

**COMMENT FAIRE CONVERGER LES
APPROCHES ECONOMIQUE ET
DEMOGRAPHIQUE DE LA DEMANDE DE
TRANSPORT ?**

G. BRESSON (*), *J-L. MADRE* (**), *A. PIROTTE* (***)

(*)ERMES (CNRS) Université Paris II

(**)DEST, INRETS

(***)GATE (CNRS) Université Lyon II

1. Introduction

L'élaboration des Plans de Déplacements Urbains (PDU), actuellement en cours dans toutes les agglomérations de plus de 100 000 habitants suite à la Loi sur l'Air et l'Utilisation Rationnelle de l'Energie, montre que les citoyens éprouvent une certaine inquiétude au sujet de l'avenir de la mobilité, ayant souvent l'impression que l'on est au bord de l'asphyxie. En raison des nuisances qu'elle génère (bruit, pollutions locales et globale,...) l'automobile ne pourra pas répondre seule à la demande de mobilité et les PDU font une large place aux modes de déplacement alternatifs. Pour le principal d'entre-eux -le transport public- cette communication vise à quantifier les principaux facteurs d'évolution de la demande: l'offre bien sûr (en quantité et en qualité), les déterminants économiques (revenus et prix), mais aussi les facteurs structurels (motorisation croissante, étalement urbain, vieillissement de la population,...).

Afin de répondre à ces besoins de planification, il faut se doter de méthodes adaptées pour cerner les évolutions à moyen et long terme. Dans ce domaine, la démographie et l'économétrie nous semblent fournir à la fois les outils adéquats et des résultats fiables sur lesquels on pourra bâtir une démarche prospective.

Depuis une dizaine d'années, nous avons développé une approche démographique de la motorisation des ménages (*Berri, Gardes, Madre et Starzec (1998)*), de l'usage de l'automobile (*Gallez et Madre (1993)*) et de la mobilité quotidienne en ville (*Boulahbal et Madre (1999)*, *Bussière, Armoogum et Madre (1995)*). Il s'agit pour l'essentiel du suivi du comportement des générations successives et de l'ajustement de modèles Age-Cohorte, où l'on peut introduire quelques variables économiques *via* les effets de périodes. Parallèlement, nous avons développé des travaux économétriques sur séries temporelles nationales et sur un panel régional afin de faire ressortir les facteurs économiques (revenu et prix) et les facteurs d'offre (parc automobile, réseaux d'infrastructure) expliquant l'évolution de la circulation des véhicules légers sur le réseau national (*Bresson, Madre et Pirotte (1997)*, *Madre et Pirotte (1997)*).

Tous ces travaux répondent à des demandes de projections à long terme. Est-il possible de faire converger les approches démographique et économétrique de la demande de transport dans un même modèle susceptible d'explicitier simultanément les facteurs structurels (vieillesse de la population, étalement urbain, motorisation croissante), les facteurs économiques (revenu et prix) et les facteurs d'offre (en quantité comme en qualité)? C'est la question que nous allons traiter en nous appuyant sur l'exemple des transports collectifs en province (recherche menée en partenariat avec le transporteur VIA-GTI).

L'objectif est de construire un modèle de tendance à long terme des trafics en Transports Collectifs Urbains (TCU) pour différents réseaux de Province (appelés PTU: Périmètres de Transports Urbains). Le modèle doit

s'articuler sur des données annuelles de Transports Collectifs Urbains (fournies par le CERTU) ainsi que sur des données communales de recensement de la population (données INSEE). Ce modèle ne peut pas correspondre aux spécifications habituelles, basées sur une représentation fine des réseaux de transports et des flux de déplacements¹. En effet, du fait du détail — dans la dimension temporelle et/ou spatiale — de la représentation sur laquelle reposent ces modèles, ils ne peuvent pas s'appliquer à des analyses rétrospectives sur longue période. Mais la modélisation envisagée doit cependant répondre à plusieurs exigences:

- On veut à la fois évaluer les effets de court terme et de long terme d'un certain nombre de variables explicatives (revenu, offre de transports, prix des transports, ...) sur les trafics TCU pour chaque PTU.
- On souhaite tenir compte de l'effet de l'évolution de la structure de la population sur les trafics TCU puisque les enquêtes de mobilité montrent une forte sensibilité de la mobilité TCU à l'âge, au sexe, à la localisation résidentielle et à la motorisation du ménage.
- Comme le nombre de d'observations relatives à chaque PTU est faible (21 observations par variable) et que le nombre de variables susceptibles d'influencer dans le temps les trafics TCU peut être important, il est statistiquement peu fiable de construire un modèle spécifique à chaque PTU. A l'aide des techniques économétriques des données de panel, on peut, en revanche, bénéficier de l'information relative à tous les PTU simultanément afin de construire un modèle plus général et statistiquement fiable.

La prise en compte explicite de l'hétérogénéité des comportements en matière de transports, *i.e* l'estimation d'élasticités différentes pour chaque PTU, nous obligent à recourir à des techniques économétriques bayésiennes qui font partie de la classe générale des modèles à coefficients aléatoires.

Ce papier se décompose en deux parties. Dans la première partie, nous soulignons l'intérêt mais également les limites de la prise en compte des facteurs structurels démographiques pour appréhender la demande à long terme de transports publics. Puis dans une seconde partie, nous proposons d'estimer un modèle de demande de transport qui tient compte à la fois des comportements démographiques *via* l'utilisation d'une variable de "mobilité structurelle" et des comportements économiques habituellement testés (revenus, prix et offre de transport).

2. L'approche démographique

2.1 La démographie: une clé pour appréhender le long terme

Les comportements de transport évoluent au cours du cycle de vie. Par exemple, un ménage acquiert sa première voiture quand il est jeune, puis éventuellement une seconde pour l'épouse ou pour les enfants quand ils atteignent leur majorité. Le mouvement de démotorisation commence quand les grands enfants quittent le

¹ tels que les modèles de coût généralisé, les modèles séquentiels ou simultanés de choix modal ou d'itinéraire (Quinet (1990)).

ménage de leurs parents; il s'achève quand les personnes âgées ne disposent plus de la capacité physique de conduire ou quand les moyens manquent pour entretenir un véhicule. Comme l'essor de l'automobile remonte aux années 60 en France, les générations successives y ont eu un accès inégal. La démographie, c'est aussi la répartition de la population sur le territoire. Or les formes urbaines marquent les comportements: l'usage des transports en commun se développe à l'intérieur des villes-centres, mais diminue là où il était le plus fort c'est-à-dire sur les liaisons radiales (*Madre et Maffre (1996)*). Le mouvement d'étalement urbain est général: les habitants des zones denses recherchent plus d'espace et de proximité avec la nature. Mais la forme urbaine n'explique pas tout; la structure de la population et le niveau de développement économique constituent également des déterminants importants de la demande de transport. Les retraités se déplacent moins souvent et moins loin que les actifs; ils utilisent de moins en moins les transports en commun car les nouvelles générations de personnes âgées ont été motorisées et gardent leur voiture jusqu'à un âge avancé (*Madre (1999)*). La population vieillit, mais ce mouvement n'est pas uniforme: il correspond au passage aux différents âges de la vie des générations pleines du baby-boom. La banlieue qui jusqu'à présent était constituée de populations plus jeunes que le centre-ville va connaître un faible renouvellement de sa population ce qui entraînera au cours de la prochaine décennie un vieillissement accéléré. Pour l'instant, le retour des personnes âgées vers un logement plus central, donc plus proche des commerces et des services, ne s'est pas vraiment manifesté dans les grandes villes mais plutôt dans le milieu rural et seulement au moment du veuvage (*Cribier (1997)*). Les différentes facettes de la démographie fournissent donc de bonnes clés pour décrire et expliquer les différentes composantes de la demande de transport: équipement des ménages en automobile et comportements de mobilité. Elle peut également fournir les données de base et les méthodes pour se tourner vers l'avenir. La prospective s'est trop souvent appuyée sur l'extrapolation des tendances observées, quitte à les ralentir un peu quand on aboutit à des situations peu vraisemblables dans le long terme. Les schémas explicatifs manquent généralement. Or, la démographie décrit des évolutions à long terme qui sont rarement uniformes et qui connaissent parfois des inversions de tendance. Ainsi, l'accélération de la croissance de la population mondiale au milieu du XXème siècle n'est plus considérée comme un processus cumulatif explosif menant tout droit à la famine et à la guerre nucléaire, mais comme une transition entre un régime de forte natalité et de forte mortalité vers un autre régime stable de faible natalité et de faible mortalité. Par analogie, *Wiel (1998)* décrit l'étalement urbain comme une transition entre la ville dense de jadis et la ville desserrée du futur.

