis-à-vis de l'emploi (voir par exemple Akerlof et Main 1980; Heckman et Borjas 1980; Lynch 1989, Corak 1993). Heckman et Borjas ont indiqué que l'étude de cette question est assez difficile à cause des biais de sélection, des erreurs de réponse et de l'hétérogénéité non observée. Ces sources de confusion rendent peut-être compte des résultats peu cohérents de la documentation. Ainsi, à l'aide de données de

la CPS, Akerlof et Main (1980) ont pu indiquer que la probabilité d'un futur chômage dépend du nombre de situations antérieures de chômage de même que de la durée de ces situations. Toutefois, dans une étude d'hommes ayant terminé leurs études secondaires, Heckman et Borjas (1980) n'ont trouvé aucune indication que les situations antérieures de chômage ou leur durée ont un effet sur le futur comportement vis-à-vis de l'activité lorsque l'on tient compte du biais de tirage de l'échantillon et du biais dû à l'hétérogénéité. Les résultats tirés de la documentation sont également incohérents et ambigus relativement à la mesure dans laquelle l'hypothèse markovienne exprimée en (2) risque d'être enfreinte pour la CPS et d'autres enquêtes sur la population active. Néanmoins, dans la présente section, nous cherchons à fournir une réponse au moins partielle à la question de savoir comment les changements non markoviens de la situation vis-à-vis de l'activité influencent les estimations MLCA de l'erreur de classification.

Afin de pouvoir étudier l'effet des infractions à l'hypothèse markovienne en (2) pour la présente application, nous avons mené une étude de simulation limitée. Afin d'orienter la recherche tout en simplifiant le cadre de la simulation,

nous avons considéré des structures latentes comportant uniquement deux classes ou situations à un moment donné : chômage, noté X, Y ou $Z=1$, et autre (employé ou inactif), noté X, Y ou $Z=2$ avec des définitions semblables pour les situations observées A, B et C. Afin d'établir une population en vue de la simulation, les probabilités latentes $\pi_x, \pi_{y|x}$, et $\pi_{z|xy}$ et les probabilités de réponse $\pi_{a|x} = \pi_{b|y} = \pi_{c|z}$ ont été précisées comme concordant avec les ensembles combinés de données de 1993, 1995 et 1996.

Nous avons alors défini deux paramètres, λ_1 et λ_2 , variant dans la simulation, où

$$\lambda_1 = \frac{\pi_{z=1|x=2,y=1}}{\pi_{z=1|x=1,y=1}} \quad (17)$$

et

$$\lambda_2 = \frac{\pi_{z=1|x=2,y=2}}{\pi_{z=1|x=1,y=2}}. \quad (18)$$

Ainsi, λ_1 est la probabilité qu'une personne soit « sans emploi » en mars, après avoir été « sans emploi » en février et « autre » en janvier relativement à la probabilité d'être « sans emploi » en mars après avoir été « sans emploi » au cours des deux mois précédents. En conformité avec les résultats obtenus par Akerlof et Main (1980), qui ont montré que la probabilité qu'une personne demeure sans emploi augmente à mesure que le nombre de situations sans emploi augmente, nous supposons que $0 \leq \lambda_1 \leq 1$. De même, λ_2 est la probabilité qu'une personne soit « sans emploi » en mars, après avoir été « autre » au cours des deux mois précédents, relativement à la probabilité d'être « sans emploi », après avoir été « autre » en février et « sans emploi » en janvier. Encore une fois, suivant Akerlof et Main, nous supposons que $0 \leq \lambda_2 \leq 1$. À noter que, lorsque $\lambda_1 = \lambda_2 = 1$, les changements de chômage entre février et mars sont markoviens.

Les données simulées ont été produites de façon à correspondre complètement à un modèle MLCA à probabilités de changement non stationnaires lorsque $\lambda_1 = \lambda_2 = 1$. Nous avons simulé l'échec de l'hypothèse markovienne en variant λ_1 et λ_2 entre 0 et 1. Par souci d'uniformité avec les données 1993-1996, nous avons fixé la probabilité d'une réponse « sans emploi » correcte, $\pi_{a=1|x=1}$, à 0,80 et la probabilité d'une réponse « autre » correcte, $\pi_{a=2|x=2}$, à 0,99 dans toutes les simulations. De plus, les dénominateurs de λ_1 et de λ_2 ont été fixés à leurs valeurs déterminées à partir des données combinées 1993-1996, tandis que les numérateurs ont été calculés à partir de (17) et de (18) à l'aide de valeurs de λ_1 et de λ_2 précisées pour chaque exécution d'une simulation.

