Dynamique du niveau de motorisation automobile des ménages français

Roger Collet

(Proposition de communication pour les 24èmes JMA, Fribourg 2007)

Dynamique du niveau de motorisation automobile des ménages français.

ROGER COLLET *
INRETS , Centre d'Economie de la Sorbonne.

(Version intermédiaire 01/07)

Résumé : Ce document propose un modèle catégoriel probit ordonné dynamique avec hétérogénéité pour analyser le degré de motorisation automobile choisi par les ménages en France. Trois états d'équipement sont considérés : la non, mono, et multi motorisation. L'estimation utilise les observations pondérées du panel Parc Auto 1999-2001. Nous situons notre étude dans le cadre statistique bayésien et résolvons le modèle par la méthode de l'échantillonneur de Gibbs. Nous traitons par ailleurs l'approximation des conditions initiales avec une méthode inspirée de HECKMAN (1981), et de l'estimateur de BLUNDELL et SMITH (1991). En spécifiant un modèle latent autoregressif simple, nous adaptons quelques méthodes issues des modèles linéaires dynamiques pour interpréter notre modèle de choix ordonnés. Notamment, nous évaluons pour des ménages moyens les effets à court et long terme d'un changement résidentiel vers une zone francilienne sur leurs probabilités de motorisation, ainsi que les élasticités au revenu de court et long terme des trois probabilités d'équipement.

Mots clés : Motorisation, Automobile, Probit dynamique, Analyse bayésienne, Echantillonnage de Gibbs.

Classification JEL : C11, C25.

* Correspondance à :
Roger Collet, 39 boulevard Henri Martin, 02200 Soissons (FR).
E-mail : roger.collet@laposte.net

Table des matières

Introduction

Le nombre d'automobiles dans un ménage a un impact incontestable sur sa mobilité, tant sur un plan qualitatif, comme le choix des modes de déplacement (UNCLES, 1987), que sur un plan quantitatif, comme le nombre et les distances de déplacements (MEURS, 1990). C'est pourquoi la modélisation du comportement d'équipement en automobiles des ménages devient une nécessité lorsque l'on veut comprendre et prévoir leur demande de déplacement. Il s'agit là d'une préoccupation de premier rang pour la gestion de l'environnement et des infrastructures routières, pour l'industrie automobile et pour les services d'assurances... La place de l'automobile dans les pays développés et ses problématiques d'intérêt croissant ont accéléré la constitution de sources de données et ont rendu encore plus fort ce besoin de modélisation. Quelques unes sont structurées en panel, c'est à dire qu'elles permettent le suivi d'un échantillon de ménages au cours du temps : le « Dutch National Mobility Panel » aux Pays-Bas, le « Sydney Automobile Panel » en Australie, ou encore « Parc Auto » en France en sont quelques exemples pour le transport. NOBILE et al. (1997), ou encore KITAMURA et BUNCH (1990) ont ainsi modélisé la possession automobile sur ce type de données. Cette double dimension des panels est un atout d'analyse certain par rapport aux enquêtes « unidimensionnelles » (séries temporelles ou coupes transversales) puisqu'elle permet de considérer conjointement la dynamique et l'hétérogénéité observable et non observable des comportements (SEVESTRE, 2003, p. 3). Cette structure de données semble donc avantageuse pour modéliser le taux de motorisation individuel puisqu'il nous vient plusieurs motifs élémentaires faisant dépendre le niveau courant d'équipement automobile du niveau de motorisation passé. Par exemple l'automobile est à l'échelle du ménage un bien durable dont le parc à disposition n'est pas fréquemment remis en question. C'est aussi un bien relativement peu flexible, dont les quantités observées peuvent s'ajuster avec retard aux quantités désirées, notamment à cause du temps de recherche pour l'acquisition ou la cession (GOODWIN et MOGRIDGE, 1981). Egalement, l'automobile est potentiellement un bien dont on prend l'habitude de l'usage, ce qui tend à maintenir le taux d'équipement individuel... Du point de vue statistique, cette potentielle dépendance temporelle des comportements de motorisation requiert l'usage d'une spécification économétrique dynamique, qui intègre donc des variables retardées. Parmi celles-ci, la plus naturelle est une variable de motorisation passée qui devient endogène s'il existe par ailleurs un facteur d'hétérogénéité inobservable. Dans les modèles statistiques de réponses qualitatives ordonnées à base de variables latentes, l'existence de ces deux déterminants réclame des outils d'estimation relativement complexes, comme le maximum de vraisemblance approximée utilisant des points de quadrature (BUTLER et MOFFIT, 1982), ou le maximum de vraisemblance simulée (GOURIEROUX et MONFORT, 1994). Dans la littérature des transports, quelques exemples de modélisation dynamique portant sur l'état de motorisation d'observations désagrégées et utilisant ces techniques sont KITAMURA et BUNCH (1990), HANLY et DARGAY (2000), DARGAY et al. (2006). Une solution pratique pour conserver les méthodes d'estimation simples consiste à ne considérer qu'un seul des deux effets, en posant l'hypothèse

que l'autre n'existe simplement pas. Par exemple, HANLY et DARGAY (2000) estiment également des modèles sans effet d'hétérogénéité inobservable. Mais idéalement, la modélisation doit tenir compte conjointement de ces deux effets.

Dans ce document, nous utilisons un modèle probit ordonné dynamique avec hétérogénéité pour décrire le degré de motorisation automobile individuel en France. Trois catégories de ménages sont considérées : les ménages non motorisés, ceux mono motorisés, et les multi motorisés. Le modèle est ajusté sur les données du panel d'observations pondérées, couvrant la période 1999-2001 de l'enquête « Parc Auto ».

Alors que les études précédentes utilisent un ensemble d'indicatrices du niveau (observé) de motorisation passée comme facteurs explicatifs dynamiques, notre modélisation probit s'en démarque et s'appuie sur une spécification latente linéaire autorégressive, inspirée de PAAP et FRANSES (2000). Cette formulation particulière du modèle latent nous permet d'étendre les propriétés d'analyse des modèles linéaires dynamiques aux outils d'interprétation du probit ordonné, sous la forme d'effets probabilistes marginaux de court et long terme. Par ailleurs, une forme flexible réduite, empruntée à l'approche d'HECKMAN (1981) et à l'estimateur linéaire de BLUNDELL et SMITH (1991), est utilisée pour approximer les conditions initiales du modèle latent.

Dans une première section, nous présentons notre source de données et justifions le modèle statistique utilisé pour notre étude. Celui-ci est formellement exposé, puis ses outils d'interprétation sont présentés dans la deuxième et troisième section. Nous situons dans une quatrième partie dans un cadre économétrique bayésien, et l'estimons à l'aide de l'échantillonneur de Gibbs (augmenté), une méthode de Monte Carlo par chaînes de Markov particulière. Nous observerons ensuite quelques statistiques descriptives du panel Parc Auto relatives à la motorisation des ménages dans une cinquième section, avant de commenter dans une sixième partie l'ensemble des résultats d'estimation. La dernière section est consacrée à leur exploitation. Nous mesurons la vitesse d'ajustement du comportement des ménages aux variations de leurs caractéristiques, et nous examinons les effets de court et long terme d'un déménagement dans une zone francilienne (Paris, petite ou grande couronne) sur les probabilités de motorisation de ménages synthétiques. Enfin, les élasticités au revenu des niveaux d'équipement automobile à court et long terme sont présentées.

1 Données et choix du modèle

Les données de « Parc Auto » nous ont servi de support empirique pour conduire notre analyse du niveau d'équipement automobile des ménages. Exploitée à l'INRETS, cette enquête est réalisée annuellement par SOFRES auprès d'un échantillon de 10000 ménages français, avec un renouvellement d'environ 1/3. Son ambition est la connaissance générale du parc automobile à la disposition des ménages, notamment en termes d'équipement et d'usage. Ainsi, « Parc Auto » constitue une source de données très ciblée, qui fait une description approfondie des

attributs, de la qualité et de l'utilisation des automobiles, mais aussi de leurs utilisateurs (individus et ménages).

Pour notre étude, il convient en premier lieu de choisir dans la panoplie des modèles économétriques celui qui semble le mieux adapté pour décrire une variable quantitative d'équipement. Usuellement, ce choix est dicté par la nature de la variable à expliquer. Dans l'enquête « Parc Auto », nous disposons individuellement pour chaque ménage du nombre d'automobiles à sa disposition, et pour caractériser cette variable quantitative discrète, à valeurs positives, un modèle de comptage de type Poisson semble théoriquement le plus adapté. Mais au regard de la distribution des ménages par nombre de voiture, il apparaît que 95% des ménages entre 1999 et 2001 possèdent aucune, une ou deux automobiles dans leur parc. Autrement dit, les effectifs de ménage possédant trois automobiles et davantage sont relativement réduits : 4.7% en 1999, 4.6% en 2000, 5.1% en 2001 (HIVERT (2001), INRETS (2001, 2002)), et l'ajustement d'un modèle de Poisson, théoriquement défini sur $\mathbb{N}=\{0,1,...,+\infty\}$, sur des observations quasi totalement distribuées sur les trois premières modalités de comptage $\{0,1,2\}$ peut sembler faiblement approprié. C'est pourquoi un modèle économétrique dont le nombre possible d'output est ajustable et limité parait plus adapté à nos observations : cela nous amène à rassembler en tranches des modalités de comptages mitoyennes avec des effectifs trop faibles.

Pratiquement, nous avons conservé la classe des ménages non équipés, celle des ménages équipés d'une seule voiture, et nous avons agrégés les ménages disposant de deux voitures ou plus en une classe de ménages multi motorisés. La variable que nous souhaitons modéliser est donc ici catégorielle avec trois modalités, les ménages ayant été ventilés en trois ensembles mutuellement exclusifs : non motorisés, mono motorisés, et multi motorisés. Compte tenu de l'ordonnancement logique de ces trois niveaux de motorisation, selon la quantité d'automobiles à disposition, un calque statistique adapté pour la représentation de cette variable (discrète et ordonnée) peut être choisi dans la famille des modèles de choix discrets ordonnées.

2 Le modèle catégoriel probit ordonné dynamique

Une bonne façon d'introduire les modèles catégoriels ordonnés est de partir des observations. Notons Y_{it} la variable du niveau de motorisation pour le ménage i à la période t, qui peut dans notre cas prendre trois valeurs $\left(Y_{it}=j \mid j\in\{0,1,2\}\right)$, puisque sont pris en compte trois états de motorisation :

$$\begin{aligned} &Y_{it}=0 \text{ si le ménage } i \text{ est non motorisé à la période } t \\ &Y_{it}=1 \text{ s'il est mono motorisé} \\ &Y_{it}=2 \text{ s'il est multi motorisé} \end{aligned} \qquad (1)$$

Y_{it} est donc une variable codée issue de la motorisation observée du ménage, que les modèles de choix ordonnés supposent généralement résulter de la discrétisation d'une variable latente Y_{it}^*, continue et inobservée, obéissante à la règle suivante :

$$\begin{aligned} &Y_{it} = 0 \text{ si } -\infty < Y_{it}^* \leq s_1 \\ &Y_{it} = 1 \text{ si } s_1 < Y_{it}^* \leq s_2 \\ &Y_{it} = 2 \text{ si } s_2 < Y_{it}^* \leq +\infty \,, \end{aligned} \tag{2}$$

avec $\{s_1, s_2\}$, le couple des seuils inconnus qui règlent ainsi la classification des ménages dans l'un des trois états de motorisation. Le système (2) est la clé de passage entre variables latentes inobservées Y^* et variables « d'output » observée, et pose la base d'un modèle de mesure (XIANG, 1989) dont la spécification économétrique porte uniquement sur Y^*. Les données de panel permettent l'élargissement de spécifications, usuellement statiques, à des configurations dynamiques souvent plus réalistes et dont les propriétés élargissent de fait l'éventail des conclusions. Pratiquement, une spécification est généralement rendue dynamique lorsque sont introduites dans sa formulation des variables endogènes et explicatives retardées. Au premier ordre, le modèle latent peut s'écrire sous la forme d'un modèle autorégressif à retards distribués, noté $ARDL(1,1)$:

$$Y_{it}^* = X_{it}\beta + \gamma Y_{it-1}^* + X_{it-1}\psi + \xi_{it} \,, \tag{3}$$

où X_{is} est la matrice traditionnelle des facteurs explicatifs du ménage i à la période s, Y_{it-1}^* est sa variable endogène retardée, $\{\beta, \gamma, \psi\}$ est un ensemble de paramètres, et ξ_{it} est le terme d'erreur aléatoire. Techniquement, la stationnarité du processus (3) nécessite notamment que le coefficient γ réside dans le cercle unitaire. Economiquement, l'interprétation des comportements est impossible pour $\gamma < 0$. Ces deux conditions réunies imposent finalement que $0 \leq \gamma < 1$. La spécification (3) peut être réécrite sous la forme d'un modèle vectoriel à correction d'erreur, noté VECM. En lui retranchant Y_{it-1}^* de chaque coté, puis en y ajoutant et retirant $X_{it-1}\beta$ à droite de l'égalité, nous obtenons effectivement :

$$\Delta Y_{it}^* = \Delta X_{it}\beta + \lambda\left[Y_{it-1}^* - X_{it-1}\mu\right] + \xi_{it} \,, \tag{4}$$

avec $\lambda = (\gamma - 1)$, et $\mu = -(\psi + \beta)/\lambda$. Cette formulation VECM en (4) se décompose d'une relation d'équilibre, $\Delta Y_{it}^* = \Delta X_{it}\beta$ et d'une erreur de déséquilibre à la période passée : $er_{it-1} = Y_{it-1}^* - X_{it-1}\mu$. Lorsque cette dernière est positive, elle a un effet compensatoire négatif sur Y_{it}^* puisque $\lambda < 0$. Au contraire, elle induit un effet de rattrapage positif sur Y_{it}^* lorsqu'elle est négative. Le paramètre λ agit donc comme un ressort qui permet une correction d'autant plus rapide (ou lente) que γ tend vers 0 (ou 1) : il reflète donc la vitesse d'ajustement de Y^* à X dans leur relation à long terme.

Ce format d'écriture est particulièrement avantageux puisque les paramètres β et μ mesurent respectivement les effets de court et long terme d'un changement permanent de X_i

sur Y_i^*. Par mesure de simplicité, nous posons dans ce document la restriction $\psi = 0$ et faisons ainsi l'hypothèse que le modèle (3) n'est dynamique que par la présence de Y_{it-1}^*. Littéralement, ceci revient à dresser une spécification pour Y_{it}^* qui est une relation linéaire des variables caractéristiques contemporaines, de l'endogène retardée, et du terme d'erreur :

$$Y_{it}^* = X_{it}\beta + \gamma Y_{it-1}^* + \xi_{it} \tag{5}$$

et qui pose une relation de proportionnalité entre effets de court et long terme β et μ, dont le facteur dépend de γ : $\mu = (1-\gamma)^{-1}\beta$.