2.2 Les limites de cette approche ou comment introduire les facteurs structurels dans un modèle économétrique?

Classiquement, on met en évidence l'effet des facteurs démographiques en estimant un modèle Age-Cohorte

(Kessier et Masson (1985)). Dans ce cadre, les facteurs économiques et d'offre, qui connaissent souvent des évolutions beaucoup plus rapides que celles de la structure de la population, sont considérés comme des "effets de période". Malheureusement, le volume des données disponibles est généralement insuffisant pour estimer un modèle Age-Cohorte-Période. Même quand on dispose d'enquêtes permanentes sur une vingtaine d'années, on peut difficilement introduire plus de deux facteurs (revenu et prix (Gallez (1994))). Et les villes où il existe plus de trois enquêtes échelonnées dans le temps sont exceptionnelles.

La principale limite des travaux de projection que nous avons réalisés sur des bases démographiques est qu'ils proposent un scénario unique d'évolution à long terme: ils donnent à tort l'impression d'un avenir inéluctable et ne mettent pas en lumière les "variables de commande" qui permettraient d'influer sur ce futur. C'est maintenant dans cette direction que nous allons nous engager en développant l'analyse des facteurs économiques (revenus et prix) et des effets de l'offre. Malgré la complexité relative du modèle Age-Cohorte, la liste des facteurs pris en compte reste limitée: l'âge (principalement ses composantes de cycle de vie et de génération), le sexe, la zone de résidence et la motorisation. D'autres variables sont largement déterminées par celles qui sont au coeur du modèle. Citons l'exemple du niveau d'instruction, principal moteur des progrès de productivité donc de l'élévation des revenus à long terme; il est lié à la génération puisque, pour un individu donné, il n'évolue plus guère après 25 ans. Cependant, toute l'influence du revenu ne peut se résumer à l'élévation des niveaux d'étude observée quand on va vers les jeunes générations. L'activité professionnelle figure aussi parmi les déterminants importants de la mobilité. La régularité du comportement de chaque génération en donne une première approche, mais le modèle pourrait aussi fournir un cadre pour des simulations testant l'influence d'un prolongement de la vie active ou de modifications dans l'activité professionnelle des femmes par exemple.

Une approche purement économique (reposant sur l'offre, les prix et le revenu) ne serait pas non plus satisfaisante dans un domaine où les phénomènes de structure sont importants:

- l'étalement urbain favorise l'usage de l'automobile au détriment des transports publics,
- et, contrairement à ce qu'indiquerait une analyse en coupe, les nouvelles générations de retraités gardent leur voiture et utilisent beaucoup moins les transports collectifs que leurs aînés.

L'introduction du parc automobile comme variable explicative des trafics a été une première tentative pour faire converger les deux approches. En effet, la motorisation des ménages est le principal facteur structurant de leurs comportements de transport; et nous avons développé, dans le cadre du modèle Age-Cohorte-Période, une analyse tout à la fois économique et démographique de son développement, qui sert de base à nos exercices de projection. S'il était encore besoin de preuves sur l'importance des effets d'offre, la comparaison de l'usage des

transports collectifs en périphérie entre Paris et Montréal montre bien qu'un réseau ferré performant constitue un atout important pour les transports collectifs, même dans des zones peu denses qui leur sont traditionnellement défavorables (*Boulaïhal et Madre* (1999)). Par ailleurs, il apparaît sur la base d'enquêtes permanentes que les revenus et, dans une moindre mesure les prix, ont une influence sur l'évolution du parc automobile et de la circulation. Toutefois, l'effet de ces facteurs ne peut pas être estimé dans le cadre d'un modèle âge-période-cohorte quand on dispose d'observations trop rares dans le temps; par exemple, les villes où il existe plus de trois enquêtes comparables sont exceptionnelles: c'est certainement le cas de Montréal où nos estimations s'appuient sur cinq enquêtes, mais nous n'avons pu utiliser les sources antérieures à 1975 ni à Grenoble, ni à Sao Paulo, ni à Paris (âge codé en classes, champ différent,...) (*Bussière, Armoogum et Madre* (1995)).

3. Convergence des approches démographique et économique

3.1 De la notion de mobilité "structurelle"

Comme on ne peut pas se contenter d'enquêtes ponctuelles pour estimer les effets d'offre, nous allons proposer une méthode économétrique calée sur des données de panel. Pour une soixantaine de villes, nous disposons de séries annuelles sur la période 1975-95 pour mesurer la demande de transport public dont nous cherchons à comprendre l'évolution, ainsi que les différentes composantes de l'offre (en quantité et en qualités) et les facteurs économiques (revenus et prix).

Comment introduire dans une équation les facteurs structurels dont le rôle important est mis en évidence par l'approche démographique? Les recensements donnent une photographie de la structure de la population à intervalles plus ou moins réguliers (tous les 5 ans au Canada, tous les 6 à 9 ans et bientôt de façon permanente en France,...). Afin de résumer cette structure (croisement du sexe, de l'âge, de la zone de résidence et de la motorisation) par un indicateur unique, on pondère le nombre d'individus appartenant à chacune de ses catégories par une mesure de sa demande de transport (ici le nombre de déplacements quotidiens en transports collectifs par personne) estimée en exploitant une enquête lourde (ici l'Enquête nationale INSEE-INRETS Transports).

L'indicateur est calculé sur la base de deux recensements successifs en gardant le même jeu de pondérations; on mesure ainsi l'évolution de la demande imputable aux transformations structurelles de la population: par exemple, de plus en plus de personnes motorisées et/ou âgées et/ou habitant en périphérie. Comme ces phénomènes sont tendanciels, on peut se contenter d'interpoler la série de "demande structurelle" entre deux recensements. Plus précisément, cet indicateur est obtenu en appliquant les rapports de mobilité des Trans-

ports Collectifs Urbains entre catégories de population — observés sur l'Enquête Nationale Transports en 1993 — aux données de population des recensements de 1975, 1982 et 1990 ainsi qu'aux projections de population pour 1995 (modèle Omphale de l'INSEE) recalées sur la base des premiers résultats du recensement de 1999. Ainsi, on considère que les rapports de mobilité sont constants entre les catégories de population. Pour construire cet indicateur, on suppose donc que les niveaux de mobilité, associés à chacune des catégories de population, évoluent parallèlement.

On obtient ainsi une seule variable explicative supplémentaire qui rend compte à la fois de plusieurs dimensions de structure de la population, tout en leur attribuant une pondération adaptée au type de demande étudiée. En agrégeant ces différentes dimensions de structure de la population en une seule variable, on satisfait au principe de parcimonie, tout en contrôlant les interactions entre ces facteurs et en limitant les difficultés résultant de la collinéarité entre des variables explicatives trop nombreuses.

La mobilité "structurelle" à l'intérieur de chaque PTU est donc estimée à champ géographique constant. Elle retrace l'évolution de la structure de la population. Les catégories de population sont construites sur la base du critère de l'âge, du sexe, de la motorisation du ménage, de la zone de résidence à l'intérieur du PTU et de la taille du bassin de vie. La mobilité "structurelle" est donc homogène à un trafic TCU.

3.2 ... à l'estimation de la demande de transport public

3.2.1 Le choix des spécifications

L'objet de cette étude est d'estimer les élasticités de court terme et de long terme des trafics TCU des réseaux de Province par rapport à un certain nombre de variables de revenu, de prix, d'offre et de caractéristiques démographiques.

Pour cela, nous avons choisi un modèle d'ajustement partiel (ou encore modèle à correction d'erreur) reliant la mobilité effective à :

- la quantité d'offre (places-kilomètres offertes par tête, pko), mais aussi différentes composantes de sa qualité (fréquence, densité de lignes, vitesse commerciale);
- les prix des modes de transport en concurrence (celui des TCU, mais aussi celui de l'essence afin de cerner la partie la plus sensible du coût d'usage de l'automobile);
- la structure de la population desservie en terme d'âge, de sexe, de localisation et surtout de motorisation (là encore pour tester l'effet de concurrence).
- le revenu des ménages,

Soit $Y_{i,t}$, la variable expliquée (la mobilité effective) du PTU i ($i = 1, \dots, N$) à la date t ($t = 1, \dots, T$) et soit $X_{1,i,t}, X_{2,i,t}, \dots, X_{k,i,t}$, les k variables explicatives du même PTU i ($i = 1, \dots, N$) à la date t ($t = 1, \dots, T$). La

spécification choisie est:

$$\log Y_{i,t} = \beta_{0,i} + \alpha_i \log Y_{i,t-1} + \sum_{j=1}^k \beta_{j,i} \log X_{j,i,t} + \varepsilon_{i,t} \quad (1)$$

où $Y_{i,t-1}$ est la mobilité effective observée à la période précédente dans le PTU i et $\varepsilon_{i,t}$, une perturbation aléatoire. Ce modèle traduit l'impossibilité de réaliser instantanément l'ajustement. Notre problème consiste donc à estimer la relation (1) à partir des observations issues de la base de données, à déterminer les estimateurs des élasticités de court terme puis à dériver les élasticités de long terme et leurs intervalles de confiance.

Nous avons volontairement indiqué les coefficients β_i pour préciser que l'on souhaite obtenir des estimations des élasticités pour chaque PTU. Afin de tenir compte de l'hétérogénéité des comportements en matière de transports, les estimations ne sont pas faites pour chaque PTU indépendamment des autres PTU. La prise en compte — et la mesure — de l'hétérogénéité nous obligent à recourir à des techniques économétriques bayésiennes. Les méthodes que nous développons font partie de la classe générale des modèles à coefficients aléatoires².