Le tableau 4 résume les résultats de la simulation pour $\lambda_1 = \lambda_2 = \lambda$ où l'on fait varier λ entre 0,2 et 1,0 par paliers de 0,2. À noter que pour $\lambda_1 = \lambda_2 = 1,0$, qui correspond à un modèle markovien, les probabilités estimatives de réponse correcte sont exactement comme prévu. Pour des valeurs plus petites de λ_1 et de λ_2 , les estimations manifestent un biais négatif, le biais maximal

correspondant à la valeur la plus faible envisagée, 0,2. Néanmoins, le biais absolu attribuable à des probabilités de changement non markovien ne dépasse jamais 3 points de pourcentage. Les résultats du tableau 4 correspondent à ceux de Bushery et Kindelberger (1999), qui ont utilisé une stratégie un peu différente afin d'illustrer le même caractère de robustesse des modèles MLCA pour des données de la CPS. Les deux études indiquent que le manque de validité de l'hypothèse markovienne ne semble pas être une source importante de biais dans l'estimation des probabilités d'erreur de classification de la CPS.

Tableau 4

Estimations de classification correcte pour des changements non markoviens (les entrées de cellules sont des pourcentages)

Pr (Correct)	$\lambda_1 = \lambda_2 = \lambda$				
	$\lambda = 0,2$	$\lambda = 0,4$	$\lambda = 0,6$	$\lambda = 0,8$	$\lambda = 1,0$ (Markov)
Pr (« chô. » « chô. » réel) = $\pi_{a=1 x=1}$	77,6	78,1	78,7	79,3	80
Pr (« autre » « autre » réel) = $\pi_{a=2 x=2}$	98,6	98,7	98,8	98,9	99

4. Comparaison des estimations de type mlca et de réinterview non rapprochée

4.1 Estimation de Hui-Walter

On peut obtenir un autre ensemble d'estimations des probabilités de réponse des données de réinterview de la CPS à l'aide d'un type de modèle de structure latente d'abord proposé par Hui et Walter (1980). Nous utilisons la notation mentionnée ci-dessus; soit X , la classification réelle de la situation vis-à-vis de l'activité à un moment donné; soit A et A' , les classifications d'interview et de réinterview, respectivement. Soit G , une variable de groupe ment définie comme en (4). Considérons la vraisemblance du tableau groupe \times interview \times réinterview noté GAA' . Soit $\pi_{gaa'}$ la probabilité qu'une personne relevant du groupe g soit classée dans la cellule (a, a') du tableau. Le modèle pour $\pi_{gaa'}$ proposé par Hui et Walter est le suivant :

$$\pi_{gaa'} = \sum_x \pi_g \pi_{x|g} \pi_{a|x} \pi_{a'|x} \quad (19)$$

Dans ce modèle, l'hypothèse des mesures parallèles pour les réponses d'interview et de réinterview est réduite et les probabilités de réponse pour les deux mesures, c'est-à-dire $\pi_{a|x}$ et $\pi_{a'|x}$, sont estimées séparément. L'hypothèse ICE est adoptée comme condition de l'identifiabilité. Il est également supposé que $\pi_{a|x}$ et $\pi_{a'|x}$, ne dépendent pas de la variable de groupe, G , tandis que la fréquence des trois catégories employé, sans emploi et inactif, c'est-à-dire $\pi_{x|g}$, dépend toujours de G .

Sinclair et Gastwirth (1996), dans leur analyse de l'erreur de classification de la population active de la CPS, ont utilisé le sexe comme variable de groupement, et nous utilisons également cette variable de groupement dans notre analyse. Sinclair et Gastwirth ont limité leur analyse à des

hommes blancs et des femmes blanches et à deux catégories d'activité : inactif et actif. Cette dernière catégorie est la somme de nos catégories employé et sans emploi. Dans notre analyse, nous considérons des membres d'échantillon de toutes les races et nous analysons la classification à trois catégories de la situation vis-à-vis de l'activité utilisée dans la stratégie MLCA. Ainsi, l'analyse H-W estime 16 paramètres pour chaque année, ce qui correspond au nombre de degrés de liberté disponibles du tableau $G \times A \times A'$, ne laissant aucun degré de liberté pour vérifier l'ajustement du modèle.