2.1 Décomposition de l'erreur et paramétrage

Les données de panel nous autorise la prise en considération d'un effet d'hétérogénéité inobservable, noté α_i, individuel et invariant dans le temps. Ainsi, nous pouvons détailler l'erreur globale ξ_{it} en deux composantes :

$$\xi_{it} = \alpha_i + \varepsilon_{it} \tag{6}$$

où les termes d'hétérogénéité α_i sont traités comme des effets aléatoires, tandis que ε_{it} est le terme d'erreur idiosyncratique. Les hypothèses relatives à leur distribution sur la population et à leurs propriétés d'indépendance sont supposées les suivantes :

$$\begin{array}{ll} \alpha_i \sim \mathcal{N}\left(0, \sigma_\alpha^2\right) & \alpha_i \perp \alpha_j \ \forall i \neq j \\ \varepsilon_{it} \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right) & \varepsilon_{it} \perp \varepsilon_{js} \ \forall [i,t] \neq [j,s] \end{array}. \tag{7}$$

Puisque α_i et ε_{it} sont des aléas supposés normaux et indépendants, alors le terme ξ_{it} est normalement distribué :

$$\begin{array}{l} \xi_{it} \sim \mathcal{N}\left(0, \sigma_\xi^2\right), \\ \sigma_\xi^2 = \sigma_\alpha^2 + \sigma_\varepsilon^2 \end{array} \tag{8}$$

Ce paramétrage particulier des erreurs range notre modèle de choix ordonné dans la famille des modèles « probit ». En comparaison, l'usage d'une loi logistique pour ξ_{it} l'aurait placé dans la gamme des modèles « logit ». Ici, l'utilisation d'un probit tient plutôt à des conventions dans notre domaine d'étude qu'à des préoccupations techniques. Pour les choix ordonnés effectivement, les modèles probit sont davantage utilisés en sciences sociales que leurs concurrents logit, mais inversement en biostatistiques par exemple (Powers et Xie, 2000, p. 215).

2.2 Identification du modèle

Sauf à poser quelques restrictions courantes, les modèles probit ne sont pas identifiables. Dans le cas ordonné, une première condition d'identification consiste à ancrer un seuil sur une valeur fixe pour s'affranchir du problème de translation. Ainsi, nous posons par convention :

$$s_1 = 0\,. \tag{9}$$

Pour notre analyse, cette condition impose que les variables latentes soient négatives pour les ménages non motorisés, et inversement, positives pour les ménages motorisés : ceci nous facilite la mise en lumière de comportement d'habitude de motorisation si l'estimation révèle un coefficient γ significativement positif, où autrement dit, si l'état de (non) motorisation passé encourage l'état de (non) motorisation courant. Pour s'affranchir du classique problème d'échelle, une seconde condition d'identification consiste à fixer un autre paramètre : par convention dans les modèles probit, ce problème est résolu en fixant la variance à un, ici :

$$\sigma_\xi^2 = 1 \tag{10}$$

2.3 Traitement des conditions initiales

La présence parmi les variables explicatives de l'endogène décalée Y_{it-1}^* est la source dynamique de notre spécification (5). Cela pose le problème de la période initiale puisque si t varie de la période 1 (pour l'année 1999) jusque 3 (pour l'année 2001) dans notre cas, alors la variable latente n'est définie que pour $t \geq 2$ (pour les années 2000 et 2001).

Nous nous inspirons ici d'une approche proposée par HECKMAN (1981), qui suggère d'adopter une forme flexible réduite pour approximer les conditions de départ pour $t = 1$ et amorcer le modèle général (5). On retrouve également cette idée dans l'estimateur de BLUNDELL et SMITH (SEVESTRE, 2003, p. 130) pour les modèles linéaires dynamiques. Cette forme réduite comprend une spécification « raccourcie » et nécessairement statique, donc exprimée en termes de regresseurs initiaux :

$$Y_{i1}^* = X_{i1}\beta_0 + \varepsilon_{i1} \tag{11}.$$

Les variables explicatives X_{i1} répondent ici au même descriptif de l'observation i que X_{it} dans le modèle général, β_0 est ici le vecteur de paramètres initiaux, alors que le terme d'erreur noté ε_{i1} est supposé suivre une $\mathcal{N}\left(0, \sigma_{\varepsilon_0}^2\right)$. Par analogie avec la règle du modèle général (2), le niveau de motorisation initial est déterminé par la discrétisation de la variable Y_{i1}^* :

$$\begin{aligned} &Y_{i1} = 0 \text{ si } -\infty < Y_{i1}^* \leq s_1' \\ &Y_{i1} = 1 \text{ si } s_1' < Y_{i1}^* \leq s_2' \end{aligned} \tag{12}$$

$$Y_{i1} = 2 \text{ si } s_2' < Y_{i1}^* \leq +\infty$$

Si Y_{i1}^* doit intervenir comme endogène retardée dans le modèle général écrit pour Y_{i2}^*, alors il doit être de même nature qu'une endogène du modèle principal et ainsi obéir aux même clés de passage vers l'état observé de motorisation. Ainsi, les seuils de la règle (12) sont donnés par :

$$s_1' = s_1 \ ; \ s_2' = s_2 \tag{13}$$

L'importation des deux seuils $\{s_1, s_2\}$ du modèle général vers le modèle initial constitue deux modalités d'identification suffisantes pour le modèle initial. Il n'est donc pas nécessaire d'imposer une contrainte sur la variance $\sigma_{\varepsilon_0}^2$, comme c'était le cas dans le modèle général pour σ_ξ^2.

Par ailleurs, le terme d'hétérogénéité α_i n'apparaît pas explicitement dans le modèle initial : s'il existe, alors celui-ci est absorbé par l'erreur ε_{i1}. Pour en tenir compte, nous introduisons le schéma de corrélation suivant :

$$\begin{aligned} &\alpha_i = \delta\varepsilon_{i1} + v_i \\ &\text{avec : } v_i \perp \varepsilon_{i1} \ ; \ v_i \sim \mathcal{N}\left(0, \sigma_v^2\right) \end{aligned} \tag{14}$$

2.4 Vraisemblance du modèle

Les étapes d'identification et de paramétrage effectuées d'une part, et les conditions initiales mises en place d'autre part, nous pouvons écrire la fonction de vraisemblance de notre modèle. Dans ce but, nous considérons tout d'abord l'agent i, dont la pondération dans le panel est notée w_i (avec $\sum_i w_i = W$), et faisons tout d'abord comme si nous avions la connaissance exacte de toutes ses variables inobservables, exceptés les termes d'erreurs $\{\varepsilon_{i3}, \varepsilon_{i2}, \varepsilon_{i1}\}$. Nous pouvons ainsi calculer L_{i3}^*, L_{i2}^*, L_{i1}^* qui sont respectivement les vraisemblances périodiques des variables latentes Y_{i3}^*, Y_{i2}^*, Y_{i1}^*. En observant les regroupements de paramètres $\Theta = \{\beta, \gamma, \delta, \beta_0, \sigma_v^2, \sigma_{\varepsilon 0}^2\}$ et $S = \{s_1', s_2', s_1, s_2\}$, nous avons :

$$L_{i3}^* = L\left(Y_{i3}^* \mid \Theta, Y_{i2}^*, Y_{i1}^*, X_i, v_i\right) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\left(Y_{i3}^* - X_{i3}\beta - \gamma Y_{i2}^* - \delta\left(Y_{i1}^* - X_{i1}\beta_0\right) - v_i\right)^2\right) \tag{15}$$

$$\begin{aligned} &L_{i2}^* = L\left(Y_{i2}^* \mid \Theta, Y_{i1}^*, X_i, v_i\right) = \\ &\frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}} \exp\left(-\frac{1}{2\sigma_\varepsilon^2}\left(Y_{i2}^* - X_{i2}\beta - \gamma Y_{i1}^* - \delta\left(Y_{i1}^* - X_{i1}\beta_0\right) - v_i\right)^2\right) \end{aligned} \tag{16}$$

$$L_{i1}^{*} = L\left(Y_{i1}^{*} \mid \Theta, X_i, v_i\right) = \frac{1}{\sqrt{2\pi\sigma_{\varepsilon 0}^{2}}} \exp\left(-\frac{1}{2\sigma_{\varepsilon 0}^{2}}\left(Y_{i1}^{*} - X_{i1}\beta_0\right)^2\right) \tag{17}$$

Par application de la règle de Bayes, le produit de ces trois quantités (15), (16) et (17) calcule L_i^* la vraisemblance jointe des variables latentes de l'agent i :

$$L_i^{*} = L\left(Y_i^{*} \mid \Theta, X_i, v_i\right) = \prod_{t=1}^{3} L_{it}^{*} \tag{18}$$

Le produit pondéré sur l'ensemble des agents de L_i^* nous donne la vraisemblance des variables latentes de notre échantillon (une statistique cruciale pour la suite) :

$$L\left(Y^{*} \mid \Theta, X, v\right) = \prod_{i}\left(L_i^{*}\right)^{w_i} \tag{19}$$

En retournant au niveau individuel, et sachant les hypothèses (14), on calcule pour i la vraisemblance jointe de ses variables latentes et de v_i par :

$$L\left(Y_i^{*}, v_i \mid \Theta, X_i\right) = L_i^{*} \times P\left(v_i\right) \tag{20}$$

où $P\left(v_i\right)$ est la fonction de densité d'une $\mathcal{N}\left(0, \sigma_v^2\right)$:

$$P\left(v_i\right) = \left(2\pi\sigma_v^2\right)^{-0.5} \exp\left(-0.5\sigma_v^{-2}\left(v_i - 0\right)^2\right)$$

L'intégration de (20) sur le domaine des possibles des variables latentes tout d'abord (défini par les système de décisions (2) et (12) et les seuils S), puis sur v_i ensuite, nous donne la vraisemblance de la motorisation observée de l'observation i :

$$L_i = L\left(Y_i \mid \Theta, S, X_i\right) = \int_{-\infty}^{+\infty}\int_{a_{i1}}^{b_{i1}}\int_{a_{i2}}^{b_{i2}}\int_{a_{i3}}^{b_{i3}} L\left(Y_i^{*}, v_i \mid \Theta, S, X_i\right) dY_{i3}^{*}\, dY_{i2}^{*}\, dY_{i1}^{*}\, dv_i \tag{21}$$

où les bornes d'intégration sont données par:

$$\left.\begin{array}{l} a_{it} = -\infty \\ b_{it} = \tilde{s}_1 \end{array}\right\} \text{si } Y_{it}^{*} = 0 \; ; \quad \left.\begin{array}{l} a_{it} = \tilde{s}_1 \\ b_{it} = \tilde{s}_2 \end{array}\right\} \text{si } Y_{it}^{*} = 1 \; ; \quad \left.\begin{array}{l} a_{it} = \tilde{s}_2 \\ b_{it} = +\infty \end{array}\right\} \text{si } Y_{it}^{*} = 2$$
$$\text{avec : } \tilde{s}_1 = s_1' \text{ pour } t = 1 \;\; ; \;\; \tilde{s}_1 = s_1 \text{ pour } t = \{2,3\} \tag{22}$$
$$\tilde{s}_2 = s_2' \text{ pour } t = 1 \;\; ; \;\; \tilde{s}_2 = s_2 \text{ pour } t = \{2,3\}$$

et où le jeu de contraintes d'identification est:

$$s_1 = 0 \; ; \; s_1' = s_1 \; ; \; s_2' = s_2 \; ; \; \sigma_\varepsilon^2 = 1 - \left(\delta^2\sigma_{\varepsilon 0}^2 + \sigma_v^2\right) \tag{23}$$

Par le produit pondéré des vraisemblances individuelles en (21), nous obtenons la vraisemblance des états de motorisation de notre échantillon :

$$L(Y \mid \Theta, S, X) = \prod_i (L_i)^{w_i} \tag{24}$$

La non connaissance explicite des variables latentes Y^* a pour effet d'introduire des intégrales emboîtées dans la fonction (21), ce qui rend l'estimation des paramètres par maximum de vraisemblance très complexe. Si l'ensemble des variables latentes Y_{it}^* étaient connues, la vraisemblance (21) ne contiendrait plus qu'une intégrale sur v_i, ce qui est beaucoup plus gérable, et un estimateur traditionnel pour modèles linéaires dynamiques avec erreurs composées aurait pu alors être envisagé.

C'est pourquoi une procédure d'estimation par simulations des variables inobservées peut considérablement nous simplifier la tâche. En particulier, nous savons depuis les travaux d'ALBERT et CHIB (1993) que les méthodes de Monte Carlo par chaînes de Markov (MCMC) tel que l'échantillonnages de Gibbs avec augmentation de données sont particulièrement adaptées et efficaces pour l'estimation de modèles économétriques à base de variables latentes. Ces méthodes s'accommodent si bien de la statistique bayésienne qu'elles ont été le moteur sa résurrection au cours des dix dernières années. Même si ces méthodes MCMC ne lui sont pas exclusives (TANNER, 1991), nous adoptons un cadre d'estimation bayésien pour traiter notre modèle ordonné du niveau de motorisation des ménages.

En choisissant la spécification (5) pour le modèle latent, nous souhaitons mettre les propriétés dynamiques avantageuses d'une spécification linéaire autorégressive basique et ses outils d'analyse au service de notre modèle de choix ordonné. Cette méthodologie nous a été inspirée par les travaux relatifs à l'estimation du modèle probit multinomial dynamique de PAAP et FRANSES (2000). Dans la littérature des transports, il existe quelques travaux de modélisation dynamique portant sur l'état de motorisation d'observations désagrégées. Parmi ceux cités en introduction, aucune étude à notre connaissance n'utilise une spécification latente autorégressive, la raison tenant à la nature des termes dynamiques utilisés. Toutes en effet utilisent au mieux un ensemble d'indicatrices du niveau (observé) de motorisation passée comme facteurs explicatifs. Dans la pratique, il est évidemment plus commode d'estimer les paramètres d'une spécification lorsque l'ensemble des regresseurs est connu, plutôt que d'utiliser une spécification incorporant de telles variables latentes non observées comme dans notre cas. Hormis quelques conclusions attestant du rôle significatif et positif de la motorisation passée sur la motorisation courante, l'interprétation et l'exploitation des résultats dynamiques dans ces études ne pouvaient, semble-t-il, pas être aisément approfondies. Notre modèle latent autorégressif (5) permet des conclusions dynamiques plus intéressantes. Notamment, nous verrons ultérieurement que son estimation nous donne la capacité de mesurer le temps nécessaire aux agents d'ajuster leur décision de motorisation aux changements de caractéristiques, mais aussi d'évaluer individuellement les variations/élasticités de probabilités à court et long terme des trois états de motorisation que ces changements induisent.

3 L'interprétation du modèle

Selon le modèle probit ordonné, la modalité observée d'un ménage traduit que sa variable latente se situe dans la fourchette correspondante. Un modèle dont la spécification retenue aurait été statique, qui n'aurait fait dépendre la variable latente que de variables contemporaines impliquerait un ajustement instantané total aux variations de ses déterminants. Autrement dit, une variation de revenu ou un changement de résidence d'un ménage par exemple seraient complètement et immédiatement répercutés sur le niveau de sa variable latente, et donc de son niveau de motorisation. On se doute évidemment qu'une telle flexibilité n'est que théorique, surtout concernant l'équipement automobile. En effet, l'acquisition ou la cession d'une automobile par un agent sont des actes généralement précédés d'étapes et d'activités coûteuses en temps, les plus évidentes étant la réflexion, la recherche (d'informations, de prix, des modalités de stationnement et d'assurances…), voire l'attente (d'une meilleure opportunité, de conception du modèle désiré, ou l'hésitation). Il est donc pratiquement acquis que la décision individuelle de motorisation revêt une dimension temporelle et repose sur une spécification dynamique de la variable latente. Même si ce n'était pas le cas, notre spécification (5) n'en serait pas pour autant inopérante puisque le cas statique lui est un cas particulier, pour $\gamma = 0$.