3.2.2 La méthodologie retenue

Une des hypothèses de cette étude est qu'il peut exister des situations dans lesquelles des changements structurels impliquent que les paramètres des variables explicatives peuvent être différents selon les PTU (politiques locales différentes, mécanismes d'incitation aux modes de transports en commun, ...). Afin de tenir compte de ces différences structurelles inobservables, il est alors nécessaire de décomposer les coefficients des variables explicatives.

Brièvement, on peut dire que même si la théorie économique ne précise pas les formes fonctionnelles exactes des relations, la plupart des méthodes économétriques font cependant des hypothèses spécifiques sur ces formes. La pratique économétrique usuelle de l'additivité des perturbations aléatoires à une relation mathématique n'est pas toujours correcte (Klein (1989)). Au contraire, dans un modèle à coefficients aléatoires, on suppose que chaque coefficient est la somme de deux éléments: l'effet direct d'une variable explicative sur la variable expliquée et les effets indirects (ou biais de variables omises) dus au fait que la variable explicative affecte les variables exclues et celles-ci affectent en retour la variable explicative. Dans cette classe de modèles, l'effet direct est volontairement dissocié des effets indirects. Chaque coefficient a deux composantes: une composante

² Voir annexe.

fixe et une composante stochastique³. Si l'on reprend la spécification (1), cela signifie que:

$$\log Y_{i,t} = (\beta_0 + \nu_{0,i}) + (\alpha + \nu_i) \log Y_{i,t-1} + \sum_{j=1}^k (\beta_j + \nu_{j,i}) \log X_{j,i,t} + \varepsilon_{i,t} \quad (2)$$

C'est-à-dire que tout coefficient $\beta_{j,i}$ associé à la variable j relative au PTU i est la somme d'une composante fixe β_j associée à la variable j et d'une composante stochastique $\nu_{j,i}$:

$$\beta_{j,i} = (\beta_j + \nu_{j,i}) \quad (3)$$

Une première raison d'utiliser cette approche est, qu'en cas d'erreur de spécification (*ie* variables omises, forme fonctionnelle incorrecte, ...), il n'est pas raisonnable de supposer que le "bruit" affecte uniquement le terme d'erreur. Il peut contaminer tous les coefficients de l'équation. Une seconde raison est qu'il n'existe pas, à notre connaissance, d'autre méthode permettant d'évaluer numériquement l'hétérogénéité des comportements par l'intermédiaire des coefficients des variables explicatives (autres que la constante).

Afin d'estimer ce type de spécification, on a recours à des estimateurs bayésiens itératifs appelés "estimateurs à rétrécisseur" ("*shrinkage estimators*"). Cette notion d'estimateur à rétrécisseur est souvent utilisée dans la littérature relative à la prévision parce que les prédictions effectuées, avec ce type d'estimateur, sont bien meilleures que celles réalisées avec d'autres estimateurs⁴.

La spécification que l'on souhaite estimer est donc de la forme:

$$\begin{aligned} \log(\text{Mob_eff})_{i,t} = & \beta_{0,i} + \alpha_i \log(\text{Mob_eff})_{i,t-1} + \beta_{1,i} \log(\text{Mob_std})_{i,t} \\ & + \beta_{2,i} \log(\text{Re_v_pop})_{i,t} + \beta_{3,i} \log(\text{Pko_pop})_{i,t} \\ & + \beta_{4,i} \log(\text{Freq_gen})_{i,t} + \beta_{5,i} \log(\text{Densite})_{i,t} \\ & + \beta_{6,i} \log(\text{Vitesse})_{i,t} + \beta_{7,i} \log(\text{Prix_TC})_{i,t} \\ & + \beta_{8,i} \log(\text{Prix_Carb})_{i,t} + \varepsilon_{i,t} \end{aligned} \quad (4)$$

où

³ Un des intérêts de cette approche est de pouvoir utiliser une série telle que celle du prix des carburants invariante entre PTU. Tout comme une constante, le coefficient associé à cette variable sera néanmoins différent d'un PTU à l'autre. C'est la seule approche sur données de panel qui permet d'obtenir de tels résultats.

⁴ Pour une présentation plus complète de cette classe d'estimateurs, se reporter à l'annexe.

(Mob_eff) _{i,t}	est la mobilité effective dans le PTU <i>i</i> à l'année <i>t</i>
(Mob_std)	est la mobilité structurelle
(Re v_pop)	est le revenu net d'impôt par tête
(Pko_pop)	est le nombre de places-km offertes par tête
(Freq_gen)	est la fréquence généralisée, mesurée par le rapport des véhicules*km à la longueur des lignes
(Densite)	est la densité du réseau, mesurée par le rapport de la longueur des lignes à la surface des communes desservies
(Vitesse)	est la vitesse commerciale, mesurée par le réseau (donnée peu fiable et parfois extrapolée)
(Prix_TC)	est l'indice des prix des transports en commun, mesuré par la recette directe moyenne par voyage
(Prix_Carb)	est l'indice des prix des carburants, plus précisément c'est l'indice du coût du carburant par kilomètre parcouru

et où tous les coefficients $\beta_{j,t}$ ($i = 1, \dots, 8$) sont les élasticités de court terme de la mobilité effective par rapport aux variables *j*. Ainsi, nous obtiendrons autant d'élasticités de court terme pour une variable qu'il y a de PTU.

3.2.3 Les bases de données utilisées

Nous disposons d'une base de données de trafics TCU dans 61 réseaux de Province (appelés PTU: Périmètres de Transports Urbains) couvrant la période 1975-1995.

Le PTU est le territoire sur lequel une (ou un ensemble) de commune(s) qui se déclare organisatrice d'un réseau de transports urbains. Cette déclaration permet à l'autorité organisatrice:

- d'avoir entière autorité sur l'organisation, la gestion du réseau, sa concession ou non à des exploitants privés, ...
- de prélever une taxe auprès des employeurs de son territoire (taxe proportionnelle à la masse salariale): le Versement Transport.

Les communes rurales péri-urbaines (où seuls les systèmes de ramassage scolaire sont viables) restent généralement en dehors des PTU. De plus, au sein d'une même agglomération (au sens de l'INSEE), il peut exister plusieurs PTU.

La liste et la composition communale des PTU est susceptible de varier dans le temps. Dans notre étude, nous ne disposons que de la composition communale en 1995. Nous avons cependant éliminé les PTU dont l'emprise territoriale (obtenue auprès du CERTU sur la période 1975-1993) a trop fluctué durant la période étudiée. La liste des PTU — qui fait l'objet de la modélisation — est donc composée de PTU:

- qui ont existé de 1975 à 1995,
- qui n'ont pas (ou peu) varié en surface entre 1975 et 1995,
- et qui sont presque complètement renseignés du point de vue des variables retenues sur la période 1975-

1995. Un traitement relativement lourd des données nous a permis d'obtenir un échantillon de 61 PTU observés sur 21 ans (1975-1995) et comprenant 27 variables par PTU. Donc, pour chacune des variables de la base de données, nous avons $NT = N \times T = 61 \times 21 = 1281$ observations.

La variable expliquée est la mobilité effective annuelle (exprimée en nombre déplacements par tête) et les variables explicatives couvrent quatre domaines:

- l'offre de transport
 - le nombre de véhicules-km (en km)
 - le nombre de places-km offertes (en km)
 - la vitesse commerciale (en km/h)
 - la longueur du réseau (en km)
 - la surface du PTU (en km²)
- le prix des transports,
 - la recette moyenne par voyageur (en Francs base 95)
 - le prix des carburants (base 1 en 1995)
- le revenu net d'impôt des ménages (en Francs base 95)
- la mobilité "structurelle" (en déplacements par tête).

Le tableau I retrace, pour l'ensemble des observations, les valeurs minimales, moyennes, maximales ainsi que le rapport (max / min). Ce coefficient est une indication de la variabilité plus ou moins importante de certaines séries. Ainsi, le nombre de véhicules-kilomètres varie de 1 à 230, la longueur des réseaux varie de 1 à 88, la surface des PTU varie de 1 à 41, etc. ... Comme les grands réseaux sont généralement les plus performants, "l'effet de masse" renforce donc "l'effet par tête". Les figures 1a, 1b et 1c représentent les moyennes individuelles des mobilités effective et structurelle. On constate des disparités tout à fait significatives entre les PTU, les moyennes de la mobilité effective (resp. de la mobilité structurelle et du rapport mobilité effective/mobilité structurelle) variant de 1 à 14.2 (resp. de 1 à 15.6 et de 1 à 4.27).

Pour les séries de mobilité, l'importance de la variabilité inter-individuelle (*i.e.* entre PTU) est confirmée par la décomposition de la variance (tableaux 2 et 3). Cela se manifeste par une grande proximité de toutes les observations concernant le même PTU vis-à-vis des distances inter-PTU.