Le logiciel ℓ EM a été utilisé pour l'ajustement du modèle H-W aux données d'interview et de réinterview non rapprochée pour trois périodes de temps coïncidant avec les trois périodes de notre stratégie MLCA : avant 1994, 1995, 1996. Nous avons tenté de limiter l'analyse au premier trimestre seulement de ces périodes de temps. Malheureusement, à cause de la faible taille des échantillons, les estimations ont été plutôt instables. Il a donc fallu utiliser les données de réinterview de l'ensemble des quatre trimestres de ces périodes de temps. Les données antérieures à 1994 ont été recueillies de 1985 à 1988 grâce à un échantillon de réinterviews non rapprochées.

Les résultats de cette comparaison des estimations MLCA et H-W sont résumés dans le tableau 5. Les estimations MLCA sont les mêmes que celles des lignes du tableau 2 étiquetées « Total ». Les estimations H-W sont les probabilités de classification associées à l'interview originale, c'est-à-dire la mesure A en (19). Le tableau montre la comparaison pour les trois années. Puisque le taux d'erreur le plus élevé de la stratégie MLCA est survenu pour la catégorie sans emploi, celle-ci revêt un intérêt particulier dans la comparaison MLCA/H-W.

Tableau 5
Comparaison d'estimations modélisées MLCA et H-W des probabilités de réponse de la CPS selon l'année (écarts-types entre parenthèses)

Classification		1993		1995		1996	
Réelle	Observée	H-W	MLCA	H-W	MLCA	H-W	MLCA
Employé	Employé	99,3 (0,3)	98,8 (0,1)	99,5 (0,7)	98,7 (0,1)	99,6 (0,1)	98,8 (0,1)
	Sans emploi	0,0 (0,0)	0,3 (0,1)	0,0 (n/a)	0,5 (0,1)	0,4 (0,1)	0,4 (0,1)
	Inactif	0,7 (0,3)	0,9 (0,1)	0,5 (0,7)	0,8 (0,1)	0,0 (n/a)	0,8 (0,1)
Sans emploi	Employé	11,1 (1,0)	7,1 (0,7)	11,5 (2,3)	7,9 (0,9)	4,6 (15,2)	8,6 (1,0)
	Sans emploi	74,3 (2,7)	81,8 (1,1)	67,9 (6,1)	76,1 (1,3)	67,6 (11,1)	74,4 (1,4)
	Inactif	14,7 (2,9)	11,1 (0,9)	20,6 (6,5)	16,0 (1,2)	27,9 (5,3)	17,0 (1,2)
Inactif	Employé	2,0 (0,5)	1,4 (0,1)	2,5 (1,5)	1,1 (0,1)	2,6 (1,5)	1,1 (0,1)
	Sans emploi	1,2 (0,3)	0,8 (0,1)	0,5 (0,6)	0,7 (0,1)	0,0 (n/a)	0,9 (0,1)
	Inactif	96,8 (0,6)	97,8 (0,1)	97,0 (1,6)	98,2 (0,1)	97,4 (1,1)	98,0 (0,1)

En général, les deux ensembles d'estimations sont en assez bon accord. Les années 1995 et 1996 ne manifestent aucune différence statistiquement significative (au niveau de 5 %) entre les estimations MLCA et H-W pour ce qui est de la population des sans-emploi. Les estimations antérieures à 1994 affichent des différences appréciables; toutefois, celles-ci s'expliquent peut-être par le fait que les données de réinterview antérieures à 1994 ont été recueillies de 1985 à 1988, plutôt qu'en 1993. Ces différences sont étudiées plus en profondeur dans la section qui suit.