En substituant les endogènes retardées par leur expression, on peut réécrire le processus générateur de la variable latente (5) et le faire apparaître comme une réponse « lissée » aux modifications abruptes des déterminants X, qui nous servira pour les calculs d'effets marginaux de court et long terme :

$$\begin{aligned} Y^*_{it+h} &= \gamma^t Y^*_{ih} + \sum_{s=0}^{t-1} \gamma^s \left(X_{it+h-s}\beta + \alpha_i + \varepsilon_{it+h-s} \right) \\ &= \gamma^h Y^*_{it} + \sum_{s=0}^{h-1} \gamma^s \left(X_{it+h-s}\beta + \alpha_i + \varepsilon_{it+h-s} \right) \end{aligned} \tag{25}$$

Puisque les termes aléatoires ε_{iz} sont normaux, homoscédastiques, et d'espérance nulle quelque soit z, alors l'espérance, la variance, et la distribution de Y^*_{it} sachant les déterminants présents et passés $X_{i,\leq t} = \{X_{it},...,X_{i1}\}$ nous sont donnés dans la relation (25) par :

$$\begin{aligned} \mu_{it} &= \mathrm{E}\left[Y^*_{it} \mid X_{i,\leq t} \right] = \gamma^t Y^*_{i0} + \sum_{s=0}^{t-1} \gamma^s \left(X_{it-s}\beta \right) \\ \omega_{it} &= \mathrm{V}\left[Y^*_{it} \mid X_{i,\leq t} \right] = \mathrm{V}\left(\sum_{s=0}^{t-1} \gamma^s \left(\alpha_i + \varepsilon_{it-s} \right) \right) = \left(\frac{1-\gamma^t}{1-\gamma} \right)^2 \sigma^2_\alpha + \frac{1-\gamma^{2t}}{1-\gamma^2} \sigma^2_\varepsilon \\ \left(Y^*_{it} \mid X_{i,\leq t} \right) &\sim \mathcal{N}\left(\mu_{it}, \omega_{it} \right) \end{aligned} \tag{26}$$

Dans la partie suivante, nous examinons les propriétés dynamiques des statistiques en (26), et mesurons l'effet de court et long terme qu'induit une modification de déterminant sur la variable latente avec les multiplicateurs linéaires standards.

3.1 Effets latents et multiplicateurs

a) Dans le court terme

La relation (25) établie que l'impact d'une modification de la caractéristique k en $t-s$ sur la variable latente courante se calcule par :

$$\frac{dY_{it}^{*}}{dX_{ikt-s}} = \gamma^{s}\beta_{k} \tag{27}$$

Nous retrouvons un résultat illustrant cette idée ordinaire que l'influence sur le présent des évènements passés diminue avec le temps. En effet, pour $0 \leq \gamma < 1$, nous obtenons que l'impact absolu est décroissant avec le temps s qui les sépare :

$$\lim_{s\to\infty}\left(\left|\frac{dY_{it}^{*}}{dX_{ikt-s}}\right|\right) = 0 \tag{28}$$

L'effet de court terme sur la variable latente Y_{it}^{*} correspond à l'effet instantané d'une variation de caractéristique X_{ikt}, et s'obtient a partir de (27) pour $s=0$:

$$dY_{it}^{*} = m_{k}^{\mathrm{CT}} \times dX_{ikt} \tag{29}$$

ou $m_{k}^{\mathrm{CT}} = \beta_{k}$ est le multiplicateur de court terme de X_{ik} sur la variable latente.

b) Dans le long terme

Nous supposons que la période t est suffisamment ancienne pour que ses effets sur la variable latente en $t+h$ soient dissous, ou formellement que h est suffisamment grand pour que $\gamma^{h} \approx 0$ dans l'équation (25). Si de plus nous considérons le gel des caractéristiques tel que $X_{i,[t;t+h]} = \{X_{it},...,X_{it+h}\} = \underline{X}_{i}$, alors (25) se simplifie, et les moments limites de Y_{it+h}^{*} sont donnés par:

$$Y_{it+h}^{*} = \sum_{s=0}^{h}\gamma^{s}\left(\underline{X}_{i}\beta + \alpha_{i} + \varepsilon_{it+h-s}\right) \tag{30}$$

$$\begin{aligned}
\mu_{(\underline{X}_{i})}^{\mathrm{LT}} &= \lim_{h\to\infty}\mathrm{E}\left[Y_{it+h}^{*} \mid X_{i,[t;t+h]} = \underline{X}_{i}\right] = \frac{1}{1-\gamma}\underline{X}_{i}\beta \\
\omega^{\mathrm{LT}} &= \lim_{h\to\infty}\mathrm{V}\left[Y_{it+h}^{*} \mid X_{i,[t;t+h]} = \underline{X}_{i}\right] = \frac{1}{(1-\gamma)^{2}}\sigma_{\alpha}^{2} + \frac{1}{1-\gamma^{2}}\sigma_{\varepsilon}^{2} \\
\left(Y_{i\infty}^{*} \mid X_{i,\geq t} = \underline{X}_{i}, \beta, \alpha_{i}\right) &\sim \mathcal{N}\left(\mu_{(\underline{X}_{i})}^{\mathrm{LT}}, \omega^{\mathrm{LT}}\right)
\end{aligned} \tag{31}$$

Nous constatons dans ces conditions que les statistiques caractérisant la distribution de Y_{it+h}^{*} en (31) sont indépendantes des périodes : celle ci est donc stationnaire et l'équilibre latent

de long terme est atteint. Le multiplicateur de long terme, ou d'équilibre (GREENE, 2003, p.562) associé à la k-ième variable explicative de $\underline{X}_i$ nous est donné par :

$$m_k^{\mathrm{LT}} = \lim_{h\to\infty}\left(\frac{dY^*_{it+h}}{d\underline{X}_{ik}}\right) = \frac{\beta_k}{1-\gamma} \tag{32}$$

L'effet latent de long terme résultant d'un changement permanent de la caractéristique k correspond à la variation induite sur $Y^*_{i\infty}$:

$$dY^*_{i\infty} = d\underline{X}_{ik} \times m_k^{\mathrm{LT}} \tag{33}$$

Notons qu'une autre manière d'aboutir aux multiplicateurs de long terme (32) est de cumuler les multiplicateurs de court terme et décalés de l'équation (27) pour $s>0$: $m_k^{\mathrm{LT}} = \beta_k + \sum_{s=1}^{\infty}\gamma^s\beta_k = (1-\gamma)^{-1}\beta_k$.

Les effets latents de court et long terme que nous venons de dériver ne sont que des étapes intermédiaires : ils interviennent au niveau de la variable latente Y_i^*, et n'entretiennent ainsi qu'un rapport indirect avec les états de motorisation Y_i. Il est donc nécessaire d'étendre l'analyse aux effets marginaux de court et long terme sur les probabilités des états de motorisation.

3.2 Effets marginaux à court et long terme

a) Dans le court terme

Les effets marginaux sont l'outil d'interprétation habituel du modèle probit ordonné, et évaluent dans notre étude les variations de probabilité $\Pr[Y_{it}=j \mid X_{i,\le t}]$ induites par la modification de caractéristiques. Une façon pratique de calculer ces probabilités est de « centrer-réduire » la distribution de $Y^*_{it} \mid X_{i,\le t}$ et d'utiliser la fonction de répartition Φ d'une $\mathcal{N}(0,1)$:

$$\begin{aligned}
\Pr\left[Y_{it}=0 \mid X_{i,\le t}\right] &= \Phi\left(\left(s_1-\mu_{it}\right)\Big/\sqrt{\omega_{it}}\right)\\
\Pr\left[Y_{it}=1 \mid X_{i,\le t}\right] &= \Phi\left(\left(s_2-\mu_{it}\right)\Big/\sqrt{\omega_{it}}\right) - \Phi\left(\left(s_1-\mu_{it}\right)\Big/\sqrt{\omega_{it}}\right)\\
\Pr\left[Y_{it}=2 \mid X_{i,\le t}\right] &= 1-\Phi\left(\left(s_2-\mu_{it}\right)\Big/\sqrt{\omega_{it}}\right)
\end{aligned} \tag{34}$$

L'effet à court terme d'une variation de X_{it}, notée ΔX_{it}, est une variation de μ_{it} qui fait intervenir les multiplicateurs de court terme dans (29), et mesurée par $\Delta\mu_{it} = \Delta X_{it} m^{\mathrm{CT}}$. Il résulte de cette variation un nouveau système de probabilité :

$$\Pr\left[Y_{it}=0 \mid X_{i,\leq t},\triangle X_{it}\right]=\Phi\left(\frac{s_1-\left(\mu_{it}+\triangle X_{it}m^{\mathrm{CT}}\right)}{\sqrt{\omega_{it}}}\right)$$
$$\Pr\left[Y_{it}=1 \mid X_{i,\leq t},\triangle X_{it}\right]=\Phi\left(\frac{s_2-\left(\mu_{it}+\triangle X_{it}m^{\mathrm{CT}}\right)}{\sqrt{\omega_{it}}}\right)-\Phi\left(\frac{s_1-\left(\mu_{it}+\triangle X_{it}m^{\mathrm{CT}}\right)}{\sqrt{\omega_{it}}}\right) \qquad (35)$$
$$\Pr\left[Y_{it}=2 \mid X_{i,\leq t},\triangle X_{it}\right]=1-\Phi\left(\frac{s_2-\left(\mu_{it}+\triangle X_{it}m^{\mathrm{CT}}\right)}{\sqrt{\omega_{it}}}\right)$$

Les effets marginaux de court terme, notés $\Lambda_{ij}^{CT}\left(\Delta X_{it}\right)$ se déduisent en différenciant les systèmes de probabilités (35) et (34), et s'interprètent pour chaque état de motorisation $Y_{it}=j$ comme la variation de probabilité résultant d'un changement de caractéristique $\triangle X_{it}$:

$$\Lambda_{ij}^{CT}\left(\Delta X_{it}\right)=\Pr\left[Y_{it}=j \mid X_{i,\leq t},\triangle X_{it}\right]-\Pr\left[Y_{it}=j \mid X_{i,\leq t}\right] \qquad (36)$$

b) Dans le long terme

Dans une démarche identique, la conséquence à long terme d'une variation permanente de $\underline{X}_i$ mesurée par $\Delta\underline{X}_i$ sur états de motorisation est le système de probabilités suivant, qui fait intervenir les multiplicateurs d'équilibre m^{LT} de (33) :

$$\Pr[Y_{i\infty}=0 \mid X_{i,\geq t}=\underline{X}_i,\Delta\underline{X}_i]=\Phi\left(\frac{s_1-(\mu_{(\underline{X}_i)}^{\mathrm{LT}}+\Delta\underline{X}_i m^{\mathrm{LT}})}{\sqrt{\omega^{\mathrm{LT}}}}\right)$$
$$\Pr[Y_{i\infty}=1 \mid X_{i,\geq t}=\underline{X}_i,\Delta\underline{X}_i]=\Phi\left(\frac{s_2-(\mu_{(\underline{X}_i)}^{\mathrm{LT}}+\Delta\underline{X}_i m^{\mathrm{LT}})}{\sqrt{\omega^{\mathrm{LT}}}}\right)-\Phi\left(\frac{s_1-(\mu_{(\underline{X}_i)}^{\mathrm{LT}}+\Delta\underline{X}_i m^{\mathrm{LT}})}{\sqrt{\omega^{\mathrm{LT}}}}\right) \qquad (37)$$
$$\Pr[Y_{i\infty}=2 \mid X_{i,\geq t}=\underline{X}_i,\Delta\underline{X}_i]=1-\Phi\left(\frac{s_2-(\mu_{(\underline{X}_i)}^{\mathrm{LT}}+\Delta\underline{X}_i m^{\mathrm{LT}})}{\sqrt{\omega^{\mathrm{LT}}}}\right)$$

En différenciant les systèmes de probabilités (37) et (34), nous obtenons les effets marginaux de long terme $\Lambda_{ij}^{\mathrm{LT}}\left(\Delta\underline{X}_i\right)$, qui mesurent la variation limite de probabilité résultant d'un changement de caractéristiques permanent $\triangle\underline{X}_i$:

$$\Lambda_{ij}^{\mathrm{LT}}\left(\Delta\underline{X}_i\right)=\Pr\left[Y_{i\infty}=j \mid X_{i,>t}=\underline{X}_i,\Delta\underline{X}_i\right]-\Pr\left[Y_{it}=j \mid X_{i,\leq t}\right] \qquad (38)$$

4 Estimation bayésienne MCMC

Les techniques d'estimation à l'aide de méthodes de Monte Carlo par Chaînes de Markov (MCMC) constituent un apport majeur pour l'analyse des données. Pour JACKMAN (2000), elles sont « probablement le développement statistique le plus excitant » depuis les années 1990. Ces techniques viennent notamment secourir des méthodes bayésiennes jusqu'alors peu attractives face aux méthodes classiques, tant leurs manipulations algébriques

sont souvent lourdes et décourageantes[1]. La réunion du cadre bayésien et de l'outil MCMC rend désormais accessible l'estimation de modèles traditionnellement complexes, et autorise des spécifications larges pour une bonne exploitation des données de panel notamment. Dans la littérature, ALBERT et CHIB (1993) font figure de pionniers dans l'usage de ces méthodes pour l'estimation de modèles probit. Notamment sur ces bases, les travaux de McCULLOCH et ROSSI (1994) ont exposé la première estimation bayésienne calculable du modèle probit multinomial. Par la suite, ce modèle a été exploité par NOBILE et al. (1997) pour modéliser le degré de motorisation aux Pays-Bas, en utilisant un algorithme MCMC hybride.

4.1 Quelques concepts bayésiens d'économétrie

La pièce maîtresse de l'analyse bayésienne est la célèbre règle de Bayes en personne : $P(A \mid B) = \big(P(B \mid A) \times P(A)\big) / P(B)$. Elle fait ici apparaître deux types de probabilité : marginales $(P(A), P(B))$ d'une part, puis conditionnelles $(P(A \mid B), P(B \mid A))$ d'autre part. La lecture usuelle du produit de cette règle, $P(A \mid B)$, est la probabilité de voir se réaliser l'évènement A sachant que l'évènement B s'est réalisé. L'adaptation économétrique de la règle de Bayes constitue le cœur de l'économétrie bayésienne. En général, l'économètre porte un grand intérêt à l'estimation de paramètres θ et utilise pour cela des données y à sa disposition comme source d'information. Autrement dit, l'estimation traite de l'état des connaissances concernant θ, après l'examen de (ou « *sachant* ») y, et que l'on note $P(\theta \mid y)$: telle est la préoccupation fondamentale du chercheur bayésien. En décomposant avec la règle de Bayes précédente, nous obtenons :

$$P(\theta \mid y) = \frac{P(y \mid \theta) \times P(\theta)}{P(y)}$$

Les données y étant constituées d'observations, elles ne sont pas modulables, et leur probabilité $P(y)$ est donc constante quelque soit la valeur des paramètres θ. Autrement dit, puisque l'intérêt ne porte que sur θ, qui n'intervient pas dans $P(y)$, ce dernier terme peut être ignoré. Dans l'usage, les données y ont une valeur informative et servent l'actualisation des paramètres. Ainsi il vient que :

$$P(\theta \mid y) \propto P(y \mid \theta) \times P(\theta) \tag{39}$$

Dans la terminologie bayésienne, l'élément $P(\theta \mid y)$ renvoie à la distribution ou densité de probabilité des paramètres *a posteriori*, le terme $P(y \mid \theta)$ est la fonction de densité des observations sachant les paramètres, alors que $P(\theta)$ concerne la distribution ou fonction de densité des paramètres *a priori*.

[1] La lecture de l'analyse bayésienne du modèle de régression classique dans Greene (2003, p430), est en ce sens assez convaincante.

La densité *a priori* des paramètres, $p(\theta)$, traduit les croyances initiales du modélisateur, ou toute l'information qu'il détient avant l'exploration des données. Elle peut être ajustée pour en refléter diverses nuances : si aucune information n'est *a priori* disponible sur la valeur des paramètres, alors il est usuel et pratique d'assigner à θ une distribution *a priori* $p(\theta)$ conjuguée[2] et très diffuse, c'est-à-dire caractérisée par une variance élevée et une allure (quasiment) « plate ». Au contraire, s'il existe de l'information statistique concernant θ dans la littérature, provenant par exemple de travaux antérieurs ayant utilisé une autre source de données, alors le modélisateur pourra s'en servir pour spécifier une distribution $p(\theta)$ informative.