Tableau 1 - Statistiques descriptives sur l'ensemble des observations 1975-1995 ($NT = 1281$)

Variabie	Minimum	Moyenne	Maximum	$\frac{\text{Maximum}}{\text{Minimum}}$	Taux de croissance annuel moyen %
Mobilité effective	8.415	74.749	203.834	24.222	2.656
Mobilité structurelle	5.775	43.679	112.280	19.442	-1.084
Mob. effect. / mob. struct.	0.656	1.931	5.392	8.219	4.581
Revenu fiscal par tête	23298.58	35906.65	49956.86	2.144	1.476
PKO par tête	279.98	2211.65	5700.39	20.359	4.490
Fréquence généralisée	2.453	25.263	56.469	23.019	-0.278
Densité des réseaux	0.134	1.212	5.916	44.149	3.484
Vitesse commerciale	10	15.725	31.919	3.192	1.015
Longueur des réseaux	15.2	195.162	1345.986	88.55	4.230
Nbre de Véhicules-km	234	5396.339	53784	229.846	3.378
Surface des PTU	15	173.966	618.26	41.217	0.637
Prix relatif (TC/carburants)	0.332	0.775	1.669	5.027	1.068
Prix des TC	0.493	0.985	1.909	3.872	-1.539
Prix des Carburants	1	1.307	1.581	1.581	-0.474

D'après le tableau 2, on constate l'extrême prédominance de la variabilité inter-individuelle dans la variabilité totale des observations. Ainsi, on observe que près de 84% de la variabilité de la mobilité effective est attribuable aux différences inter-PTU. Cette part s'élève à 98,5 % pour la mobilité structurelle. Ces résultats paraissent tout à fait cohérents car on perçoit bien que, pour un PTU, le niveau de mobilité effective (ou structurelle) change d'une année sur l'autre mais dans des plages de variation relativement étroites. La variabilité temporelle du niveau de mobilité pour chaque PTU est donc faible. En revanche, le panel de PTU est marqué par une très forte hétérogénéité. C'est un cas de figure tout à fait classique en économétrie des données de panel. Cette variabilité inter-PTU prédomine largement sur la variabilité intra-PTU

On peut également envisager la décomposition de la variabilité totale en une variabilité inter-temporelle et une variabilité intra-temporelle (tableau 3). Si l'on raisonne uniquement dans la dimension temporelle, on s'aperçoit maintenant que 90,3 % de la variabilité de la mobilité effective est attribuable aux différences intra-périodes. Cette part s'élève à 98,8 % pour la mobilité structurelle. Cette fois-ci, la variabilité intra l'emporte sur la variabilité inter. Cette prédominance s'explique par l'importance de la dimension individuelle. En effet, comme le cas le plus fréquent pour un panel est que le nombre d'individus de l'échantillon est grand ($N \rightarrow \infty$) et que le nombre de périodes est généralement faible (T fini), la variabilité inter-périodes est faible puisqu'elle dépend du nombre d'années disponibles. En revanche, la variabilité intra-périodes est forte puisqu'elle dépend directement du nombre d'individus (ici, les PTU) présents dans l'échantillon⁵.

⁵ Plus précisément, pour une année donnée t et si $N \rightarrow \infty$, alors l'information disponible pour l'année t est extrêmement riche puisque l'on observe une infinité de comportements pour cette date.

Nous pouvons constater, toujours d'après le tableau 2, des proportions inverses dans la structure des variabilités lorsque les variables sont exprimées en ratios. C'est le cas notamment du prix relatif (transports en commun / carburants), du prix relatif des transports en commun⁶ ou bien du revenu fiscal par tête (la variabilité inter-individuelle représente respectivement 28.3 %, 53.2 % et 45.6 %).

Tableau 2 - Décomposition individuelle de la variance des observations ($NT = 1281$)

Variable	% variabilité inter-individuelle	% variabilité intra-individuelle
Mobilité effective	83.250	16.750
Mobilité structurelle	98.442	1.558
Mob.effect. / mob. struct.	50.474	49.526
Revenu fiscal tête	45.628	54.372
PKO par tête	67.108	32.892
Fréquence généralisée	89.114	10.886
Densité des réseaux	71.777	28.333
Vitesse commerciale	49.333	50.667
Longueur des réseaux	87.044	12.956
Nbre de Véhicules-km	95.490	4.510
Surface des PTU	97.192	2.808
Prix relatif (TC/carburants)	28.342	71.658
Prix des TC	53.245	46.755
Prix des Carburants	0	100

Tableau 3 - Décomposition temporelle de la variance des observations ($NT = 1281$)

Variable	% variabilité inter-temporelle	% variabilité intra-temporelle
Mobilité effective	9.715	90.285
Mobilité structurelle	1.203	98.797
Mob.effect. / mob. struct.	29.359	70.641
Revenu fiscal par tête	49.866	50.134
PKO par tête	23.993	76.007
Fréquence généralisée	0.519	99.481
Densité des réseaux	10.723	89.277
Vitesse commerciale	21.714	78.286
Longueur des réseaux	5.433	94.567
Nbre de Véhicules-km	1.780	98.220
Surface des PTU	0.339	99.661
Prix relatif (TC/carburants)	49.034	50.966
Prix des TC	3.673	96.327
Prix des Carburants	100	0

⁶ La variabilité inter-individuelle (resp. intra-période) du prix relatif des carburants est nulle puisque ce prix relatif est le même pour tous les PTU.

3.2.4 Estimation des élasticités de court terme et de long terme

L'évolution du prix des carburants sur la période 1975-1995 fait apparaître deux sous-périodes distinctes. L'une avant le contre-choc pétrolier 1975-1985 où l'indice des prix des carburants est croissant (1.39 à 1.55, soit +0.9% par an) et l'autre de 1986 à 1995 où l'indice des prix des carburants est nettement plus bas et décroissant (1.25 à 1, soit -2.17% par an en moyenne). On constate donc, sur l'ensemble de la période, des niveaux différents ainsi que des pentes différentes pour les prix des carburants. Ces ruptures de pente et de niveau perturbent les estimations et fournissent des coefficients difficilement interprétables. Pour tenir compte de ce phénomène, il est nécessaire d'introduire un mécanisme à changement de régimes captés par des constantes différentes selon les périodes. Le modèle général est donc :

$$\begin{aligned} \log(\text{Mob_eff})_{i,t} = & \alpha_i \log(\text{Mob_eff})_{i,t-1} + \beta_{1,i} \log(\text{Mob_std})_{i,t} + \beta_{2,i} \log(\text{Rev_pop})_{i,t} \\ & + \beta_{3,i} \log(\text{Pko_pop})_{i,t} + \beta_{4,i} \log(\text{Freq_gen})_{i,t} + \beta_{5,i} \log(\text{Densite})_{i,t} \\ & + \beta_{6,i} \log(\text{Prix_TC})_{i,t} + \beta_{7,i} \log(\text{Prix_Carb})_{i,t} \\ & + \begin{cases} \beta_{0,i,75-85} & \text{si } t < 1985 \\ \beta_{0,i,86-95} & \text{si } t \geq 1986 \end{cases} \end{aligned} \quad (5)$$

Une première série de tests a consisté à estimer le modèle (5) sans la variable "mobilité structurelle", sans le prix des carburants et avec une seule variable d'offre (pko) afin de comparer nos résultats avec les estimations obtenues en Grande-Bretagne par *Dargay et Hanly* (1999). Pour les zones métropolitaines anglaises (sauf Londres) entre 1987 et 1996, ces auteurs ont régressé la "demande de transports en commun" (*bus demand* \equiv "*bus passenger journeys per capita*") sur le revenu des ménages par tête, le prix des TC (*bus fare*) ainsi qu'un indicateur de qualité de service (*service* \equiv le nombre de véhicules-kilomètres par tête). A titre d'information, les élasticités de la demande de "bus" estimées par *Dargay et Hanly* (1999) sont les suivantes⁷:

	court terme	long terme
<i>bus fare</i>	-0.23	-0.44
<i>Income</i>	-0.84	-1.59
<i>service</i>	0.35	0.67

Par comparaison, nous testons la même spécification⁸ (A-1), (voir tableau 4). Du fait de l'énorme quantité d'informations (2 * 61 coefficients de court terme et de long terme par variable), nous ne donnons, par variable

⁷ Voir *Dargay et Hanly* (1999) page 47.