4.2 Comparaison des indices d'incohérence

Comme il a été décrit à la section 3.1, nous estimons l'indice d'incohérence pour chaque période de temps à l'aide des estimations MLCA modélisées des probabilités de réponse. Essentiellement, nous estimons le tableau de tri croisé prévu des interviews-réinterviews à partir des estimations MLCA des probabilités de réponse, et nous appliquons ensuite la formule pour l'indice à ce tableau comme si le tableau était observé. Un deuxième tableau de la classification prévue des interviews-réinterviews peut être estimé à l'aide des estimations H-W des probabilités de réponse. Nous comparons ensuite ces deux ensembles d'estimations à l'estimation de l'indice obtenue directement des données de réinterview de la CPS à l'aide de méthodes traditionnelles (U.S. Bureau of the Census 1985). L'accord entre les trois estimations confirme la validité des trois méthodes.

Le tableau 6 montre les trois méthodes d'estimation de l'indice d'incohérence pour les trois périodes de temps. Comme auparavant, la catégorie sans emploi revêt un intérêt particulier à cause de son taux d'erreur élevé. Puisque les erreurs-types ne sont pas disponibles pour les estimations MLCA ou H-W de l'indice, il n'est pas possible de vérifier l'hypothèse formellement. Toutefois, il existe des erreurs-types pour les estimations traditionnelles que l'on peut utiliser comme approximation des erreurs-types pour les estimations H-W.

Dans l'ensemble, le profil général des estimations MLCA aussi bien que l'ordre de grandeur des estimations MLCA sont en assez bon accord généralement avec les estimations H-W et traditionnelles pour les trois années. Toutefois, pour ce qui est de la catégorie inactif en 1995 et en 1996, les estimations traditionnelles de I sont un peu plus grandes que l'une ou l'autre des estimations modélisées de structure latente. Une analyse plus poussée indique que cette différence est attribuable à un biais de la stratégie d'estimation traditionnelle résultant de l'invalidité de l'hypothèse des mesures parallèles.

Le Bureau of the Census des États-Unis (1985) a montré que si les processus d'interview et de réinterview comportent des fiabilités différentes, l'estimation traditionnelle de l'indice est biaisée. Si, par exemple, la fiabilité des données de réinterview est inférieure à celle des données d'interview, l'estimateur traditionnel de la fiabilité test-retest

sous-estime la fiabilité réelle des données de la CPS; l'indice d'incohérence de la CPS est donc trop élevé.

L'interview et la réinterview de la CPS comporteront des fiabilités différentes si la répartition des erreurs n'est pas égale pour les deux interviews. Il est possible de le vérifier en comparant l'ajustement d'un modèle de type H-W en présence et en l'absence de la restriction $\pi_{a|x} = \pi_{a'|x}$. L'hypothèse d'une fiabilité égale est rejetée si la différence entre les chi-carrés du rapport des vraisemblances pour les deux modèles est supérieure à un chi-carré ayant 6 degrés de liberté. Ce test a été rejeté pour 1995 et 1996 au niveau de signification de 10 %. La différence entre les estimations de la catégorie inactif pour 1995 et 1996 semble donc attribuable en partie au biais des estimations traditionnelles de I .

Tableau 6

Comparaison des estimations MLCA, H-W et traditionnelles de l'indice d'incohérence selon l'année et la classification de la population active

Méthode d'estimation	Classification de la population active			Indice global
	Employé	Sans emploi	Inactif	
1993				
Estimation traditionnelle	8,16 (0,24)	33,49 (1,16)	9,96 (0,27)	11,05 (0,26)
H-W	7.37	34.93	10.07	10.78
MLCA	6.35	28.04	7.63	8.73
1995				
Estimation traditionnelle	6,69 (0,44)	36,28 (2,85)	10,80 (0,56)	10,42 (0,53)
H-W	6.82	37	8.98	9.7
MLCA	6.06	36.19	7.2	8.72
1996				
Estimation traditionnelle	5,93 (0,39)	35,97 (2,68)	11,95 (0,56)	10,61 (0,51)
H-W	5.67	39.46	7.55	8.56
MLCA	5.99	37.39	7.76	9.06

À noter également que les indices H-W et MLCA concordent assez bien pour 1995 et 1996, même s'ils diffèrent quelque peu en 1993. Toutefois, comme il a été noté à la discussion du tableau 5, les comparaisons entre les estimations MLCA et H-W pour cette année-là sont obscurcies par les différentes périodes de temps utilisées pour la construction de l'ensemble des données d'interview-réinterview antérieur à 1994. Cela pourrait expliquer au moins partiellement l'écart entre les estimations pour cette année-là.