La fonction de densité des observations sachant les paramètres, notée $P(y/\theta)$, n'est autre que la fonction de vraisemblance classique : elle contient l'information issue des données y sur les paramètres θ.

La distribution *a posteriori* des paramètres est selon la relation (39) proportionnelle au produit de la fonction de vraisemblance avec la distribution *a priori* des paramètres. (ou La densité de la distribution des paramètres « *a posteriori* » $P(\theta \mid y)$ est donc proportionnelle au produit de l'information *a priori* et de l'information courante détenue par l'échantillon (fonction de vraisemblance). Cette distribution, que le bayésien cherche à caractériser, focalise l'intérêt puisqu'elle condense toute l'information - celle *a priori* et celle contenue dans les données - concernant les paramètres θ.

Le passage de la distribution *a priori* vers celle *a posteriori* peut être vu comme un processus bayésien d'apprentissage, où la mise à jour des connaissances initiales s'effectue à lumière des données. D'une façon imagée, la vraisemblance agit telle une boite, qui reçoit en input la distribution *a priori* $p(\theta)$, la modifie selon l'information dont elle dispose en y, et renvoie en ouput la distribution *a posteriori*.

Ces concepts bayésiens maintenant définis, il est important de noter que l'usage de distributions *a priori* très diffuses revient à approximer $p(\theta)$ par une constante. Il en ressort que $P(\theta/y)$ devient proportionnelle à $P(y/\theta)$ et que la vraisemblance décide seule de l'allure de la distribution *a posteriori*. Dans ce cas d'*a priori*, dits non informatifs, les méthodes classiques ou bayésiennes donnent lieu à des inférences sur θ équivalentes.

Cependant, la densité *a posteriori* des paramètres $P(\theta \mid y)$ n'est caractérisable par ses moments qu'au prix de résolutions d'intégrales, et de manipulations algébriques fort décourageantes. Cela explique en particulier pourquoi l'économétrie bayésienne a été si longtemps laissée en disgrâce face aux méthodes classiques, beaucoup plus accessibles pour une large majorité de modèles (cf. note 1). Fort heureusement, et grâce aux propriétés de l'échantillonneur de Gibbs, il est possible de reconstituer empiriquement la forme de la fonction de densité *a posteriori* par échantillonnages successifs dans les lois *a posteriori* conditionnelles des paramètres, et de calculer les moments de leurs distributions marginales.

---

[2] Une distribution *a priori* est dite « conjuguée » lorsque son produit avec la vraisemblance donne naissance à une distribution *a posteriori* de même nature que la distribution *a priori*. Par exemple, le produit d'une vraisemblance normale avec une distribution *a priori* normale sera proportionnel à une distribution *a posteriori* normale.

4.2 Une méthode MCMC : l'échantillonneur de Gibbs[3]

L'échantillonneur de Gibbs est une méthode de Monte Carlo par chaîne de Markov qui nous vient de la science physique, dont on attribue généralement la paternité à GEMAN et GEMAN (1984), et l'introduction dans le domaine statistique à GELFAND et SMITH (1990). A l'origine, GEMAN et GEMAN ont utilisé cet outil pour la restauration d'images. Les auteurs ont été inspiré par un résultat de BESAG (1974), qui montre l'équivalence de détenir une distribution jointe de $m \geq 2$ variables aléatoires, ou détenir m distributions conditionnelles aux autres $m-1$ variables. Dans leur contexte d'imagerie, la distribution jointe des valeurs d'un ensemble de m pixels est compliquée, car de dimension très élevée. Au contraire, les distributions conditionnelles de chaque pixel sont nettement plus manipulables, et sont des distributions de Gibbs. C'est pourquoi l'algorithme mis au point par GEMAN et GEMAN porte le nom d'échantillonneur de Gibbs, bien que leur technique puisse être appliquée dans un autre contexte avec d'autres types de distributions conditionnelles.

a) Fonctionnement

L'algorithme d'échantillonnage de Gibbs est une méthode de Monte Carlo par chaîne de Markov (MCMC) qui constitue un cas particulier de l'algorithme, également MCMC, de Metropolis-Hastings (ROBERT, 1996, p.166). Concrètement, il permet la simulation de distribution jointe à partir de tirages dans des distributions conditionnelles. Lorsque le vecteur de paramètres θ peut être partitionné en $(\theta_1, \theta_2, \theta_3, ...\theta_K)$, et d'une manière telle que l'on puisse faire des tirages dans les distributions *a posteriori* conditionnelles suivantes :

$$P(\theta_i \mid y, \theta_{-i}) \text{ pour tout } i$$

avec $\theta_{-i} = \{\theta_1, ..., \theta_{i-1}, \theta_{i+1}, \theta_K\}$, alors des tirages successifs et continûment réactualisés des paramètres « fraîchement » échantillonnés (dans les lois conditionnelles supra) donneront asymptotiquement des tirages aléatoires dans la distribution jointe *a posteriori* des paramètres $P(\theta \mid y)$. De cette façon, l'échantillonneur de Gibbs simule θ_i à partir des ces distributions conditionnelles, récursivement, et en mettant systématiquement à jour les distributions conditionnelles par la valeur du dernier paramètre échantillonné (CHIB, 1995). C'est-à-dire que l'on effectue une boucle avec les tirages suivants :

$$\begin{array}{ll}
1: & \theta_1^{(g+1)} \text{ dans } P\left(\theta_1 \mid y, \theta_2^{(g)}, \theta_3^{(g)}, ..., \theta_K^{(g)}\right), \\
2: & \theta_2^{(g+1)} \text{ dans } P\left(\theta_2 \mid y, \theta_1^{(g+1)}, \theta_3^{(g)}, ..., \theta_K^{(g)}\right), \\
3: & \theta_3^{(g+1)} \text{ dans } P\left(\theta_3 \mid y, \theta_1^{(g+1)}, \theta_2^{(g+1)}, ..., \theta_K^{(g)}\right), \\
\vdots: & \quad \vdots \qquad \vdots \qquad\qquad \vdots \\
\text{K}: & \theta_K^{(g+1)} \text{ dans } P\left(\theta_K \mid y, \theta_1^{(g+1)}, \theta_2^{(g+1)}, ..., \theta_{K-1}^{(g+1)}\right), \\
\text{K+1}: & g \leftarrow g+1 \text{ ; retour en 1.}
\end{array} \tag{40}$$

[3] En l'honneur de Josiah Willard Gibbs (1839-1903), physicien américain de grande renommée pour ses travaux en thermodynamique et en statistique mécanique.

ou $\theta_1^{(g+1)}$ est le $(g+1)$ *ième* échantillonnage de θ_1, etc… Ce processus génère une chaîne de Markov (c'est la partie « Markov Chain » de l'acronyme MCMC) dont les éléments convergent en distribution vers des tirages de $P(\theta \mid y)$, la fonction de densité jointe des paramètres *a posteriori*. La convergence atteinte, la distribution de la chaîne de Markov est alors dite stationnaire sur $P(\theta \mid y)$ (TIERNEY, 1994). L'avantage de détenir *in fine* un échantillon empirique de distribution stationnaire $P(\theta \mid y)$ à la place d'une fonction de densité est qu'il permet de calculer très aisément toutes les statistiques (moments, écarts-types, quantiles,…) la caractérisant. Il n'est plus nécessaire de calculer le moment postérieur de $z(\theta)$ par[4] :

$$\tilde{H} = \int_{\theta} z(\theta) \times P(\theta \mid y)\, d\theta$$

puisque la moyenne de la fonction z des éléments de la chaîne de Markov converge (presque sûrement) vers l'espérance de la densité *a posteriori* de $z(\theta)$. C'est la partie « Monte Carlo » de MCMC) :

$$\hat{H} = \frac{1}{G}\sum_{g=1}^{G} z(\theta^g) \xrightarrow{p.s.} E\left(z(\theta)\right)$$

Sous les conditions de Lebesgue et lorsque G est grand, $\hat{H}$ est une estimation consistante de $\tilde{H}$ en vertu du théorème ergodique (STOUT (1974), TIERNEY (1994), ROBERT (1996)).

b) Une extension pratique : l'étape « augmentation data »

Il s'agit d'une étape astucieuse que l'on doit à TANNER et WONG (1987) qui consiste à augmenter l'espace des paramètres avec le vecteur de variables latentes Y^*. « *Pour les bayésiens, toutes les quantités inobservables peuvent faire l'objet d'inférence sans se soucier qu'elles sont celles de paramètres ou de variables latentes* » (ROSSI et ALLENBY, 2000, p.309). En incluant cette étape dans l'algorithme de Gibbs, on contourne l'évaluation de $\int P\left(\theta, Y^* \mid Y\right) dY^*$ pour calculer la distribution marginale *a posteriori* $P\left(\theta \mid Y\right)$. L'idée reste simple : si l'on parvient à récupérer de l'échantillonneur de Gibbs une suite de tirages $\{\theta^{(g)}, Y^{*(g)}\}$ issus de la loi jointe $P(\theta, Y^* \mid Y)$, alors il suffit de supprimer l'ensemble des $Y^{*(g)}$ échantillonnés pour marginaliser la distribution des paramètres θ. Formellement, l'échantillonneur de Gibbs avec augmentation de données devient une succession de tirages dans des distributions conditionnelles suivantes:

[4] z est un simple fonction de θ. Par exemple, pour $z(\theta) = \theta$, nous avons la formulation de l'espérance.

Etape 1 : Elle correspond aux K premières étapes de l'algorithme (40).
Pour $k = 1 \rightarrow K$,
k : θ_k^{g+1} tirés dans $P\left(\theta_k \mid \theta_{<k}^{(g+1)}, \theta_{>k}^{(g)}, Y^{*(g)}, y\right)$
Etape 2 : Augmentation de données
Pour $i = 1 \rightarrow N$,
$K + i$: $Y_i^{*(g+1)}$ tiré dans $P\left(Y_i^* \mid \theta^{(g+1)}, y\right)$,

$K + N + 1$: $g \leftarrow g + 1$. Retour en l'étape 1.

Si l'étape d'augmentation de données nous affranchi de toute intégrale ou calcul de probabilité, elle est venue agrandir l'espace d'échantillonnage, ce qui généralement s'accompagne d'une corrélation en série des tirages accrue, et par conséquence d'une convergence plus lente vers la distribution stationnaire $P(\theta \mid y)$. Pour contrôler la convergence, nous utiliserons dans ce document le diagnostique de GEWEKE (1992). Sans en détailler tous les calculs, nous explicitons dans le prochain paragraphe les distributions conditionnelles *a posteriori* des paramètres de notre modèle, nécessaires à la bonne marche de l'échantillonnage de Gibbs.

4.3 Distributions conditionnelles a posteriori du modèle

Toutes les distributions conditionnelles sont formellement obtenues à partir de la relation de proportionnalité décrite en (39). En augmentant l'espace des paramètres avec celui des variables latentes, nous pouvons réécrire cette relation de proportionnalité avec les notations de notre modèle :

$$P\left(\Theta, S, Y^*, v \mid Y, X\right) \propto L\left(Y, Y^* \mid \Theta, S, v, X\right) \times P(\Theta) \times P(S) \times P(v) \tag{41}$$

En opérant une décomposition conditionnelle sur $L\left(Y, Y^* \mid \Theta, S, v, X\right)$, l'équation précédente est équivalente à :

$$\begin{aligned} &P\left(\Theta, S, Y^*, v \mid Y, X\right) \propto \\ &\quad L\left(Y \mid Y^*, \Theta, S, v, X\right) L\left(Y^* \mid \Theta, S, v, X\right) \times P(\Theta) \times P(S) \times P(v) \end{aligned} \tag{42}$$

En vertu de la règle décisionnelle décrite en (2) et (12), les variables latentes Y^* et les seuils $S = \{ s_1, s_2, s_1', s_2' \}$ sont des éléments dont la connaissance est suffisante à déterminer sans équivoque les niveaux de motorisation Y . Par ailleurs, ces seuils n'apportent aucune information sur le niveau des variables latentes si leur vraisemblance n'est pas conditionnée par les choix Y . Ces deux remarques contribuent à simplifier la relation (42) en :

$$P\left(\Theta, S, Y^*, v \mid Y, X\right) \propto L\left(Y \mid Y^*, S\right) \times L\left(Y^* \mid \Theta, v, X\right) \times P(\Theta) \times P(S) \times P(v) , \tag{43}$$

et fait apparaître la vraisemblance $L\left(Y^* \mid \Theta, v, X\right)$ détaillée en (19).Toutes les distributions conditionnelles nécessaires à la mise en route de l'échantillonneur de Gibbs découlent de la relation précédente. Pour chaque paramètre (ou variable latente) considéré(e), ces distributions sont obtenues en réécrivant la relation (43) comme une simple fonction de celui (ou celle-ci), les autres prenant la valeur de leur dernier tirage dans la boucle d'échantillonnage. En spécifiant des distributions *a priori* « conjuguées », nous obtenons les distributions conditionnelles suivantes. Nous explicitons dans un premier temps celles des paramètres, avant de nous intéresser à celles des variables individuelles. Nous adoptons les notations supplémentaires suivantes : soit π le vecteur colonne cumulé des coefficients du modèle dynamique, $\pi = (\beta', \gamma, \delta)'$, et $\Theta(\text{-}\theta), S(\text{-}\theta)$ les ensembles de paramètres Θ et S privés du paramètre θ.

a) Distributions a posteriori des paramètres

- Distribution de $\left(\pi \mid \Theta(\text{-}\pi), S, Y^*, v, Y, X\right)$

La distribution d'une loi $\mathcal{N}\left(\underline{\pi}, \underline{V}_\pi\right)$ est utilisée comme *a priori* sur π. Notons $Y_{it}^{*\pi} = Y_{it}^* - v_i$, et $Z_{it} = \left(X_{it} \sim Y_{it-1}^* \sim \left(Y_{i1}^* - X_{i1}\beta_0\right)\right)$ pour $t > 1$, où le signe $\sim$ est ici l'opérateur de la concaténation horizontale. Avec la relation (43) écrite comme une fonction de π, nous avons après quelques manipulations d'algèbre :

$$P\left(\pi \mid \Theta(\text{-}\pi), S, Y^*, v, Y, X\right) \propto \exp\left(-0.5\,(\pi - \overline{\pi})'\ \overline{V}_\pi^{-1}\,(\pi - \overline{\pi})\right),$$

$$\text{avec } \overline{V}_\pi = \left(\frac{1}{\sigma_\varepsilon^2}\sum_i \sum_{t=2}^{3} w_i Z_{it}' Z_{it} + \underline{V}_\pi^{-1}\right)^{-1};\ \overline{\pi} = \overline{V}_\pi\left(\frac{1}{\sigma_\varepsilon^2}\sum_i \sum_{t=2}^{3} w_i Z_{it}' Z_{it} \times \hat{\pi} + \underline{V}_\pi^{-1} \times \underline{\pi}\right) \quad (44)$$

$$\text{et où} \qquad \hat{\pi} = \left(\sum_i \sum_{t=2}^{3} w_i Z_{it}' Z_{it}\right)^{-1}\left(\sum_i \sum_{t=2}^{3} w_i Z_{it}' Y_{it}^{*\pi}\right)$$

Il en découle que la distribution conditionnelle *a posteriori* de π est normale :

$$\pi \mid \Theta(\text{-}\pi), S, Y^*, v, Y, X \sim \mathcal{N}\left(\overline{\pi}, \overline{V}_\pi\right) \quad (45)$$

- Distribution de $\left(\beta_0 \mid \Theta(\text{-}\beta_0), S, Y^*, v, Y, X\right)$