⁸ La seule différence vient de l'indicateur de qualité de service. Nous utilisons le nombre de places-kilomètres offertes par tête. Mais nous disposons également des véhicules*km (numérateur de la fréquence généralisée).

et pour l'ensemble des PTU, que les valeurs minimales et maximales des coefficients de court terme et de long terme. Les intervalles de variation de ces coefficients sont donnés dans le tableau 4. On constate que la mobilité effective peut être considérée comme un bien inférieur (l'élasticité-revenu est négative). C'est un résultat identique à celui de *Dargay et Hanly* (1999) avec cependant des amplitudes de court terme et de long terme beaucoup plus réduites. D'autre part, l'élasticité estimée du revenu beaucoup plus forte sur données anglaises que sur données françaises. La différence ne provient pas de la vitesse d'ajustement puisque les coefficients d'ajustement α_i sont proches de la valeur estimée pour l'Angleterre (0.472). Ainsi, le délai médian d'ajustement⁹, — le temps nécessaire pour que la variable "mobilité effective" parcoure la moitié de la distance séparant sa valeur initiale de sa valeur stationnaire — est d'environ 0.92 année pour l'Angleterre¹⁰ et [0.915; 1.192] année pour les réseaux de Province. Ces résultats sont tout à fait comparables. On constate un effet-prix négatif relativement important à long terme [-0.69; -0.51] et à peu près identique à l'effet-revenu. Par contre, l'effet d'offre est sensiblement plus faible (en valeur absolue) que les effets prix et revenu. La qualité de l'ajustement (mesurée par le *mapre*¹¹) est relativement bonne. Le modèle estime les observations réelles avec une erreur moyenne de $\pm 3.43\%$, les erreurs s'échelonnant, pour les différents PTU entre $\pm 1.05\%$ et $\pm 10.03\%$.

Nous avons également testé une spécification (A-2) (tableau 4) tenant compte de plus d'informations relatives à l'offre de service. L'introduction de nouvelles variables explicatives ne modifie pas significativement la vitesse d'ajustement. Le délai médian d'ajustement est environ de [0.823; 0.890] année pour les réseaux de Province. Il faut toujours à peu près une année pour que 50% de l'ajustement soit réalisé. Par contre, on constate, pour les variables *pko*, prix des TC et revenu, un rétrécissement des intervalles de variation des élasticités de court terme et de long terme. Les effets-prix et les effets d'offre sont plus importants que l'effet-revenu. Ce dernier, toujours négatif et inférieur à l'unité (en valeur absolue), signifie que la "mobilité effective" est un bien inférieur, c'est-à-dire un bien dont la consommation diminue lorsque le revenu augmente. L'élasticité au prix des carburants est positive et inférieure à l'unité. Cela veut dire que les modes de transports collectifs et transports privés sont des substituts bruts. L'impact du prix des carburants est naturellement plus faible que celui du prix des transports en commun: le rapport est environ de 1 à 2.1.

⁹ Le délai médian d'ajustement est défini par:

$$-\frac{\log 2}{\log \alpha_i}$$

¹⁰ Son intervalle de variation important [0.658; 1.336] traduit une imprécision des estimations.

¹¹ Le *mapre* est défini par:

$$\frac{100}{T} \sum_{t=1}^T \left| \frac{\hat{\epsilon}_{i,t}}{Y_{i,t}} \right|$$

Le *mapre* est la moyenne des valeurs absolues des erreurs relatives exprimées en pourcentage ("Mean Absolute Percent Relative Error"). C'est un test de qualité d'estimation traduisant une certaine erreur moyenne. Le meilleur modèle sera celui qui possède le plus petit *mapre*.

Tableau 4 - Intervalles de variation des coefficients de court terme et de long terme

	spécification A-1		spécification A-2	
	court terme	long terme	court terme	long terme
mobilité effective($t - 1$)	{0.469; 0.559}		{0.365; 0.409}	
prix des TC	{-0.304; -0.269}	{-0.691; -0.509}	{-0.292; -0.285}	{-0.526; -0.513}
pko par tête	{0.151; 0.570}	{0.343; 1.075}	{0.324; 0.355}	{0.570; 0.656}
revenu par tête	{-0.409; -0.174}	{-0.773; -0.396}	{-0.173; -0.129}	{-0.318; -0.227}
fréquence	---	---	{0.135; 0.147}	{0.248; 0.262}
prix des carburants	---	---	{0.148; 0.186}	{0.274; 0.292}
densité	---	---	{0.040; 0.050}	{0.070; 0.109}
vitesse	---	---	{0; 0.009}	{0; 0.014}
dummy 1975-85	{1.870; 2.782}		{0.397; 0.467}	
dummy 1986-95	{1.936; 2.756}		{0.467; 0.512}	
Mapre (%)	{1.050; 3.436; 10.036}		{1.326; 3.738; 9.071}	

Les variables d'offre (fréquence généralisée, densité et vitesse) ont des effets positifs et significatifs. Néanmoins, l'effet vitesse commerciale est relativement négligeable. Cela peut provenir, soit d'un manque de variabilité dans le temps et entre FTU, soit de la mesure¹² même de cette variable "vitesse commerciale". Enfin, un dernier point à noter concerne les indicatrices d'avant et d'après le contre-choc pétrolier. On observe des différences significatives entre les deux sous-périodes confirmant donc le choix du modèle à changement de régimes. Quoiqu'il en soit, l'introduction de nouvelles variables d'offre et du prix des carburants enrichit considérablement le modèle "à la *Dargay*".

3.2.5 L'impact de la mobilité structurelle

Quand on introduit la variable "mobilité structurelle" (spécification A-3) (voir tableau 5), la vitesse d'ajustement est plus rapide. Le délai médian d'ajustement est de l'ordre de [0.687; 0.775]. Il faut maintenant environ 8 à 9 mois pour que 50% de l'ajustement soit réalisé. De plus, on constate un effondrement de l'effet-revenu, celui-ci a tendance à devenir nul. L'élasticité-revenu est légèrement négative (bien faiblement inférieur) ou nulle (bien normal) selon les PTU. Le modèle n'arrive pas à estimer les effets-revenu pour les PTU d'Amiens, du Havre, de Nice et de Rouen. Cette inversion de nature du bien "transports en commun" ne s'explique pas, sinon par un défaut de spécification. Il semble donc qu'il y ait une "incompatibilité" entre les variables de revenu et de mobilité structurelle. Quand on omet la mobilité structurelle, dont la principale composante est une croissance de la motorisation liée au revenu sur la période observée, le revenu se comporte comme "proxy" de la motorisation. Quand on introduit ces deux variables, le modèle attribue le rôle principal à la motorisation (effet de concurrence avec l'automobile rendu par la mobilité structurelle) et donne un rôle résiduel au revenu.

Une fois pris en compte cet effet de concurrence de l'automobile, l'effet résiduel du revenu reste négatif, mais il est très faible et quelquefois non significatif. En moyenne sur l'ensemble des PTU, l'élasticité revenu

¹² Il semblerait que VIA GTI confirme la mauvaise qualité de mesure de cette variable.

est en effet égale à -0.035 à court terme et -0.059 à long terme. Sur le tableau suivant, on constate que les élasticités de la mobilité effective par rapport au revenu sont très faibles quelle que soit la répartition (minimum, maximum et 1er et 3ème quartiles des distributions):

	min	Q1	Q3	max
effet-revenu à court terme	-0.06	-0.04	-0.025	0.0
effet-revenu à long terme	-0.11	-0.07	-0.06	0.0

L'élasticité de la mobilité effective à la mobilité structurelle est positive et toujours inférieure à l'unité. La mobilité effective est donc assez peu élastique à la mobilité structurelle. La présence simultanée des variables de revenu et de mobilité structurelle a tendance à réduire l'impact du prix des carburants — peu variable entre PTU — et l'effet de variables d'offre telles que la fréquence généralisée et la densité du réseau. Par contre, on constate, par rapport à la spécification A-2, un accroissement de l'élasticité à la vitesse commerciale. Les différentes valeurs du test mapre (min, moyenne et max) confirment que la qualité du modèle ne s'améliore pas par rapport à celle de la spécification précédente.

Nous avons donc choisi d'éliminer la variable "revenu fiscal par tête" dont l'impact est très faible et quelquefois statistiquement non significativement différent de zéro. Si la vitesse d'ajustement ne change pas (spécification A-4), en revanche, les effets de certaines variables telles que la vitesse commerciale et la densité sont non significativement différents de zéro pour certains PTU. L'élimination de la variable revenu améliore la qualité de l'estimation (le mapre diminue), mais il faut encore supprimer l'effet vitesse de la régression puisque l'effet à long terme de cette variable est toujours nul, quel que soit le PTU

Tableau 5 - Intervalles de variation des coefficients de court terme et de long terme

	spécification A-3		spécification A-4	
	court terme	long terme	court terme	long terme
mobilité effective($t - 1$)	[0.365; 0.409]		[0.368; 0.428]	
prix des TC	[-0.340; -0.315]	[-0.536; -0.529]	[-0.355; -0.323]	[-0.567; -0.554]
pko par tête	[0.368; 0.442]	[0.581; 0.749]	[0.329; 0.461]	[0.521; 0.804]
fréquence	[0.071; 0.087]	[0.119; 0.138]	[0.039; 0.110]	[0.069; 0.174]
revenu par tête	[-0.066; 0]	[-0.111; 0]	—	—
prix des carburants	[0.113; 0.122]	[0.189; 0.194]	[0.089; 0.098]	[0.152; 0.158]
densité	[0.009; 0.038]	[0; 0.06]	[0.011; 0.039]	[0; 0.065]
vitesse	[0.013; 0.024]	[0.021; 0.041]	[0; 0.010]	[0; 0]
mobilité structurelle	[0.254; 0.317]	[0.427; 0.500]	[0.239; 0.347]	[0.414; 0.551]
dummy 1975-85	[-1.67; -1.53]		[-2.23; -1.69]	
dummy 1986-95	[-1.69; -1.51]		[-2.09; -1.73]	
Mapre (%)	[1.270; 3.661; 10.131]		[1.182; 3.642; 9.253]	

Le tableau 6 fournit les élasticités de court terme et de long terme de la meilleure spécification. Les valeurs du mapre montrent la bonne qualité de l'ajustement. Cette spécification (A-5) estime les valeurs observées de la mobilité effective avec, en moyenne pour tous les PTU, une erreur de $\pm 3.464\%$. Les erreurs moyennes

s'échelonnent entre $\pm 1.178\%$ pour le PTU d'Arras et $\pm 8.409\%$ pour le PTU d'Avignon. Par rapport aux autres spécifications, on constate un resserrement de l'intervalle de variation du test mapre.