5. Sommaire et conclusions

La présente recherche a surtout tenté d'examiner la validité des estimations MLCA de l'erreur de classification de la population active de la CPS, et de déterminer l'efficacité de la stratégie MLCA comme substitut des méthodes

traditionnelles d'évaluation de la qualité des données de la CPS. Nous avons analysé des données d'interview de la CPS pour le premier trimestre des trois années 1993, 1995 et 1996, et nous avons mené une analyse supplémentaire des données de réinterviews non rapprochées de la CPS pour les mêmes périodes de temps à peu près. Les données de réinterview ont représenté une autre stratégie d'estimation de l'erreur de classification de la CPS; comparées aux estimations MLCA, elles ont aidé à examiner la validité de la stratégie MLCA.

Nous avons abordé cinq dimensions de la validité de la stratégie MLCA comme suit :

1. **Diagnostics modélisés.** Nous avons examiné tout un choix de modèles MLCA comportant des variables de groupement définies selon l'âge, la race, le sexe, la scolarité, le mode d'interview et le type de réponse (avec ou sans procuration). Le modèle manifestant la plus grande parcimonie et le meilleur ajustement pour les trois années comporte une variable de groupement définie par le type de réponse (avec ou sans procuration) selon quatre catégories : trois cycles tous menés sans procuration, deux cycles seulement menés sans procuration, deux cycles seulement menés par procuration et trois cycles tous menés par procuration. Pour cette catégorie de modèles, le meilleur a été le modèle 4 (voir le tableau 1), qui précisait des probabilités de changement non homogènes et non stationnaires et des probabilités de réponse non homogènes. Ce modèle a fourni un ajustement adéquat aux données pour les trois années.
2. **Validité de l'ajustement du modèle d'une année à l'autre de la CPS.** Un autre indicateur de la validité du modèle est son ajustement d'un échantillon indépendant à l'autre pour la même population. À supposer que la dynamique de la population active et que la structure des probabilités de réponse pour la CPS soient stables pendant quatre ans, le même modèle général devrait manifester un ajustement adéquat pour les trois années. Le modèle 4 affiche une validité de l'ajustement sur plusieurs années (voir le tableau 1). De plus, d'autres variables de groupement ont été vérifiées dans l'étude, mais le modèle comportant la variable avec ou sans procuration a été le meilleur pour les trois années.
3. **Accord entre les estimations modélisées et de type test-retest des probabilités de réponse.** À l'aide des données d'interview-réinterview non rapprochées tirées de la CPS pour trois périodes (avant 1994, 1995, 1996), nous avons appliqué la méthode H-W à l'estimation des probabilités de réponse et nous avons comparé celles-ci aux estimations MLCA. Un bon accord a été observé pour 1995 et 1996, deux années pour lesquelles les périodes de temps pour les données de réinterview et les données de la CPS manifestaient une étroite concordance (voir le tableau 5). Pour

1993, nous avons observé des différences faibles mais significatives entre les estimations MLCA et les estimations H-W correspondantes. Ces différences pourraient s'expliquer par des différences de période de temps, puisque les données de réinterview ont précédé de quelques années les données d'interview de la CPS.

4. **Accord entre les estimations modélisées et de type test-retest du manque de cohérence.** Nous avons comparé des estimations MLCA modélisées de l'indice d'incohérence aux estimations directes correspondantes tirées du programme de réinterview de la CPS. Les deux ensembles d'estimations concordent assez bien pour les trois années, exception faite pour la catégorie inactif (voir le tableau 6). Pour 1995 et 1996, les différences s'expliquent en partie par le biais de l'estimateur traditionnel résultant de l'échec de l'hypothèse des mesures parallèles. La méthode H-W, qui n'exige pas l'hypothèse des mesures parallèles, permet de produire des estimations de l'indice qui concordent bien avec les estimations MLCA pour 1995 et 1996. Pour 1993, la différence entre les estimations MLCA et H-W est peut-être attribuable à des différences de période de temps pour les ensembles de données de réinterview et de la CPS.
5. **Plausibilité des profils d'erreur de classification.** Les estimations MLCA des probabilités d'erreur de classification semblent plausibles. Les estimations d'un groupe avec/sans procuration à l'autre concordent avec les attentes antérieures voulant que des taux d'erreur moins élevés soient observés pour les répondants sans procuration que pour les répondants par procuration. De plus, les taux d'erreur les plus élevés ont été observés pour la population des sans-emploi et l'ordre de grandeur de ces estimations concordait avec celui d'études antérieures (voir par exemple Fuller et Chua 1985, Abowd et Zellner 1985, Poterba et Summers 1986, Sinclair et Gastwirth 1996; voir le tableau 3).