Nous utilisons une distribution pour $\mathcal{N}\left(\underline{\beta}_0, \underline{V}_{\beta_0}\right)$ caractériser la connaissance *a priori* de β_0. Notons $X_{i1}^{\beta_0} = -\delta X_{i1}$; $Y_{i2}^{*\beta_0} = Y_{i2}^* - X_{i2}\beta - (\gamma + \delta)Y_{i1}^*$; $Y_{i3}^{*\beta_0} = Y_{i3}^* - X_{i3}\beta - \gamma Y_{i2}^* - \delta Y_{i1}^* - v_i$ pour alléger les prochaines écritures. La relation (43)écrite comme une fonction de β_0 donne après quelques développements :

$$P\left(\beta_0 \mid \Theta(\text{-}\beta_0), S, Y^*, v, Y, X\right) \propto \exp\left(-\frac{1}{2}\left(\beta_0 - \overline{\beta}_0\right)' \overline{V}_{\beta_0}^{-1}\left(\beta_0 - \overline{\beta}_0\right)\right) \tag{46}$$

$$\hat{\beta}_0^a = \left(\sum_i w_i X_{i1}' X_{i1}\right)^{-1}\left(\sum_i w_i X_{i1}' Y_{i1}^*\right);\ H(\hat{\beta}_0^a) = \sigma_{\varepsilon_0}^{-2}\left(\sum_i w_i X_{i1}' X_{i1}\right)$$

$$\text{Avec :}\quad \hat{\beta}_0^b = \left(\sum_i 2w_i X_{i1}^{\beta_0\prime} X_{i1}^{\beta_0}\right)^{-1}\left(\sum_i \sum_{t=2}^{3} w_i X_{i1}^{\beta_0\prime} Y_{it}^{*\beta_0}\right);\ H(\hat{\beta}_0^b) = 2\sigma_{\varepsilon}^{-2}\left(\sum_i w_i X_{i1}^{\beta_0\prime} X_{i1}^{\beta_0}\right)$$

$$\overline{V}_{\beta_0} = \left(H(\hat{\beta}_0^a) + H(\hat{\beta}_0^b) + \underline{V}_{\beta_0}^{-1}\right)^{-1}\ ;\ \overline{\beta}_0 = \overline{V}_{\beta_0}\left(H(\hat{\beta}_0^a)\hat{\beta}_0^a + H(\hat{\beta}_0^b)\hat{\beta}_0^b + \underline{V}_{\beta_0}^{-1}\underline{\beta}_0\right)$$

Par identification avec la densité de la loi normale, nous avons donc :

$$\left(\beta_0 \mid \Theta_{-\beta_0}, S, Y^*, v, Y, X\right) \sim \mathcal{N}\left(\overline{\beta}_0; \overline{V}_{\beta_0}\right)$$

- Distribution de $\left(\sigma_{\varepsilon_0}^{-2} \mid \Theta(\text{-}\sigma_{\varepsilon_0}^2), S, Y^*, v, Y, X\right)$

Posons en *a priori* que $\sigma_{\varepsilon_0}^{-2} \sim \mathcal{G}\left(p_1/2, \mathrm{T}_1/2\right)$. En écrivant la relation (43) comme une fonction de $\sigma_{\varepsilon_0}^{-2}$, nous obtenons :

$$P\left(\sigma_{\varepsilon_0}^{-2} \mid \Theta(\text{-}\sigma_{\varepsilon_0}^2), S, Y^*, v, Y, X\right) \propto \prod_i \left(\sigma_{\varepsilon_0}^{-2}\right)^{\frac{W+p_1}{2}-1} \exp\left(-\frac{\sum_i w_i\left(Y_{i1}^* - X_{i1}\beta_0\right)^2 + \mathrm{T}_1}{2}\sigma_{\varepsilon_0}^{-2}\right) \tag{47}$$

Par identification avec la fonction de densité de la loi gamma, nous avons qu'*a posteriori*, la distribution conditionnelle de $\sigma_{\varepsilon_0}^{-2}$ est de distribution :

$$\sigma_{\varepsilon_0}^{-2} \mid \Theta(\text{-}\sigma_{\varepsilon_0}^2), S, Y^*, v, Y, X \sim \Gamma\left(\frac{W+p_1}{2}, \frac{\sum_i w_i\left(Y_{i1}^* - X_{i1}\beta_0\right)^2 + \mathrm{T}_1}{2}\right) \tag{48}$$

- Distribution de $\left(\sigma_v^{-2} \mid \Theta(\text{-}\sigma_v^2), S, Y^*, v, Y, X\right)$

En choisissant une distribution $\Gamma\left(p_2/2, \mathrm{T}_2/2\right)$ comme *a priori* de σ_v^{-2}, nous écrivons la relation (43) comme une fonction de σ_v^{-2}, puis nous obtenons après simplification :

$$P\left(\sigma_v^{-2} \mid \Theta(\text{-}\sigma_v^2), S, Y^*, v, Y, X\right) \propto \left(\sigma_v^{-2}\right)^{\left(\frac{W+p_2}{2}-1\right)} \exp\left(-\frac{\sum_i w_i\left(v_i - 0\right)^2 + \mathrm{T}_2}{2}\sigma_v^{-2}\right) \tag{49}$$

Il en découle par identification avec la densité d'une loi gamma que :

$$\sigma_v^{-2} \mid \Theta(\text{-}\sigma_v^2), S, Y^*, v, Y, X \sim \Gamma\left(\frac{W + p_2}{2}, \frac{\sum_i w_i (v_i - 0)^2 + \mathrm{T}_2}{2}\right) \tag{50}$$

- Distribution de $\left(\sigma_\varepsilon^{-2} \mid \Theta(\text{-}\sigma_\varepsilon^2), S, Y^*, v, Y, X\right)$

Les conditions d'identification peuvent être interprétées comme des *certitudes*, donc non révisables *a posteriori*, traduites par des relation purement déterministes. Concernant σ_ε^{-2}, nous obtenons compte tenu de la condition (23) :

$$\left(\sigma_\varepsilon^{-2} \mid \Theta(\text{-}\sigma_\varepsilon^2), S, Y^*, v, Y, X\right) = \left(1 - \delta^2 \sigma_{\varepsilon_0}^2 - \sigma_v^2\right)^{-1} \tag{51}$$

Concernant les *a posteriori* conditionnels des seuils $S = \{ s_1, s_2, s_1', s_2' \}$, nous examinons tout d'abord ceux du modèle général $\{ s_1, s_2 \}$, avant de considérer ceux du modèle initiales $\{ s_1', s_2' \}$.

- Distribution de $\left(s_1 \mid \Theta, S(\text{-}s_1), Y^*, v, Y, X\right)$

Puisqu'il existe également une restriction d'identification portant sur s_1 en (23), alors celle-ci tient lieu d'*a posteriori* déterministe. Nous noterons toutefois qu'il s'agit d'une restriction qui ne dépend pas de la valeur d'autres paramètres :

$$\left(s_1 \mid \Theta, S(\text{-}s_1), Y^*, v, Y, X\right) = s_1 = 0 \tag{52}$$

- Distribution de $\left(s_2 \mid \Theta, S(\text{-}s_2), Y^*, v, Y, X\right)$

Le point de départ incontournable est l'écriture de la relation (43), cette fois comme une fonction de s_2. Nous choisissons un *a priori* totalement non informatif pour s_2 en lui spécifiant une distribution uniforme, c'est-à-dire que $P(s_2) \propto c \times 1(s_1 < s_2 < +\infty)$. Après quelques écritures, nous obtenons finalement :

$$s_2 \mid \Theta, S(\text{-}s_2), Y^*, v, Y, X \sim \mathcal{U}\left[\max\left(s_1; \max_{i,t \geq 2}\left(Y_{it}^* : Y_{it} = 1\right)\right); \min_{i,t \geq 2}\left(Y_{it}^* : Y_{it} = 2\right)\right] \tag{53}$$

- Distributions de $\left(s_l' \mid \Theta, S(\text{-}s_l'), Y^*, v, Y, X\right)$

Les valeurs des seuils $s' = \left\{s_1', s_2'\right\}$ de la période initiale sont dans notre modélisation soumises à des restrictions d'identification. Il découle notamment des restrictions d'identification (23) que :

$$\left(s_1' \mid \Theta, S(\text{-}s_1'), Y^*, v, Y, X\right) = s_1 = 0 \quad ; \quad \left(s_2' \mid \Theta, S(\text{-}s_2'), Y^*, v, Y, X\right) = s_2 \qquad (54)$$

Après avoir considéré les *a posteriori* des paramètres et seuils, nous devons maintenant détailler celles des variables propres à chaque ménage i, c'est-à-dire leur facteur d'hétérogénéité v_i et leurs variables latentes Y_1^*, Y_2^*, Y_3^*.

b) Distributions a posteriori des variables individuelles

- **Les variables latentes**

On explicite les distributions conditionnelles des variables latentes Y_{it}^* de la même façon que celle des paramètres précédents, c'est-à-dire que l'on commence par écrire la relation (43) comme une fonction de Y_{it}^*. Pour tout i et $\forall t$ de 1 à 3, nous aboutissons à cette expression :

$$P\left(Y_{it}^* \mid \Theta, S, Y_{i-t}^*, v_i, Y_i, X_i\right) \propto L\left(Y_{it} \mid Y_{it}^*, S\right) L\left(Y_{it}^* \mid \Theta, Y_{i,-t}^*, v_i, X_i\right), \qquad (55)$$

ou $Y_{i,-t}^*$ correspond au vecteur Y_i^* amputé de Y_{it}^*. Dans la relation (55), le résultat de $L\left(Y_{it} \mid Y_{it}^*, S\right)$ est une variable binaire qui vaut 1 si l'association $\left\{Y_{it}^*, S\right\}$ engendre effectivement la décision Y_{it} (0 sinon) en vertu de la règle de choix décrite des systèmes (2) et (12). Formellement, il s'agit du résultat suivant :

$$L\left(Y_{it} \mid Y_{it}^*, S\right) = 1\left(a_{it} < Y_{it}^* < b_{it}\right) \qquad (56)$$

avec a_{it} et b_{it} définis en (22), et où 1(arg) désigne la fonction indicatrice valant 1 si l'argument est vrai, 0 sinon. Pour notre usage, $L\left(Y_{it} \mid Y_{it}^*, S\right)$ va seulement nous servir à poser les troncatures de distributions conditionnelles des variables latentes Y_{it}^*.

- Distribution de $\left(Y_{i1}^* \mid \Theta, S, Y_{i-1}^*, v_i, Y, X\right)$

Notons : $Y_{i3}^{*1} = Y_{i3}^* - X_{i3}\beta - \gamma Y_{i2}^* + \delta\beta_0 X_{i1} - v_i$, $Y_{i2}^{*1} = Y_{i2}^* - X_{i2}\beta + \delta\beta_0 X_{i1} - v_i$. En développant le terme $L\left(Y_{i1}^* \mid \Theta, Y_{i-1}^*, v_i, X_i\right)$ en (55), nous aboutissons à :

$$Y_{i1}^* \mid \Theta, Y_{i,-1}^*, v_i, X_i \sim \mathcal{N}\left(\overline{Y}_{i1}^*, \overline{V}_{Y_1^*}\right) \qquad (57)$$

avec :

$$\hat{Y}_{i1}^{*a} = X_{i1}\beta_0 \ ; \ \hat{Y}_{i1}^{*b} = (\gamma+\delta)^{-1} Y_{i2}^{*1} \ ; \ \hat{Y}_{i1}^{*c} = \delta^{-1} Y_{i3}^{*1}$$
$$H\left(\hat{Y}_{i1}^{*a}\right) = \sigma_{\varepsilon_0}^{-2} \ ; \ H\left(\hat{Y}_{i1}^{*b}\right) = (\gamma+\delta)^2 \sigma_{\varepsilon}^{-2} \ ; \ H\left(\hat{Y}_{i1}^{*c}\right) = \delta^2 \sigma_{\varepsilon}^{-2}$$
$$\overline{V}_{Y_1^*} = \left(H(\hat{Y}_{i1}^{*a}) + H(\hat{Y}_{i1}^{*b}) + H(\hat{Y}_{i1}^{*c})\right)^{-1}$$
$$\overline{Y}_{i1}^* = \overline{V}_{Y_1^*}\left(H(\hat{Y}_{i1}^{*a})\hat{Y}_{i1}^{*a} + H(\hat{Y}_{i1}^{*b})\hat{Y}_{i1}^{*b} + H(\hat{Y}_{i1}^{*c})\hat{Y}_{i1}^{*b}\right)$$

De (56) et (57) , il vient que : $Y_{i1}^* \mid \Theta, S, Y_{i-1}^*, v_i, Y_i, X_i \sim 1\left(a_{i1} < Y_{i1}^* < b_{i1}\right) \times \mathcal{N}\left(\overline{Y}_{i1}^*, \overline{V}_{Y_1^*}\right)$. Autrement dit :

$$Y_{i1}^{*} \mid \Theta, S, Y_{i-1}^{*}, v_i, Y_i, X_i \sim \underset{[a_{i1}, b_{i1}]}{\mathcal{NT}}\left(\overline{Y}_{i1}^{*}, \overline{V}_{Y_1^{*}}\right) \tag{58}$$

- Distribution de $\left(Y_{i2}^{*} \mid \Theta, S, Y_{i,-2}^{*}, v_i, Y, X\right)$

Définissons $Y_{i2}^{*2} = Z_{i2}\pi + v_i$, $Y_{i3}^{*2} = Y_{i3}^{*} - X_{i3}\beta - \delta\left(Y_{i1}^{*} - X_{i1}\beta_0\right) - v_i$. En détaillant le terme $L\left(Y_{i2}^{*} \mid \Theta, Y_{i-2}^{*}, v_i, X_i\right)$ de (55), nous observons que :

$$Y_{i2}^{*} \mid \Theta, Y_{i,-2}^{*}, v_i, X_i \sim \mathcal{N}\left(\overline{Y}_{i2}^{*}, \overline{V}_{Y_{i2}^{*}}\right) \tag{59}$$

avec : $\overline{V}_{Y_{i2}^{*}} = \left(\left(1+\gamma^2\right)\sigma_{\varepsilon}^{-2}\right)^{-1}$; $\overline{Y}_{i2}^{*} = \overline{V}_{Y_{i2}^{*}}\left(\sigma_{\varepsilon}^{-2}Y_{i2}^{*2} + \gamma^2\sigma_{\varepsilon}^{-2}\left(Y_{i3}^{*2}/\gamma\right)\right)$

Sachant (56) et (59), nous obtenons la relation de proportionnalité $P\left(Y_{i2}^{*} \mid \Theta, S, Y_{i,-2}^{*}, v_i, Y_i, X_i\right) \propto 1\left(a_{i2} < Y_{i2}^{*} < b_{i2}\right) \times \mathcal{N}\left(\overline{Y}_{i2}^{*}, \overline{V}_{Y_{i2}^{*}}\right)$, qui nous autorise à conclure que la distribution conditionnelle de Y_{i2}^{*} est normale tronquée :

$$Y_{i2}^{*} \mid \Theta, s, Y_{i-2}^{*}, v_i, Y_i, X_i \sim \underset{[a_{i2}, b_{i2}]}{\mathcal{NT}}\left(\overline{Y}_{i2}^{*}, \overline{V}_{Y_{i2}^{*}}\right) \tag{60}$$

- Distribution de $\left(Y_{i3}^{*} \mid \Theta, S, Y_{i-3}^{*}, v_i, Y, X\right)$