Tableau 6 - Intervalles de variation des coefficients de court terme et de long terme

	spécification A-5	
	court terme	long terme
mobilité effective($t-1$)	{0.381; 0.451}	
prix des TC	{-0.355; -0.315}	{-0.582; -0.567}
pko par tête	{0.299; 0.457}	{0.483; 0.834}
fréquence	{0.041; 0.116}	{0.075; 0.188}
prix des carburants	{0.079; 0.087}	{0.127; 0.160}
densité	{0.018; 0.051}	{0.033; 0.085}
mobilité structurelle	{0.203; 0.260}	{0.370; 0.582}
dummy 1975-85	{-2.298; -1.626}	
dummy 1986-95	{-2.164; -1.659}	
Mapre (%)		{1.178; 3.464; 8.409}

3.2.6 L'hétérogénéité des variables d'offre

On peut classer les variables explicatives par ordre décroissant selon leur impact sur la mobilité effective (voir graphique ci-dessous):

- places-kilomètres offertes par tête;
- prix des transports en commun;
- mobilité structurelle;
- fréquence généralisée;
- prix des carburants;
- et densité du réseau.

On constate que les intervalles de variation des élasticités (de long terme) sont très étroits pour les variables de prix et beaucoup plus larges pour les variables d'offre. De même, les élasticités de la mobilité effective par rapport à la mobilité structurelle varient sensiblement entre PTU. Ces fluctuations retracent l'hétérogénéité des comportements des usagers des TCU dans les différents PTU. L'hétérogénéité entre PTU ne se manifeste donc pas par l'intermédiaire de la réaction aux tarifs, mais plutôt par la sensibilité à l'offre (surtout à son volume: les places-kilomètres offertes). En effet, l'hétérogénéité se traduit principalement par la variabilité des facteurs de qualité de service et de demande "structurelle". Cette dernière traduit l'évolution de la structure de la population, sur la base des critères de l'âge, du sexe, de la zone de résidence (ville-centre ou banlieue) à l'intérieur du PTU et surtout de la motorisation des ménages. Les figures 3 à 8 retracent, pour chaque variable explicative, les variabilités des élasticités de court terme et de long terme de l'ensemble des PTU. On constate des

distributions unimodales¹³ pour chaque élasticité avec un déplacement des modes et des bornes minimales et maximales entre le court terme et le long terme. Ainsi, le mode de l'élasticité à court terme (resp. à long terme) de la mobilité effective par rapport à la mobilité structurelle est de 0.30 (resp. 0.50), soit une augmentation des deux-tiers. Pour toutes les variables explicatives, on constate que le passage court terme - long terme entraîne une augmentation (en valeur absolue) des élasticités de 55 à 75%. Il n'y a donc pas doublement des élasticités comme pour la spécification A-I. Puisque le rapport court terme / long terme doit être le même pour toutes les variables sur un même PTU, cette variabilité retrace donc exclusivement une diversité entre PTU. La diversité entre variables du tableau suivant résulte donc du fait que ce ne sont pas les mêmes réseaux qui sont "au mode" pour toutes les variables explicatives.

	Mode CT	Mode LT	$\Delta\%$
mobilité structurelle	0.30	0.50	66.67
pko par tête	0.36	0.63	75
prix des TC	-0.34	-0.5738	68.76
fréquence généralisée	0.09	0.14	55.55
prix des carburants	0.0825	0.141	70.91
densité du réseau	0.033	0.055	66.67

Afin d'obtenir une classification de l'ensemble des PTU, nous avons ordonné les PTU en fonction des élasticités de long terme relatives à la mobilité structurelle et aux variables d'offre par ordre d'importance de l'influence sur la mobilité effective. Rappelons que l'ordre d'influence des variables explicatives (autres que les prix) est le suivant:

- places-kilomètres offertes par tête;
- mobilité structurelle;
- fréquence généralisée;
- et densité du réseau.

Le regroupement des PTU peut être confirmé par une analyse en composantes principales (ACP) des élasticités de long terme de la mobilité effective par rapport à la mobilité structurelle et aux variables d'offre. Les tableaux 7a et 7b donnent la matrice de corrélation des élasticités de long terme, les coordonnées des variables sur le cercle unité ainsi que le calcul des valeurs propres et du pourcentage d'inertie.

¹³ Précisons que l'on ne peut obtenir que des distributions unimodales du fait de la définition des coefficients aléatoires $\beta_{j,t}$.

Tableau 7a - Matrice de corrélation des élasticités de long terme de la mobilité effective et coordonnées des variables pour les 61 PTU

	PKO	Mobilité structurelle	Fréquence	densité	1ère composante principale réduite	2ème composante principale réduite
PKO	1	-0.9467	-0.9364	-0.3208	-0.91740	-0.39775
Mob. structurelle		1	0.9994	0.6085	0.99668	0.08128
Fréquence			1	0.6317	0.99855	0.05167
Densité				1	0.67104	-0.7414

Tableau 7b - Valeurs propres et pourcentage d'inertie

	Valeurs propres	% d'inertie
Comp.princ. 1	3.28239	82.0598
Comp.princ. 2	0.71715	17.9288
Comp.princ. 3	0.00043	0.0107
Comp.princ. 4	0.00003	0.0008

Les deux premières composantes principales reconstruisent 99.989% de la variance totale (tableau 7b, colonne 3). Nous nous bornerons donc au cercle de corrélation — sur le plan des deux premières composantes principales réduites¹⁴ (tableau 7a, colonnes 6 et 7) — pour les 4 variables et au plan principal — des deux premières composantes principales — pour les 61 PTU. Toutes les variables sont bien représentées sur le cercle de corrélation puisqu'elles sont proches de la circonférence, à l'exception peut-être de la variable densité. Il apparaît une distinction évidente, entre les pko et les variables fréquence et mobilité structurelle, représentée par des corrélations élevées et négatives (tableau 7a). L'effet pko a tendance à "s'opposer" aux effets fréquence et mobilité structurelle et la première composante principale est définie par cette opposition. La deuxième aurait plutôt tendance à séparer la densité des autres variables explicatives, la corrélation étant beaucoup plus faible entre ces variables (voir figure 9). En ce qui concerne les PTU, on retrouve bien, sur le graphique, la classification opérée précédemment. Les axes pko, mobilité structurelle, fréquence et densité représentent le sens d'évolution de la sensibilité des PTU par rapport à ces facteurs. Ainsi, pour les pko, nous retrouvons les PTU de Blois, Lille, Bourges, ... De même, on constate que Bordeaux est le PTU le moins réactif aux quatre variables testées. Enfin, notons que le PTU de Reims, au centre du graphique, est donc le PTU qui "accorde" le même poids aux effets pko, mobilité structurelle, fréquence généralisée et densité du réseau.

¹⁴ Toute variable, positionnée sur le plan des deux premières composantes principales réduites, a pour coordonnées les corrélations entre cette variable et les deux premières composantes principales.

3.2.7 Qualité de l'estimation par PTU

Du fait de la trop grande quantité d'informations, nous ne pouvons pas donner pour chaque PTU les estimations de court terme et de long terme ainsi que, les intervalles de confiance à 95% de sécurité des coefficients de court terme et de long terme par variable et pour l'ensemble des PTU. La qualité générale du modèle est donc représentée par le test mapre (voir tableau 6). Cependant, on peut, pour chaque PTU, donner la valeur de ce test qui traduit une certaine erreur moyenne entre la mobilité effective observée et celle calculée par le modèle (voir tableau 8). 50% des PTU ont un mapre inférieur à $\pm 3.13\%$ et 90% des PTU ont un mapre inférieur à $\pm 5.28\%$. Le minimum est de $\pm 1.178\%$ pour le PTU d'Arras et le maximum est de $\pm 8.409\%$ pour le PTU d'Avignon. Les 10% des PTU, ayant le mapre le plus élevé, sont: Avignon, Grenoble, Châteauroux, Quimper, Périgueux et Orléans (voir figures 10a, 10b et 10c).