En résumé, nous n'avons trouvé dans ces analyses aucune raison de mettre en doute la validité de la stratégie MLCA. La méthode a donné de bons résultats pour les cinq épreuves de validité. Nous recommandons donc que la méthode MLCA soit considérée comme une autre méthode d'évaluation de l'exactitude des estimations de la population active de la CPS. La forte concordance entre les estimations MLCA et H-W confirme également la validité de la méthode H-W. Nous recommandons que les deux méthodes soient considérées dans de futures études de la qualité des données de la CPS.

Même si la stratégie MLCA a donné de bons résultats dans nos épreuves, nous recommandons que la méthode soit appliquée avec prudence à d'autres situations. Dans notre analyse, les données de réinterview ont permis d'évaluer la validité des estimations MLCA. Toutefois, les données de réinterview ne sont typiquement pas disponibles dans les

enquêtes par panel et, par conséquent, les analystes ne pourront peut-être appliquer que les critères (1), (2) et (5) ci-dessus pour vérifier la validité du modèle. L'hypothèse markovienne est la clé de la stratégie MLCA. Il se peut que des données recueillies au moyen d'un panel portent sérieusement atteinte à cette hypothèse. Heureusement, l'échec de l'hypothèse markovienne ne semble pas être un facteur important de la validité des estimations MLCA de l'erreur de classification de la population active de la CPS (voir le tableau 4).

Bibliographie

- Abowd, J., et Zellner, A. (1985). Estimating gross labor-force flows. *Journal of Business and Economic Statistics*, 3, 3, 254-283.
- Agresti, A. (1990). *Categorical Data Analysis*. New York : John Wiley & Sons, Inc.
- Akerlof, G.A., et Main, G.M. (1980). Unemployment spells and unemployment experience. *The American Economic Review*, 70, 3, 885-893.
- Bailar, B.A. (1975). The effect of rotation group bias on estimates from panel surveys. *Journal of the American Statistical Association*, 70, 23-30.
- Biemer, P., Bushery, J. et Flanagan, P. (1997). An Application of Latent Markov Models to the CPS. Internal U.S. Bureau of the Census Technical Report.
- Biemer, P., et Forsman, G. (1992). On the quality of reinterview data with applications to the Current Population Survey. *Journal of the American Statistical Association*, 87, 420, 915-923.
- Bohrstedt, G.W. (1983). Measurement. *Handbook of Survey Research*, (Éds., P.H. Rossi, R.A. Wright et A.B. Anderson). New York : Academic Press.
- Bushery, J., et Kindelberger, K. (1999). Simulation Examples for MLC Analysis. Internal U.S. Bureau of the Census Memorandum, Washington, DC, 70-122.
- Chua, T.C., et Fuller, W.A. (1987). A model for multinomial response error applied to labor flows. *Journal of the American Statistical Association*, 82, 397, 46-51.
- Clogg, C., et Eliason, S. (1985). Some common problems in log-linear analysis. *Sociological Methods and Research*, 16, 8-14.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 210, 37-46.
- Corak, M. (1993). Is unemployment insurance addictive? Evidence from the benefit durations of repeat users. *Industrial and Labor Relations Review*, 47, 1, 62-72.
- Forsman, G., et Schreiner, I. (1991). The design and analysis of reinterview: An overview. *Measurement Errors in Surveys*, (Éds., P.P. Biemer, et al.). New York : John Wiley & Sons, Inc., 279-302.
- Fuller, W., et Chua, T.C. (1985). Gross change estimation in the presence of response error. *Proceedings of the Conference on Gross Flows in Labor Force Statistics*. Washington, D.C., U.S. Bureau of the Census and U.S. Bureau of Labor Statistics, 65-77.