La relation (55) écrite pour la dernière période $t = 3$ nous donne $P\left(Y_{i3}^{*} \mid \Theta, S, Y_{i-3}^{*}, v_i, Y_i, X_i\right) \propto L\left(Y_{i3} \mid Y_{i3}^{*}, S\right) L\left(Y_{i3}^{*} \mid \Theta, Y_{i-3}^{*}, v_i, X_i\right)$. En posant $\overline{Y}_{i3}^{*} = Z_{i3}\pi + v_i$, nous obtenons notamment :

$$P\left(Y_{i3}^{*} \mid \Theta, S, Y_{i-3}^{*}, v_i, Y_i, X_i\right) \propto 1\left(a_{i3} < Y_{i3}^{*} < b_{i3}\right) \times \mathcal{N}\left(\overline{Y}_{i3}^{*}, \sigma_{\varepsilon}^{2}\right)$$

La distribution conditionnelle de Y_{i3}^{*} est normale tronquée :

$$Y_{i3}^{*} \mid \Theta, S, Y_{i-3}^{*}, v_i, Y, X \sim \underset{[a_{i3}, b_{i3}]}{\mathcal{NT}}\left(\overline{Y}_{i3}^{*}, \sigma_{\varepsilon}^{2}\right)$$

- Les termes d'hétérogénéité

- Distribution de $\left(v_i \mid \Theta, S, Y_i^{*}, Y_i, X_i\right)$

Notons $Y_{it}^{*v} = Y_{it}^{*} - Z_{it}\pi$ pour $t \geq 2$. L'écriture de la relation (43) comme seule fonction de v_i nous donne après développements :

$$P\left(v_i \mid \Theta, S, Y_i^{*}, Y_i, X_i\right) \propto \exp\left(-0.5\left(v_i - \overline{v}_i\right)\overline{V}_v^{-1}\left(v_i - \overline{v}_i\right)\right)$$

Avec : $\overline{V}_v^{-1} = 2\sigma_\varepsilon^{-2} + \sigma_v^{-2}$; $\overline{v}_i = \overline{V}_v^{-1}\left(2\sigma_\varepsilon^{-2}\left(\frac{1}{2}\sum_{t=2}^{3} Y_{it}^{*v}\right)\right)$; $Y_{it}^{*v} = Y_{it}^{*} - Z_{it}\pi$

Il apparaît donc que : $v_i \mid \Theta, S, Y_i^*, Y_i, X_i \sim N\left(\overline{v}_i, \overline{V}_v\right)$ (61)

5 Sélection des variables et statistiques descriptives

Nous présentons quelques statistiques descriptives du panel pondéré Parc auto 1999-2001, qui font état de corrélations entre la motorisation des ménages et un ensemble de leurs caractéristiques. Le codage et le libellé des variables retenues pour notre étude, ainsi que l'information statistique synthétisée par des couples moyenne-variance nous sont donnés dans la table 2.1.

TABLE 2.1 : Variables d'intérêt, libellés, et statistiques descriptives (1/2)

Variable	**Libellé**	**Moyenne**	**(Ec-type)**
nbvoi	nombre de voitures à la disposition du ménage	1.186	(0.813)
ageche	âge du chef de famille	52.003	(16.483)
Nombre dans le ménage de (CDF excepté)...			
nbactoc	actifs occupés	0.394	(0.533)
nbretrai	retraités	0.130	(0.341)
nbfem	femmes adultes	0.983	(0.479)
Nombre d'adultes âgés (CDF excepté)...			
nbad1840	de 18 à 40 ans (hors enfants majeurs)	0.265	(0.443)
nage4160	de 41 à 60 ans	0.242	(0.428)
nage6170	de 61 à 70 ans	0.108	(0.313)
nage71p	de plus de 70 ans	0.068	(0.251)
Nombre d'enfants âgés...			
nage05	de moins de 6ans	0.150	(0.441)
nage611	de 6 à 11 ans	0.171	(0.473)
nage1217	de 12 à 17 ans	0.181	(0.504)
enfmaj	de 18 à 30 ans	0.181	(0.505)
Indicatrices : le chef de famille est...			
chefact	Actif occupé par un emploi	0.593	(0.491)
chefret	retraité	0.350	(0.477)
chefchom	Actif sans emploi	0.057	(0.232)

Notes : écart types entre parenthèse

Source : Panel pondéré 1999-2001 de Parc Auto

TABLE 2.1 : Variables d'intérêt, libellés, et statistiques descriptives (2/2)

Variable	Libellé	Moyenne	(Ec-type)
Indicatrices : le revenu annuel net du ménage est...			
r12	inférieur à 75KF	0.155	(0.362)
rev3	compris dans [75KF, 100KF[	0.125	(0.331)
rev4	compris dans [100KF, 125KF[	0.144	(0.351)
rev5	compris dans [125KF, 150KF[	0.134	(0.340)
rev6	compris dans [150KF, 175KF[	0.102	(0.303)
rev7	compris dans [175KF, 200KF[	0.098	(0.297)
r89	compris dans [200KF, 300KF[	0.174	(0.379)
r1013	supérieur ou égal à 300KF	0.067	(0.251)
Indicatrices: le ménage réside...			
paris	à Paris	0.050	(0.218)
pcour	en petite couronne francilienne	0.060	(0.237)
gcour	en grande couronne francilienne	0.068	(0.252)
llm	à Lille Lyon ou Marseille	0.056	(0.229)
ruraux	en province, en zone rurale	0.031	(0.173)
periur	en province, en zone périurbaine	0.330	(0.470)
banl	en province, en banlieue	0.150	(0.358)
centre	en province, en centre ville	0.254	(0.435)
Indicatrices : le ménage compte...			
perm0	aucun permis de conduire	0.091	(0.288)
perm1	un permis de conduire	0.357	(0.479)
perm2	Deux permis de conduire	0.473	(0.499)
perm3p	Trois permis de conduire ou plus	0.079	(0.269)

Notes : écart types entre parenthèse

Source : Panel pondéré 1999-2001 de Parc Auto

6 Résultats d'estimation

La mise en pratique de l'échantillonnage de Gibbs réclame de fixer les paramètres des distributions *a priori* de la section 4.3. Ils ont été choisis de façon telle que les *a priori* soient très diffus. Les valeurs suivantes ont notamment été utilisées : $\underline{\theta} = \underline{\beta}_0 = 0$; $\underline{V}_\theta = 100 \times I_{35}$; $\underline{V}_{\beta_0} = 100 \times I_{33}$; $p_1 = p_2 = 2$; $\mathrm{T}_1 = \mathrm{T}_2 = 0.5$.

Puis, l'algorithme d'échantillonnage doit être calé sur des valeurs d'initialisation. Nous avons estimé par maximum de vraisemblance les paramètres d'un modèle probit ordonné statique sur la vague d'enquête Parc Auto 1999, et avons utilisé les valeurs obtenues pour initialiser les paramètres β, β_0 et s_2 dans notre boucle d'échantillonnage de Gibbs (table

2.12). Les paramètres dynamiques γ , δ ont été fixés à 0.1, alors que les termes des variances initialement choisis sont : $\sigma^2_{\varepsilon 0} = 1$, $\sigma^2_v = 0.1$. Les variables individuelles ont été initialisées en $v_i = 0$, $Y_{i1} = X_{i1}\beta_0$ et $Y_{it} = X_{it}\beta + \gamma Y_{it-1}$ pour $t > 1$.

TABLE 2.12: Initialisation des paramètres β, β_0, s_2

Variable	**coefficient**	**Variable**	**coefficient**	**Variable**	**coefficient**
constante	-4.767	**nage71p**	0.062	**gcour**	1.257
ageche	0.085	**nage05**	0.210	**llm**	1.181
ageche²/100	-0.084	**nage611**	-0.090	**ruraux**	2.391
chefret	-0.004	**nage1217**	0.075	**periur**	1.865
chefchom	-0.521	**rev3**	0.420	**banl**	1.723
nbactoc	0.179	**rev4**	0.566	**centre**	1.424
nbretrai	0.067	**rev5**	0.650	**perm1**	1.742
nbfem	-0.177	**rev6**	0.727	**perm2**	2.845
enfmaj	0.098	**rev7**	0.833	**perm3p**	3.679
nbad1840	0.317	**r89**	0.923		
nage4160	0.067	**r1013**	1.021	s_2	2.547
nage6170	0.042	**pcour**	0.683		

Notes : Lexique des variables en table 2.1. Estimation du modèle probit ordonné statique.
Source : Vague 1999 du panel pondéré Parc Auto 1999-2001

Ainsi amorcé, notre algorithme a effectué 400000 cycles d'échantillonnage, dont nous avons éliminé les 100000 premiers qui correspondent à une phase de convergence. Parmi les 300000 derniers, nous avons conservé les paramètres avec une fréquence de un tous les dix cycles, puisque des valeurs successives échantillonnées présentaient une très forte corrélation en série. Au final, nous disposons pour chaque paramètre d'une chaîne de 30000 tirages, dont nous avons tout d'abord contrôlé les propriétés de convergence. Notamment, nous avons appliqué le diagnostique de GEWEKE (1992), dont la statistique associée $|CD|$ est reportée. Pour chaque paramètre scalaire, celle ci est inférieure à 1.96, suggérant que les chaînes ont individuellement toutes convergé. Nous pouvons donc considérer que nos tirages de paramètres sont effectivement issus de leur distribution a posteriori, et les exploiter.

La table 2.13 présente les résultats d'estimation du modèle dynamique (5). La première colonne montre les moyennes marginales empiriques a posteriori, la seconde présente les intervalles de crédibilité à 95%. La troisième donne les résultats de la statistique $|CD|$. Les mêmes statistiques pour le modèle initial (11) sont disponibles sur demande.

Le modèle observe un pic de la motorisation pour les ménages dont le chef de famille est âgé de 37 ans. Toutes choses égales, notamment le revenu, le modèle montre que les ménages sont moins motorisés lorsque le chef est sans emploi que lorsqu'il est actif. Par contre, les estimations ne montrent pas de différence significative en comparant un ménage dont le chef est actif avec un autre retraité.

Concernant la structure du ménage autour de son chef, le modèle mesure que l'effet d'un adulte supplémentaire sur sa motorisation est dégressif avec l'âge. Notamment, l'ajout d'un adulte âgé de 18 à 40 ans augmente significativement la motorisation du ménage, l'effet d'un adulte supplémentaire âgé de 41 à 60 ans est moindre, mais n'est plus significatif qu'au seuil de crédibilité à 90%. Au-delà de 61 ans, le modèle ne distingue plus d'effet notable. Que l'adulte supplémentaire soit actif, retraité ou sans emploi, ce détail n'influe pas significativement sur la motorisation du ménage.

Au contraire, le genre des adultes est une variable déterminante : relativement à un homme, l'effet d'une femme est une diminution significative de la motorisation du ménage. Par ailleurs, les résultats montrent qu'à toutes choses égales et quelque soit leur âge, le nombre d'enfants n'influence pas significativement le niveau de motorisation des ménages.

Le revenu annuel se révèle un déterminant décisif du niveau de motorisation. En comparaison avec un ménage doté d'un budget inférieur à 75KF/an, les ménages plus riches sont significativement plus motorisés. Fort logiquement, le modèle constate un effet positif et croissant à mesure que le revenu augmente.

Concernant les disparités géographiques, le modèle estime que les ménages ruraux sont significativement plus motorisés que les urbains résidents des centres villes ou des banlieues. De façon plus générale, les résultats suggèrent que les niveaux de motorisation des ménages sont décroissants avec la densité de population de leur zone de résidence. C'est ainsi qu'en Ile de France et toute chose égale, un ménage parisien est statistiquement moins motorisé qu'un ménage francilien de petite couronne, à son tour moins motorisé qu'un ménage francilien de petite couronne. Ceci traduit également qu'une diminution de la densité de circulation ainsi qu'un moins bonne couverture de transport collectif rendent plus attractive la possession automobile. Notamment en province, les coefficients de la table 2.13 montrent qu'en moyenne, les résidents des métropoles Lille-Lyon-Marseille (dotés d'un métro notamment) sont les moins motorisés.

Mais surtout, le nombre de permis de conduire détenus par les ménages est le facteur le plus explicatif de leur motorisation. Plus particulièrement, le modèle estime que les passages d'aucun permis à un permis, d'un permis à deux permis, et de deux permis à trois permis s'accompagnent toujours d'une augmentation significative de leur motorisation. On remarque également une augmentation dégressive du coefficient avec le nombre de permis de conduire, pouvant suggérer que l'équipement d'un ménage augmente moins vite que son nombre de permis. Ce constat n'est pas totalement surprenant puisque l'automobile est un bien dont on peut partager l'usage à plusieurs permis au sein du ménage.

Le coefficient de l'endogène retardée γ est positif et très significatif, traduisant un phénomène de dépendance temporelle de la motorisation. Explicitement, un ménage aura d'autant plus de chances d'être motorisé (resp. non motorisé) à la période t qu'il était motorisé (resp. non motorisé) en $t-1$. La motorisation passée tend donc à retenir dans son état le choix de motorisation courant des ménages. Cet « effet mémoire » révèle la rigidité des ménages à ajuster leur niveau d'équipement automobile à ses déterminants courants, potentiellement due au recours à l'habitude pour le déterminer.

Enfin, la part de variance totale de l'erreur du modèle qui est imputable au terme d'hétérogénéité individuelle est estimé à $\hat{\sigma}_{\alpha}^{2}/\sigma_{\xi}^{2} = (1-\hat{\sigma}_{\varepsilon}^{2})/1 = 37\%$. Près d'un tiers du comportement non expliqué par les régresseurs utilisés tient donc, en France, à des spécificités individuelles plutôt qu'à des causes accidentelles.

TABLE 2.13: Résultats du modèle probit ordonné dynamique (1/2)

Coefficients	**Moyenne**	**Int. de crédibilité**	$\|CD\|$
constante	-2.511	[-3.331 ;-1.708]	0.79
Variable d'âge du chef de ménage			
ageche	0.025	[0.001 ; 0.051]	0.81
ageche²/100	-0.034	[-0.058 ;-0.009]	0.78
Indicatrices du statut d'activité du chef de ménage (Référence : chefact)			
chefret	0.082	[-0.094 ; 0.259]	0.66
chefchom	-0.311	[-0.564 ;-0.065]	0.67
Structure et composantes du ménage			
nbactoc	0.099	[-0.023 ; 0.222]	0.54
nbretrai	-0.018	[-0.185 ; 0.146]	1.27
nbfem	-0.141	[-0.265 ;-0.018]	0.59
enfmaj	0.099	[-0.044 ; 0.240]	0.93
nbad1840	0.289	[0.089 ; 0.487]	0.96
nage4160	0.155	[-0.019 ; 0.331]	0.43
nage6170	0.124	[-0.077 ; 0.325]	0.82
nage71p	0.105	[-0.129 ; 0.338]	0.60
nage05	0.061	[-0.057 ; 0.180]	0.39
nage611	-0.033	[-0.132 ; 0.067]	0.06
nage1217	0.002	[-0.088 ; 0.091]	0.27
Indicatrices de la fourchette de revenu du ménage (Référence : r12)			
rev3	0.257	[0.076 ; 0.436]	0.46
rev4	0.388	[0.211 ; 0.566]	0.79
rev5	0.515	[0.330 ; 0.701]	0.86
rev6	0.637	[0.433 ; 0.845]	0.76
rev7	0.650	[0.448 ; 0.852]	0.86
r89	0.739	[0.540 ; 0.940]	0.84
r1013	0.942	[0.701 ; 1.187]	0.89

Notes : Lexique des variables en table 2.1. Moyennes marginales empiriques a posteriori, intervalles de crédibilité à 95%, et statistique $|CD|$ de Geweke.