Tableau 8 - Qualité de l'estimation: valeurs du MAPRE en % pour les 61 PTU

1 Aix-en-Provence	4.589	32 Lyon	2.893
2 Amiens	3.779	33 Mans (Le)	2.757
3 Angers	5.035	34 Marseille	2.108
4 Angoulême	3.175	35 Metz	2.481
5 Annecy	2.391	36 Mulhouse	2.215
6 Annemasse	3.485	37 Nancy	2.382
7 Arras	1.178	38 Nantes	3.217
8 Avignon	8.409	39 Nevers	4.903
9 Bayonne	1.942	40 Nice	2.467
10 Belfort	4.006	41 Nîmes	3.313
11 Besançon	2.668	42 Niort	3.695
12 Blois	5.285	43 Orléans	6.144
13 Bordeaux	2.635	44 Périgueux	6.444
14 Boulogne-sur-Mer	3.358	45 Perpignan	2.972
15 Bourges	3.315	46 Poitiers	3.494
16 Brest	4.074	47 Quimper	6.517
17 Caen	3.117	48 Reims	1.837
18 Calais	2.363	49 Rennes	2.326
19 Cannes	2.269	50 Roanne	3.130
20 Châteaunoux	6.525	51 Rochelle (La)	2.948
21 Cherbourg	2.252	52 Rouen	3.360
22 Clermont-Ferrand	2.803	53 Saint-Etienne	2.559
23 Creil	4.161	54 Saint-Quentin	1.556
24 Creusot (Le) - Montceau-les-Mines	5.109	55 Strasbourg	2.268
25 Dijon	1.500	56 Tarbes	4.686
26 Dunkerque	3.329	57 Toulon	3.393
27 Evreux	4.343	58 Tours	3.063
28 Grenoble	7.348	59 Troyes	2.877
29 Havre (Le)	2.521	60 Valence	2.369
30 Lille	4.262	61 Valenciennes	2.310
31 Limoges	2.433		

4. Conclusion

L'objectif de cette étude est d'estimer les tendances à long terme des trafics des transports collectifs urbains de 61 villes de France sur la période 1975-1995. Pour représenter simultanément les évolutions à court terme et à long terme, la spécification retenue est une relation dynamique stochastique entre la mobilité effective et un certain nombre de variables relatives à l'offre de transport, aux prix des transports, aux revenus des citoyens et à la structure de la population.

Afin de tenir compte de l'hétérogénéité potentielle des comportements des usagers selon les PTU, l'approche utilisée bénéficie de l'information relative à tous les PTU simultanément. Elle permet de révéler des comportements hétérogènes entre PTU par l'intermédiaire d'élasticités différentes pour une même variable. La bonne qualité des estimations conduit à un certain nombre de conclusions tout à fait significatives.

On peut classer les variables explicatives par ordre décroissant selon leur impact à long terme sur la mobilité effective:

- places-kilomètres offertes par tête [48%, 83%];
- prix des transports en commun [-57%, -56%];
- mobilité structurelle [37%, 52%];
- fréquence généralisée [7%, 18%];
- prix des carburants [12%, 16%];
- et densité du réseau [3%, 8.5%].

D'un réseau à l'autre, les comportements sont donc beaucoup plus homogènes vis-à-vis des prix que vis-à-vis de l'offre et des facteurs structurels. L'élasticité de la mobilité effective à la mobilité structurelle est positive et toujours inférieure à l'unité. La mobilité effective est donc assez peu élastique à la mobilité structurelle. Cette dernière traduit l'évolution de la structure de la population, sur la base des critères de l'âge, du sexe, de la zone de résidence (ville-centre ou banlieue) à l'intérieur du PTU et surtout de la motorisation des ménages. Les facteurs structurels, qui sont défavorables aux transports collectifs (motorisation croissante, étalement urbain,...), ont donc une importance moyenne.

Quant au revenu, son influence est nettement négative si l'on omet le facteur structurel: le transport public est alors un "bien inférieur". En fait, l'élasticité-revenu devient nulle (ou très faiblement négative) quand on introduit la mobilité structurelle. En effet, la progression de la motorisation est une composante importante des

effets de structure. En raison de la forte corrélation entre croissance de la motorisation et croissance du revenu au cours de la période observée — corrélation qui devrait s'estomper à l'avenir quand on va approcher d'une certaine saturation — on risque d'attribuer à tort au revenu une partie de l'effet de concurrence de l'automobile (l'autre partie étant l'élasticité croisée au prix des carburants).

De plus, on constate que le volume d'offre compte plus que les différentes composantes de qualité. enfin, il semble exister des compensations entre l'effet du volume d'offre d'une part, de la fréquence et de la mobilité structurelle d'autre part.

En conclusion, les facteurs d'offre au sens large (prix relatifs et qualité de service des différents modes) peuvent donc bien être introduits explicitement dans le même modèle que les facteurs structurels. Cependant, faute d'avoir pu estimer la mobilité par un modèle Age-cohorte dans chaque ville de province, les décalages entre générations, dont nous avons montré le rôle moteur dans les évolutions à long terme (*Bussière, Armoogum et Madre* (1995)), n'ont pas encore été introduits dans le calcul de la "mobilité structurelle".

Bien souvent, l'influence des facteurs économiques et d'offre est estimée sur des séries temporelles agrégées, ce qui conduit à surestimer leur effet faute de donner un rôle suffisant aux facteurs structurels qui sont au coeur de notre approche; la confusion entre les effets de la motorisation et du revenu en fournit un bon exemple. En s'appuyant sur une estimation correcte de l'influence de ces facteurs, les projections pourront être modulées en fonction de différents scénarios de croissance économique et de politique des transports.

Afin de répondre aux besoins de planification, il semble que l'utilisation conjointe de la démographie et de l'économétrie fournisse à la fois les outils adéquats et des résultats fiables sur lesquels on pourra bâtir une démarche prospective. A cet égard, l'utilisation plus systématique des modèles à coefficients aléatoires sur données de panel permet de mesurer précisément les différences de comportement entre individus même si la variable étudiée est commune à tous les individus. C'est le cas notamment du prix d'un bien tel que l'essence. De plus, la mesure de l'hétérogénéité des comportements en matière de transport ne passe plus nécessairement par l'estimation d'une constante (modèle à effets fixes, *Madre et Pirotte* (1991), *Bresson, Madre et Pirotte* (1997)) ou par la décomposition de la variance des perturbations (modèle à erreurs composées). Au delà de l'étude des transports urbains, l'application des méthodes d'estimation Bayésiennes sur panel semble avoir une portée générale en économétrie de la consommation.

5. Annexe: Modèles à coefficients aléatoires et “estimateurs à rétrécisseur”

Il existe des situations dans lesquelles des changements structurels impliquent que les paramètres des variables explicatives peuvent être différents selon les individus. Il est alors nécessaire de décomposer les coefficients des variables explicatives.

Les années récentes ont montré un intérêt croissant pour l'estimation de modèles à coefficients aléatoires (*Raj et Ullah (1981), Swamy et Tavlas (1995)*). Brièvement, on peut dire que même si la théorie économique ne précise pas les formes fonctionnelles exactes des relations, la plupart des méthodes économétriques font cependant des hypothèses spécifiques sur ces formes. Si la procédure en soit n'est pas critiquable, les résultats peuvent cependant dépendre de la forme fonctionnelle utilisée et/ou de la façon dont la forme fonctionnelle est insérée dans un cadre stochastique. La pratique économétrique usuelle de l'additivité des perturbations aléatoires à une relation mathématique n'est peut-être pas toujours correcte (*Klein (1989)*). Dans un modèle à coefficients aléatoires, on suppose que chaque coefficient est la somme de deux éléments: l'effet direct d'une variable explicative sur la variable expliquée et les effets indirects (ou biais de variables omises) dus au fait que la variable explicative affecte les variables exclues et celles-ci affectent en retour la variable explicative. Dans cette classe de modèles, l'effet direct est volontairement dissocié des effets indirects. Chaque coefficient a deux composantes: une composante fixe et une composante stochastique. Cette dernière peut être un bruit blanc ou peut suivre un processus plus complexe. Une des raisons d'utiliser cette approche est, qu'en cas d'erreur de spécification (*i.e* variables omises, forme fonctionnelle incorrecte, ...), il n'est pas raisonnable de supposer que le “bruit” affecte uniquement le terme d'erreur. Il peut contaminer tous les coefficients de l'équation.