- Heckman, J.J., et Borjas, G.J. (1980). Does unemployment cause future unemployment? Definitions, questions, and answers from a continuous time model of heterogeneous and state dependence. *Economica*, 47, 247-283.
- Hess, J., Singer, E. et Bushery, J. (2000). Predicting test-retest reliability from behavior coding. *International Journal of Public Opinion Research*, II, 4, 346-360.
- Hui, S.L., et Walter, S.D. (1980). Estimating the error rates of diagnostic tests. *Biometrics*, 36, 167-171.
- Kleinbaum, D.G., Kupper, L.L. et Muller, K.E. (1988). *Applied Regression Analysis and Other Multivariate Methods*, Boston : PWS-KENT Publishing Co.
- Liu, T.H., et Dayton, C.M. (1997). Model selection information criteria for non-nested latent class models. *Journal of Educational and Behavioral Statistics*, 22, 249-264.
- Lynch, L.M. (1989). The youth labor market in the eighties: Determinants of re-employment probabilities for young men and women. *The Review of Economics and Statistics*, 37-45.
- Meyers, B.D. (1988). Classification-error models and labor-market dynamics. *Journal of Business and Economic Statistics*, 6, 3, 385-390.
- Moore, J.C. (1988). Self/proxy response status and survey response quality. *Journal of Official Statistics*, 4, 2, 155-122.
- O'muircheartaigh, C. (1991). Simple response Variance: Estimation and Determinants. *Measurement Errors in Surveys*, (Éds., P. Biemer, *et al.*). New York : John Wiley & Sons, Inc., 551-574.
- Poterba, J., et Summers, L. (1986). Reporting errors and labor market dynamics. *Econometrics*, 54, 6, 1319-1338.
- Poterba, J., et Summers, L. (1995). Unemployment benefits and labor market transitions: A multinomial logit model with errors in classification. *The Review of Economics and Statistics*, 77, 207-216.
- Poulsen, C.S. (1982). Latent Structure Analysis with Choice Modeling Applications. Thèse de doctorat, Wharton School, University of Pennsylvania.
- Rothgeb, J. (1994). Revisions to the CPS Questionnaire: Effects on Data Quality, U.S. Bureau of the Census. CPS Overlap Analysis Team Technical Report 2, avril 6.
- Schreiner, I. (1980). Reinterview Results from the CPS Independent Reconciliation Experiment (Second Quarter 1978 through Third Quarter 1979). Internal U.S. Bureau of the Census Report.
- Shockey, J. (1988). Adjusting for response error in panel surveys, a latent class approach. *Sociological Methods and Research*, 17, 1, 65-92.
- Sinclair, M., et Gastwirth, J. (1996). On procedures for evaluating the effectiveness of reinterview survey methods: Application to labor force data. *Journal of the American Statistical Association*, 91, 961-969.
- Sinclair, M., et Gastwirth, J. (1998). Estimations des erreurs de classification dans l'enquête sur la population active et analyse de leur incidence sur les taux de chômage publiés. *Techniques d'enquête*, 24, 2, 171-183.
- Singh, A.C., et Rao, J.N.K. (1995). On the adjustment of gross flow estimates for classification error with application to data from the canadian labour force survey. *Journal of the American Statistical Association*, 90, 430, 478-488.
- U.S. Bureau of the Census (1985). Evaluating Censuses of Population and Housing, STD-ISP-TR-5. Washington, D.C. : U.S. Government Printing Office.
- U.S. Bureau of the Census (2000). Current Population Survey: Design and Methodology. U.S. Bureau of the Census Technical Paper 63, Washington, D.C. : Government Printing Office.
- Van De Pol, F. et De Leeuw, J. (1986). A latent markov model to correct for measurement error. *Sociological Methods and Research*, 15, 1-2, 118-141.
- Van De Pol, F., et Langeheine, R. (1997). Separating change and measurement error in panel surveys with an application to labor market data. *Survey Measurement and Process Quality*, (Éds., L. Lyberg, *et al.*). New York : John Wiley & Sons, Inc.
- Vermunt, J. (1997). EM: A General Program for the Analysis of Categorical Data. Tilburg University.
- Wiggins, L.M. (1973). *Panel Analysis, Latent Probability Models for Attitude and Behavior Processing*, Amsterdam : Elsevier S.P.C.