Source : Panel pondéré 1999-2001 de Parc Auto

TABLE 2.13: Résultats du modèle probit ordonné dynamique (2/2)

Coefficients	Moyenne	Int. de crédibilité	$\|CD\|$
Indicatrices de localisation du ménage (Référence : paris)			
pcour	0.477	[0.195 ; 0.760]	0.29
gcour	1.085	[0.793 ; 1.383]	0.45
llm	0.914	[0.616 ; 1.221]	0.61
ruraux	1.970	[1.569 ; 2.388]	0.68
periur	1.501	[1.207 ; 1.809]	0.61
banl	1.285	[0.996 ; 1.581]	0.65
centre	1.132	[0.867 ; 1.406]	0.61
Indicatrices du nombre de permis dans le ménage (Référence : perm0)			
perm1	1.341	[1.088 ; 1.601]	0.77
perm2	2.126	[1.823 ; 2.435]	0.79
perm3p	3.082	[2.681 ; 3.498]	0.78
Coefficients dynamiques			
γ	0.516	[0.458 ; 0.570]	0.75
δ	0.229	[0.183 ; 0.275]	0.83
Seuils			
s_1	0	.	.
s_2	4.067	[3.920 ; 4.205]	0.77
Variances			
σ_v^2	0.234	[0.105 ; 0.411]	0.74
$\sigma_{\varepsilon 0}^2$	2.280	[2.012 ; 2.571]	0.75
σ_ε^2	0.646	[0.465 ; 0.787]	0.73

Notes : Lexique des variables en table 2.1. Moyennes marginales empiriques a posteriori, intervalles de crédibilité à 95%, et statistique $|CD|$ de Geweke.

Source : Panel pondéré 1999-2001 de Parc Auto

Dans sa globalité, le modèle révèle de bonnes propriétés d'ajustement puisque la seule connaissance des variables explicatives X_i lui permet de prévoir correctement 74,5% des observations en 2001. Dans le détail, le modèle prédit respectivement 65%, 83% et 68% des cas de ménages non, mono et multi motorisés en 2001.

7 Exploitation des résultats

7.1 Temps d'ajustement des comportements de motorisation

Nous présentons tout d'abord quelques statistiques issues de la distribution marginale a posteriori du coefficient γ. La proportion en $t+s$ de l'effet de long terme latent induit par un changement permanent de caractéristiques à partir de t se calcule par $(1-\gamma)\sum_{h=0}^{s}\gamma^h$. En moyenne, le modèle estime ainsi que la moitié (48.4%) de l'effet de long terme est pris en compte par les ménages dès le court terme dans l'ajustement de leur comportement d'équipement. Au bout d'un an, les ¾ de cet effet de long terme sont déjà absorbés, et il est assimilé à plus de 95% par les ménages dans leur décision de motorisation après quatre années (table 2.14).

TABLE 2.14: Proportion de l'effet de long terme d'un changement permanent considérée par les ménages

s	**C**	**[95%]**
0	48.4%	[43.0 ; 54.2]
1	73.4%	[67.5 ; 79.0]
2	86.3%	[81.5 ; 90.4]
3	92.9%	[89.4 ; 95.6]
4	96.3%	[94.0 ; 98.0]
5	98.1%	[96.6 ; 99.1]
6	99.0%	[98.0 ; 99.6]

Lecture : la part moyenne de l'effet à long terme d'un changement permanent de caractéristiques pris en considération par les ménages après s années est de C% . Intervalles de crédibilité à 95% entre crochets. Source : Estimation du probit ordonné dynamique.

Par ailleurs, la statistique γ^s mesure le taux de survie après s périodes de l'effet initial latent d'un changement de caractéristiques accidentel[5]. Le modèle estime donc en moyenne que la survie de l'effet initial n'est plus que d'environ 25% deux ans après le changement, et qu'il aura disparu à 95% entre la quatrième et la cinquième année après celle du changement (table 2.15).

[5] c'est à dire non reproduit les périodes suivantes.

TABLE 2.15: Taux de survie de l'effet initial d'un changement de caractéristiques occasionnel

s	S	[95%]
0	100%	.
1	51.6%	[45.8 ; 57.0]
2	26.6%	[21.0 ; 32.5]
3	13.7%	[9.6 ; 18.5]
4	7.1%	[4.4 ; 10.6]
5	3.7%	[2.0 ; 6.0]
9	0.3%	[0.1 ; 0.6]

Lecture : la part moyenne de l'effet d'un changement de caractéristiques accidentel encore pris en considération par les ménages après s années est de S% . Intervalles de crédibilité à 95% entre crochets. Source : Estimation du probit ordonné dynamique.

7.2 L'effet du changement résidentiel

Les propriétés dynamiques de notre modèle sont ensuites mises en avant par la simulation des effets de court et long terme d'un changement résidentiel des ménages. Pour cela, nous avons définis des ménages synthétiques calculés aux caractéristiques moyennes des zones géographiques, pour lesquels nous avons individuellement simulé un déménagement vers une zone francilienne (Paris, petite ou grande couronne). On considère que ces ménages moyens ont des caractéristiques stables dans le temps, si bien que la distribution de leur variable latente est à l'équilibre. Les probabilités de motorisation associées découle donc du système (34) écrit pour $X_{i,\leq t} = \overline{X}_z$, où $\overline{X}_z$ est le vecteur de caractéristiques du ménages synthétique de la zone z. Nous obtenons les systèmes de probabilités en table 2.16 suivants. En Ile de France, le ménage synthétique parisien présente logiquement la plus forte probabilité d'être non équipé (0.64), alors que les ménages moyens de petite et de grande couronne montrent une forte propension au mono équipement. Chez ce dernier, le non équipement devient presque nul, alors que la multi motorisation apparaît statistiquement. En province non urbaine, la non motorisation est peu répandue dans comportement d'équipement des ménages synthétiques. Le ménage moyen des agglomérations de Lille, Lyon ou Marseille et celui des centres-villes révèlent les plus fortes propensions au non, et mono équipement automobile. A mesure que l'on s'éloigne de ces secteurs urbains, en passant successivement par les zones de banlieue urbaine, périurbaine, puis rurale, la probabilité de mono motorisation diminue au profit de la probabilité de multi équipement, et lui devient même inférieure pour le ménage rural.

TABLE 2.16: Probabilités de motorisation des ménages synthétiques

	Etat de motorisation		
Zone résidentielle	**non**	**mono**	**multi**
Paris	0.64 [0.52 ; 0.74]	0.36 [0.26 ; 0.47]	0.00 [0.00 ; 0.00]
Petite couronne francilienne	0.24 [0.17 ; 0.32]	0.73 [0.66 ; 0.79]	0.03 [0.01 ; 0.05]
Grande couronne francilienne	0.05 [0.03 ; 0.08]	0.79 [0.74 ; 0.82]	0.16 [0.11 ; 0.22]
Lille, Lyon, Marseille	0.15 [0.10 ; 0.21]	0.79 [0.75 ; 0.83]	0.06 [0.03 ; 0.09]
Rurale	0.00 [0.00 ; 0.01]	0.48 [0.34 ; 0.61]	0.52 [0.38 ; 0.66]
Périurbain	0.01 [0.01 ; 0.01]	0.60 [0.56 ; 0.64]	0.39 [0.35 ; 0.43]
Banlieue	0.03 [0.02 ; 0.04]	0.74 [0.70 ; 0.78]	0.23 [0.19 ; 0.28]
Centre ville	0.10 [0.08 ; 0.13]	0.81 [0.79 ; 0.83]	0.08 [0.06 ; 0.11]

Notes : Lexique des variables en table 2.1. Ménages synthétiques évalués aux caractéristiques moyennes de leur zone résidentielle. Intervalle de crédibilité à 95% entre crochets.

Source : Panel pondéré 1999-2001 de Parc Auto.

Sur la base des statistiques de la table 2.16, nous simulons pour les ménages synthétiques un changement permanent de résidence vers la région Ile de France. Instantanément pour chacun, les multiplicateurs de court terme en (29) modifient la variable latente, et donne lieu à un nouveau système de probabilité évalué via les formules (35). Individuellement, le changement résidentiel permanent n'atteint son plein effet que dans le long terme, et fait intervenir ses multiplicateurs (32) pour déterminer la valeur stabilisée de leur variable latente. Sous l'hypothèse du gel des autres facteurs, on utilise les formules (37) qui calculent les probabilités de motorisation de long terme induit par le déménagement. Par différence de ses deux systèmes de probabilité avec le système initial synthétisé dans la table 2.16, nous obtenons les effets marginaux de court et long terme (formules (36) et (38)). Ils sont reportés dans la table 2.17 pour les changements résidentiels « intra » Ile de France, puis dans la table 2.18 pour les déménagements de la province vers une zone francilienne.

TABLE 2.17: Effets marginaux de court et long terme d'un déménagement intra Ile de France

			Destination					
			paris		**pcour**		**gcour**	
			CT	**LT**	**CT**	**LT**	**CT**	**LT**
Origine	**paris**	**non**			-0.12 [-0.19 ;-0.05]	-0.25 [-0.38 ;-0.11]	-0.27 [-0.34 ;-0.20]	-0.50 [-0.61 ;-0.38]
		mono			0.12 [0.05 ; 0.18]	0.24 [0.10 ; 0.37]	0.26 [0.19 ; 0.34]	0.44 [0.32 ; 0.54]
		multi			0.00 [0.00 ; 0.00]	0.01 [0.00 ; 0.02]	0.01 [0.01 ; 0.02]	0.06 [0.04 ; 0.10]
	pcour	**non**	0.10 [0.04 ; 0.17]	0.23 [0.10 ; 0.37]			-0.11 [-0.16 ;-0.06]	-0.18 [-0.26 ;-0.10]
		mono	-0.09 [-0.14 ;-0.04]	-0.21 [-0.33 ;-0.09]			0.07 [0.03 ; 0.12]	0.07 [0.02 ; 0.14]
		multi	-0.01 [-0.03 ; 0.00]	-0.02 [-0.04 ;-0.01]			0.04 [0.02 ; 0.05]	0.11 [0.06 ; 0.16]
	gcour	**non**	0.12 [0.08 ; 0.17]	0.37 [0.26 ; 0.49]	0.05 [0.03 ; 0.08]	0.15 [0.09 ; 0.23]		
		mono	-0.01 [-0.06 ; 0.05]	-0.22 [-0.33 ;-0.11]	0.02 [0.00 ; 0.06]	-0.03 [-0.09 ; 0.03]		
		multi	-0.12 [-0.16 ;-0.08]	-0.16 [-0.21 ;-0.11]	-0.08 [-0.12 ;-0.04]	-0.13 [-0.19 ;-0.07]		

Notes : Lexique des variables en table 2.1. Effets évalués sur les probabilités d'équipement des ménages synthétiques. CT et LT pour court et long terme. Intervalles de crédibilité à 95% entre crochets. Source : Panel pondéré 1999-2001 de Parc Auto.

Selon la table 2.17, le ménage synthétique parisien qui viendrait s'excentrer en Ile de France vers ses couronnes ressent une très faible amélioration de sa multi motorisation, dont la probabilité estimé ne « plafonne » qu'à 6% en grande couronne. Au contraire, sa propension à se mono motoriser augmente franchement au détriment de la non motorisation : elle croit en moyenne de +0.12 dans le court terme si sa destination est en petite couronne, et de +0.26 si elle est en grande couronne. A long terme, ces mesures respectives sont de +0.24 et +0.44. En observant le chemin inverse, le ménage de petite couronne renonce presque sûrement à la multi motorisation. Il diminue par ailleurs sa probabilité de mono équipement de 0.09 en moyenne dans le court terme et 0.21 dans le long terme s'il s'installe à Paris. Lorsqu'il se « recentre », le ménage de grande couronne francilienne renonce rapidement à la multi motorisation. A court terme, cela profite tout d'abord à sa mono motorisation s'il déménage en petite couronne, puis exclusivement à sa non motorisation dans le long terme. S'il déménage à Paris, alors il augmentera sa probabilité de non équipement dès le court terme de 12 points (37 à long terme). D'une façon générale, l'hétérogénéité des ménages synthétiques parisiens montre que les effets marginaux ne sont pas symétriques lorsque l'on compare le déménagement d'un ménage z vers une zone z', et celui d'un ménage z' vers une zone z.

De la table 2.18 découlent les constatations suivantes. Quelque soit l'horizon temporel, le ménage synthétique des métropoles Lille, Lyon et Marseille conserve sensiblement le même comportement de motorisation lorsque l'on simule son déménagement en grande couronne francilienne.

TABLE 2.18 : Effets marginaux de court et long terme d'un déménagement vers l'Ile de France

			Destination					
			paris		**pcour**		**gcour**	
			CT	**LT**	**CT**	**CT**	**LT**	**CT**
Origine	**llm**	**non**	0.18 [0.12 ; 0.24]	0.42 [0.29 ; 0.55]	0.07 [0.03 ; 0.11]	0.17 [0.08 ; 0.28]	-0.02 [-0.06 ; 0.01]	-0.05 [-0.12 ; 0.02]
		mono	-0.13 [-0.19 ;-0.08]	-0.37 [-0.48 ;-0.25]	-0.05 [-0.08 ;-0.02]	-0.13 [-0.22 ;-0.06]	0.01 [0.00 ; 0.04]	0.02 [-0.01 ; 0.05]
		multi	-0.04 [-0.07 ;-0.02]	-0.05 [-0.09 ;-0.03]	-0.03 [-0.05 ;-0.01]	-0.04 [-0.07 ;-0.02]	0.01 [-0.01 ; 0.03]	0.03 [-0.01 ; 0.08]
	ruraux	**non**	0.08 [0.04 ; 0.12]	0.48 [0.36 ; 0.60]	0.04 [0.02 ; 0.06]	0.24 [0.17 ; 0.33]	0.01 [0.01 ; 0.02]	0.06 [0.04 ; 0.10]
		mono	0.33 [0.20 ; 0.45]	0.03 [-0.15 ; 0.22]	0.30 [0.19 ; 0.40]	0.25 [0.10 ; 0.40]	0.20 [0.12 ; 0.29]	0.33 [0.19 ; 0.46]
		multi	-0.41 [-0.52 ;-0.29]	-0.51 [-0.65 ;-0.37]	-0.34 [-0.44 ;-0.24]	-0.49 [-0.63 ;-0.35]	-0.22 [-0.30 ;-0.13]	-0.39 [-0.54 ;-0.24]
	periur	**non**	0.07 [0.05 ; 0.10]	0.35 [0.24 ; 0.47]	0.04 [0.02 ; 0.05]	0.15 [0.10 ; 0.22]	0.01 [0.00 ; 0.01]	0.03 [0.01 ; 0.05]
		mono	0.21 [0.17 ; 0.25]	0.03 [-0.09 ; 0.13]	0.18 [0.14 ; 0.22]	0.19 [0.13 ; 0.24]	0.09 [0.05 ; 0.13]	0.16 [0.10 ; 0.22]
		multi	-0.29 [-0.33 ;-0.24]	-0.38 [-0.42 ;-0.34]	-0.22 [-0.26 ;-0.17]	-0.34 [-0.39 ;-0.30]	-0.10 [-0.14 ;-0.06]	-0.19 [-0.26 ;-0.11]
	banl	**non**	0.12 [0.08 ; 0.16]	0.40 [0.29 ; 0.52]	0.06 [0.04 ; 0.08]	0.18 [0.12 ; 0.25]	0.01 [0.00 ; 0.02]	0.02 [0.00 ; 0.05]
		mono	0.06 [0.01 ; 0.10]	-0.18 [-0.30 ;-0.07]	0.07 [0.04 ; 0.10]	0.01 [-0.05 ; 0.07]	0.03 [0.00 ; 0.06]	0.05 [0.00 ; 0.09]
		multi	-0.17 [-0.21 ;-0.13]	-0.22 [-0.27 ;-0.18]	-0.13 [-0.16 ;-0.09]	-0.19 [-0.24 ;-0.15]	-0.04 [-0.07 ; 0.00]	-0.07 [-0.14 ; 0.00]
	centre	**non**	0.20 [0.14 ; 0.26]	0.50 [0.38 ; 0.61]	0.10 [0.07 ; 0.13]	0.25 [0.16 ; 0.34]	0.01 [-0.02 ; 0.03]	0.01 [-0.03 ; 0.06]
		mono	-0.13 [-0.19 ;-0.08]	-0.41 [-0.53 ;-0.30]	-0.05 [-0.08 ;-0.03]	-0.18 [-0.26 ;-0.10]	0.00 [-0.01 ; 0.00]	0.00 [-0.02 ; 0.00]
		multi	-0.07 [-0.08 ;-0.05]	-0.08 [-0.10 ;-0.06]	-0.05 [-0.07 ;-0.03]	-0.07 [-0.09 ;-0.05]	0.00 [-0.02 ; 0.01]	-0.01 [-0.04 ; 0.03]

Notes : Lexique des variables en table 2.1. Effets évalués sur les probabilités d'équipement des ménages synthétiques. CT et LT pour court et long terme. Intervalles de crédibilité à 95% entre crochets. Source : Panel pondéré 1999-2001 de Parc Auto.

Par ailleurs, on remarque que le ménage rural, pourtant le plus enclin à la multi motorisation dans sa zone d'origine, consent à abandonner dès le court terme une très forte propension à être multi équipé, en venant résider en Ile de France. Dans les couronnes, il tend à devenir essentiellement mono motorisé, et le même constat s'impose pour le ménage synthétique périurbain. Ce dernier montre des effets marginaux à court et long terme très voisins lorsqu'il s'installe en petite couronne. Lorsqu'on le déplace en grande couronne, le ménage moyen de la zone banlieue tend à favoriser le mono équipement et ne recours pas à la non motorisation. Excepté le ménage moyen des centres-villes et celui des métropoles Lille Lyon et Marseille, les autres ménages provinciaux dont on a simulé le déplacement à Paris augmentent dans le court terme leur probabilité d'être mono équipé au détriment du multi équipement, puis la diminue dans le long terme au profit du non équipement automobile. Dans ces cas de figure, l'état de mono équipement tient lieu d'étape intermédiaire entre multi et non motorisation.

7.3 Les élasticités au revenu

Comme HANLY et DARGAY (2000), nous souhaitons mettre en lumière les effets de revenu sur le niveau d'équipement des ménages, tout en tirant profit de notre spécification dynamique par le calcul d'élasticités de court et long terme. A partir de WINSTON (1981) notamment, les élasticités revenu de l'équipement automobile sont définies par les variations relatives des probabilités moyennes de choisir chaque état de motorisation, rapportées au taux de croissance du revenu. Avec nos notations, nous définissons donc les élasticités-revenu de court et long terme, $\varsigma_R^{CT}(j)$ et $\varsigma_R^{LT}(j)$, suivantes :

$$\varsigma_R^{CT}(j) = \frac{\frac{1}{W}\sum_i w_i \Lambda_{i,j}^{CT}(rR_{it})}{\frac{1}{W}\sum_i w_i \Pr[Y_{it} = j \mid X_{i,\leq t}]} \times \frac{1}{r} \tag{62}$$

$$\varsigma_R^{LT}(j) = \frac{\frac{1}{W}\sum_i w_i \Lambda_{i,j}^{LT}(rR_{it})}{\frac{1}{W}\sum_i w_i \Pr[Y_{it} = j \mid X_{i,\leq t}]} \times \frac{1}{r} \tag{63}$$

avec r le taux de croissance de la variable continue de revenu R. Malheureusement, notre information sur les revenus des ménages n'est constituée que d'indicatrices de tranches, et non de R directement. Pour calculer les élasticités (62) et (63), nous devons donc reconstituer les revenus. Pour cet exercice, nous utilisons une méthode d'imputation simple, en tirant individuellement les revenus dans une loi uniforme, dont l'intervalle de définition correspond aux bornes de la tranche de revenu déclarée. Les revenus simulés sont ensuite augmentés de 20% ($r = 0.2$), puis sont à nouveau discrétisées dans les tranches de départ. Les statistiques

(62) et (63) sont calculées aux caractéristiques des ménages en 2001, et sont reportées dans la table 2.19.

TABLE 2.19: Elasticités au revenu des probabilités d'équipement automobile

	Non équipement	**Mono équipement**	**Multi équipement**
Court terme	-0.19 [-0.23 ;-0.15]	-0.05 [-0.07 ;-0.02]	0.18 [0.13 ; 0.23]
Long terme	-0.38 [-0.45 ; -0.31]	-0.11 [-0.16 ;-0.05]	0.37 [0.27 ; 0.46]

Notes : Calculs des formules (62) et (63) au caractéristiques X des ménages en 2001. Intervalles de crédibilité à 95% entre crochets.

Nos résultats d'estimation faisaient apparaître que la croissance du revenu des ménages profitait à l'augmentation de leur motorisation automobile. A court ou à long terme, il en découle que l'élasticité au revenu de la probabilité de non équipement est négative, et réciproquement, que celle de l'équipement est positive. L'automobile appartient donc, en France, à la catégorie des biens micro économiques dits « normaux ». Plus précisément, les sensibilités à court terme des probabilités de non et multi motorisation sont de signe opposés comme attendu, et d'amplitude très similaires (respectivement de -0.19 et +0.18, table 2.19). Que celle de la mono probabilité soit peu affectée par le revenu à court terme n'est pas surprenant puisqu'il s'agit de la modalité d'équipement la plus répandue dans la population. A long terme, les effets revenus sont plus élevés et correspondent assez précisément au double des élasticités de court terme. Ce rapport tient largement au fait que le coefficient γ est proche de 0.5, et en conséquence, que les multiplicateurs latents de long terme sont globalement deux fois plus élevés que ceux de court terme.

Conclusion

Ce document a traité du comportement quantitatif d'équipement automobile des ménages en France. Dans un environnement statistique bayésien, le modèle probit ordonné a été ajusté sur les observations du panel pondéré Parc Auto 1999-2001, pour caractériser trois états : la non motorisation, la mono motorisation, et la multi motorisation. Grâce à la méthode d'estimation MCMC par échantillonnages de Gibbs, qui reconstitue notamment les variables latentes par augmentation de données, il nous a été possible de faire reposer notre modèle sur une forme latente autorégressive. Pour en approximer les conditions initiales, nous avons inspiré notre méthode de l'estimateur linéaire dynamique de Blundell et Smith. Originale dans ce contexte, cette forme autorégressive du modèle linéaire latent se démarque des études précédentes qui, au mieux, introduisaient la dynamique dans leur modèle latent par des

indicatrices du niveau d'équipement passé. En comparaison, notre spécification permet d'adapter les outils d'analyse des séries temporelles à notre problème de choix catégoriel du niveau de motorisation, élargissant ainsi l'éventail de conclusions dynamique.

Si, comme les études précédentes, notre modèle conclue également à la dynamique des comportements de motorisation automobile, l'inférence statistiques est enrichie par l'analyse de court et long terme. Par exemple, notre modèle mesure que les ménages confrontés à un changement permanent de caractéristiques considèrent dès à court terme environ la moitié de son effet total de long terme. Par ailleurs, le calcul des effets marginaux simulant un déménagement des ménages moyens franciliens vers une autre zone francilienne montrent que le ménage parisien augmente sa probabilité d'équipement (initialement de 36%) de 27 points à court terme, puis de 50 points à long terme lorsqu'il s'implante dans la grande couronne d'Ile de France. Inversement, ménage moyen de grande couronne francilienne qui s'installe a paris diminue en moyenne sa probabilité d'équipement de 12 points a court terme, et 37 points a long terme. D'une façon générale et quelque soit l'horizon, les effets marginaux ne sont pas symétriques lorsque l'on observe deux ménages franciliens moyens, l'un déménageant dans la zone de l'autre et inversement, la raison tenant à leur hétérogénéité observée. Enfin, les résultats ne décèlent aucune modification significative à court ou long terme du comportement d'équipement lorsque le ménage moyen des agglomérations Lille-Lyon-Marseille et celui des centres-villes s'établissent en grande couronne francilienne.

La reconstitution du revenu annuel des ménages puis la simulation de sa croissance a permis de dériver les élasticités au revenu à court et long terme des probabilités d'équipement. Il est apparu que l'automobile est un bien normal, et que les probabilités de motorisation extrêmes avaient à court terme des sensibilités au revenu opposées et d'amplitude équivalente. Ce constat tient également pour les sensibilités de long terme, qui sont globalement deux fois plus fortes que celles de court terme.

Outre le revenu et la zone de résidence, les résultats de la modélisation mettent également en évidence le rôle prédominant du nombre de permis de conduire dans le ménage pour expliquer son niveau d'équipement.

La qualité de la motorisation des ménages n'est pas traitée dans cette étude. Pourtant, on pourrait, en parallèle avec une saturation des besoins en automobiles, soupçonner son effet lorsqu'on observe la croissance globalement concave des coefficients de revenus (table 2.13). Associé aux élasticités-revenu inférieures à un[6] (table 2.19), ceci suggère que le niveau de motorisation augmente moins vite que le revenu, et que les ménages le compensent par exemple en améliorant la qualité de leur parc. Intuitivement, il n'y a aucune difficulté à présumer que les augmentations de revenu des ménages s'accompagne d'une augmentation qualitative de leur voitures lorsque le nombre d'automobile devient égale celui de permis de conduire. Un prolongement nécessaire pour comprendre le comportement d'équipement automobile des ménages réside donc dans la prise en compte que l'automobile est un bien très différencié (COLLET, 2006).

[6] en valeur absolue.

Bibliographie

ALBERT, J. and S. CHIB, 1993, Bayesian analysis of binary and polychotomous data, *Journal of the American Statistical Association*, 88, 669-679.

BESAG, J. E., 1974, Spatial interaction and the statistical analysis of lattice systems, *Journal of the Royal Statistical Society*, Series B, 36, pp. 192–225.

BLUNDELL, R. W. and R. J. SMITH, 1991, Conditions initiales et estimation efficace dans les modèles dynamiques sur données de panel : une application au comportement d'investissement des entreprises, *Annales d'Economie et de Statistique*, 20-21, pp. 109-123.

BROOKS, S.P. and G.O. ROBERTS, 1998, Convergence assessment techniques for Markov chain Monte Carlo, *Statistics and Computing*, 8, pp. 319-335.

BUTLER, J. S., and R. MOFFIT, 1982, A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model, *Econometrica*, 50, pp. 761–764.

CHIB, S., 1995, Inference in panel data models via Gibbs sampling, in : P. Sevestre and L. Matyas ed., *The Econometrics of Panel Data : A Handbook of the Theory with Applications*, 2nd revised edition, Kluwer Academic Publishers, London, Great Britain.

COLLET, R., 2006, Etude du choix d'acquisition d'automobiles, *23ème Journée de Microéconomie Appliquées*, Nantes, France, 2 Juin.

COWLES, M. K. and CARLIN, B. P.,1996, Markov Chain Monte Carlo convergence diagnostics: a comparative review, *Journal of the American Statistical Association*, 91, pp. 883--904.

DARGAY, J., L. HIVERT and D. LEGROS, 2006, An investigation of car ownership in Europe based on the European Community Household Panel, *11th International Conference on Travel Behaviour Research*, Kyoto, Japan, August 16-20.

GELFAND, A.E. and A.F.M. SMITH, 1990, Sampling based approaches to calculating marginal densities, *Journal of the American Statistical Association*, 85, pp. 398-409.

GEMAN, S. and D. GEMAN, 1984, Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images, *IEEE Transactions on pattern analysis and machine intelligence*, 6, 721-741.

GEWEKE, J., 1992, Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, In J. M. Bernardo, A. F. M. Smith, A. P. Dawid and J. O. Berger eds., *Bayesian Statistics* 4, pp. 169-193, Oxford University Press.

GOODWIN, P.B. and M.J.H. MOGRIDGE, 1981, Hypotheses for a fully dynamic model of car ownership, *International Journal of Transport Economics*, 8, 313-327.

GOURIÉROUX, C. and A. Monfort, 1994, *Simulation Based Econometric Methods*, Core lectures series, Oxford University Press.

GREENE, W. H., 2003, *Econometric Analysis*, 5th edition, Prentice Hall, New York, USA.

HANLY, M., and J. DARGAY, 2000, Car ownership in Great Britain : Panel data analysis, *Transportation Research Record*, 1718, 83-89.

HECKMAN, J.J., 1981, Heterogeneity and state dependence, in Sherwin Rosen ed., *Studies In Labour Markets*, University of Chicago Press, 91-139.

HIVERT, L., 2001, *Le parc automobile des ménages : Etude en fin d'année 1999 à partir de la source PARC AUTO - SOFRES*, Rapport de convention INRETS-ADEME, Arcueil, France.

INRETS (Institut National de Recherche sur les Transports et leur Sécurité), 2001, Le Parc Automobile 2000 - Module Ménage, Arcueil, France.

INRETS (Institut National de Recherche sur les Transports et leur Sécurité), 2002, Le Parc Automobile 2001 - Module Ménage, Arcueil, France.

JACKMAN, S., 2000, Estimation and Inference via Bayesian Simulation: an introduction to Markov Chain Monte Carlo, *American Journal of Political Science*. 44-2, pp. 375-404.

KITAMURA, R and, D. BUNCH, 1990, Heterogeneity and state dependence in household car ownership: A panel analysis using ordered-response probit models with error components, in : M. Koshi ed., *Transportation And Traffic Theory*, Elsevier Science Publishing Co., Amsterdam, Netherlands, 476-496.

MCCULLOCH R. and P. ROSSI, 1994, An exact likelihood analysis of the multinomial probit model, *Journal of Econometrics*, 64, 207-240.

MEURS, H., 1990, Trip generation models with permanent unobserved effects, *Transportation Research*, 24B, 145-158.

NOBILE, A., C.R. BHAT and PAS E.I., 1997, A Random Effects Multinomial Probit Model of Car Ownership Choice, *Case Studies in Bayesian Statistics 3*, eds. C. Gatsonis, et al., pp. 419 - 434, Springer, New York, USA.

PAAP R. and FRANSES P.H., 2000, A Dynamic Multinomial Probit Model for Brand Choice with Different Long-run and Short-run Effects of Marketing-Mix Variables, *Journal of Applied Econometrics*, 15-6, pp. 717-744.

POWERS, D. A., and Y. XIE, 2000, *Statistical Methods for Categorical Data Analysis*, Academic Press, San Diego, USA.

RIPLEY, B.D., 1987, *Stochastic simulation*, J. Wiley and Sons, New York, USA.

ROBERT, C., 1996, *Méthodes de Monte Carlo par Chaines de Markov*, Economica, Paris, France.

ROSSI, P. E. and G. M. ALLENBY, 2003, Bayesian Statistics and Marketing, *Marketing Science*, 22, pp. 304-328.

SEVESTRE, P., 2003, *Econométrie des Données de Panel*, Dunod, Paris, France.

STOUT, W. F., 1974, *Almost sure convergence*, Academic Press, New York, USA.

TANNER, M.A., 1991, *Tools for Statistical Inference*, Springer-Verlag, New York, USA.

TANNER, M.A. and W.H WONG, 1987, The calculation of posterior distributions by data augmentation , *Journal of American Statistical Association*, 82, 528-540.

TIERNEY L., 1994, Markov chains for exploring posterior distributions, *Annals of statistics*, 22, 1701-1762.

UNCLES, M. D., 1987, A beta logisticmodel of mode choice : Goodness of fit and intertemporal dependence, *Transportation Research*, 21B, 195-205.

WINSTON, C.,1981, A disaggregate model of the demand for intercity freight transportation, *Econometrica*, 49, 981-1006.