Soit $y_{i,t}$, $i = 1, \dots, N$ et $t = 1, \dots, T$, l'observation de la variable y (par exemple, la mobilité effective (*i.e*, le trafic par tête)) du PTU i à la période t et soit $X_{j,i,t}$, l'observation de la variable X_j (par exemple, l'offre par tête) du PTU i à la période t . Supposons que l'on observe tous les individus (*i.e*, tous les PTU) sur les mêmes périodes. En empilant les N individus, on obtient un vecteur colonne y de taille $(NT, 1)$ soit:

$$y_{(NT,1)} = \begin{pmatrix} y_1 \\ (T,1) \\ y_2 \\ (T,1) \\ \vdots \\ y_N \\ (T,1) \end{pmatrix} \text{ avec } y_i = \begin{pmatrix} y_{i,1} \\ y_{i,2} \\ \vdots \\ y_{i,T} \end{pmatrix} \text{ pour } i = 1, \dots, N$$

De même, pour les k variables explicatives du modèle, on définit une matrice X de taille (NT, k) soit:

$$X_{(NT,k)} = \begin{pmatrix} X_{11} & X_{12} & X_{1k} \\ (T,1) & (T,1) & (T,1) \\ X_{21} & X_{22} & X_{2k} \\ (T,1) & (T,1) & (T,1) \\ \vdots & \vdots & \vdots \\ X_{N1} & X_{N2} & X_{Nk} \\ (T,1) & (T,1) & (T,1) \end{pmatrix}$$

On suppose que la spécification étudiée est un modèle linéaire dans sa forme la plus générale:

$$y_{i,t} = \sum_{j=1}^k \beta_{j,i,t} X_{j,i,t} + \varepsilon_{i,t} \quad (6)$$

soi, matriciellement:

$$y_i = X_i \beta_i + \varepsilon_i, \quad i = 1, \dots, N \quad (7)$$

Si tous les coefficients étaient supposés fixes et différents entre individus de différentes périodes, il y aurait alors NTk paramètres avec seulement NT observations. De manière évidente, il n'y a aucune façon d'obtenir un estimateur de $\beta_{j,i,t}$. Nous sommes obligés de faire quelques hypothèses sur la décomposition des coefficients.

D'une manière générale, chaque coefficient de régression peut être considéré comme une variable aléatoire ayant une distribution de probabilité (Hurwicz (1950), Klein (1953), Zellner (1969)). L'intérêt de la spécification à coefficients aléatoires est qu'elle réduit le nombre de paramètres à estimer tout en permettant aux paramètres de varier entre individus.

Les modèles à coefficients aléatoires supposent que le vecteur de coefficients β_i a une moyenne constante β et une matrice de variances-covariances constante:

$$\beta_i = \beta + \nu_i, \quad V[\nu_i] = E[\nu_i \nu_j^T] = \begin{cases} \Delta & \text{pour } i = j \\ 0 & \text{sinon} \end{cases}$$

où β est un vecteur commun de coefficients moyens $\beta = (\beta_1, \dots, \beta_k)^T$ et $\nu_i = (\nu_{1,i}, \dots, \nu_{k,i})^T$ est un vecteur de variables aléatoires de moyenne nulle et de matrice de variances-covariances Δ constante retraçant les caractéristiques spécifiques à chaque individu. Dans ce type de modèle, nous sommes intéressés par l'estimation des vecteurs β_i , du vecteur des coefficients moyens β et par la prédiction de chaque composante individuelle ν_i .

Le modèle (6) peut donc s'écrire:

$$y_{i,t} = \sum_{j=1}^k \beta_{j,i,t} X_{j,i,t} + \varepsilon_{i,t} = \sum_{j=1}^k (\beta_j + \nu_{j,i}) X_{j,i,t} + \varepsilon_{i,t} \quad (8)$$

Swamy (1970) a supposé que:

$$\begin{aligned} E[\nu_i] &= \mathbf{0} \\ E[\nu_i \nu_j^\top] &= \begin{cases} \Delta & \text{pour } i = j \\ 0 & \text{sinon} \end{cases} \\ E[X_{i,t} \nu_j^\top] &= \mathbf{0} \\ E[\varepsilon_i \varepsilon_j^\top] &= \begin{cases} \sigma_i^2 I_T & \text{pour } i = j \\ 0 & \text{sinon} \end{cases} \end{aligned} \quad (9)$$

Si l'on désire obtenir des estimations du vecteur des coefficients β_i , il faut spécifier la forme de sa distribution. Il est alors nécessaire d'utiliser des estimateurs tels que celui du Maximum de Vraisemblance ou bien les estimateurs bayésiens.

5.1 L'estimateur du Maximum de Vraisemblance

On suppose que le vecteur des perturbations ε_i ($i = 1, \dots, N$) et le vecteur de coefficients β_i ($i = 1, \dots, N$) sont normalement distribués. Puisque $y_i = X_i \beta_i + \varepsilon_i$ et $\beta_i = \beta + \nu_i$, alors le vecteur y_i suit une loi normale de moyenne $X_i \beta$ et de matrice de variances-covariances $\Phi_i = X_i \Delta X_i^\top + \sigma_i^2 I_T$. Les vecteurs y_1, \dots, y_N sont supposés être mutuellement indépendants. Sous ces conditions, l'estimateur du Maximum de Vraisemblance est celui qui maximise la fonction log-vraisemblance¹⁵:

$$\begin{aligned} \text{Log } L(\beta_i, \beta, \sigma_i^2, \Delta | y, X) &= \text{Cte} - \frac{T}{2} \sum_{i=1}^N \text{Log} \sigma_i^2 - \frac{N}{2} \text{Log} |\Delta| \\ &\quad - \frac{1}{2} \sum_{i=1}^N \frac{1}{\sigma_i^2} (y_i - X_i \beta_i)^\top \cdot (y_i - X_i \beta_i) \\ &\quad - \frac{1}{2} \sum_{i=1}^N (\beta_i - \beta)^\top \cdot \Delta^{-1} \cdot (\beta_i - \beta) \end{aligned} \quad (10)$$

soit

$$\hat{\sigma}_i^2 = \frac{1}{T} (y_i - X_i \beta_i)^\top \cdot (y_i - X_i \beta_i) \quad (11)$$

¹⁵ Voir Li, Maddala et Trost (1996).

$$\widehat{\Delta} = \frac{1}{N} \sum_{i=1}^N (\widehat{\beta}_i - \widehat{\beta}) \cdot (\widehat{\beta}_i - \widehat{\beta})^T \quad (12)$$

$$\widehat{\beta} = \frac{1}{N} \sum_{i=1}^N \widehat{\beta}_i \quad (13)$$

et

$$\widehat{\beta}_i = \left[\widehat{\Delta}^{-1} + \frac{1}{\widehat{\sigma}_i^2} \mathbf{X}_i^T \mathbf{X}_i \right]^{-1} \left[\widehat{\Delta}^{-1} \widehat{\beta} + \frac{1}{\widehat{\sigma}_i^2} \mathbf{X}_i^T \mathbf{X}_i \widehat{\mathbf{b}}_i \right] \quad (14)$$

qui peut encore s'écrire:

$$\widehat{\beta}_i = \widehat{\beta} + \widehat{\Delta} \mathbf{X}_i^T \left[\mathbf{X}_i \widehat{\Delta} \mathbf{X}_i^T + \widehat{\sigma}_i^2 \mathbf{I}_T \right]^{-1} \left[\mathbf{y}_i - \mathbf{X}_i \widehat{\beta} \right] \quad (15)$$

La méthode du Maximum de Vraisemblance nous fournit donc directement une estimation du vecteur β_i en plus de celles de β , σ_i^2 et Δ . Afin d'obtenir les estimateurs du Maximum de vraisemblance $\widehat{\beta}_i$, $\widehat{\Delta}$ et $\widehat{\sigma}_i^2$, il est nécessaire d'effectuer une première étape à l'aide des estimateurs des MCO:

$$\begin{aligned} \widehat{\sigma}_i^2 &= \frac{1}{T} (\mathbf{y}_i - \mathbf{X}_i \widehat{\mathbf{b}}_i)^T \cdot (\mathbf{y}_i - \mathbf{X}_i \widehat{\mathbf{b}}_i) \\ \widehat{\beta} &= \frac{1}{N} \sum_{i=1}^N \widehat{\mathbf{b}}_i, \text{ où } \widehat{\mathbf{b}}_i = (\mathbf{X}_i^T \mathbf{X}_i)^{-1} \mathbf{X}_i^T \mathbf{y}_i \\ \widehat{\Delta} &= \frac{1}{N} \sum_{i=1}^N (\widehat{\mathbf{b}}_i - \widehat{\beta}) \cdot (\widehat{\mathbf{b}}_i - \widehat{\beta})^T \end{aligned}$$

et

$$\widehat{\beta}_i = \left[\widehat{\Delta}^{-1} + \frac{1}{\widehat{\sigma}_i^2} \mathbf{X}_i^T \mathbf{X}_i \right]^{-1} \left[\widehat{\Delta}^{-1} \widehat{\beta} + \frac{1}{\widehat{\sigma}_i^2} \mathbf{X}_i^T \mathbf{X}_i \widehat{\mathbf{b}}_i \right]$$

Une fois, ces estimateurs calculés, on peut déterminer $\widehat{\sigma}_i^2$, $\widehat{\Delta}$ et $\widehat{\beta}_i$:

$$\begin{aligned} \widehat{\sigma}_i^2 &= \frac{1}{T} (\mathbf{y}_i - \mathbf{X}_i \widehat{\beta}_i)^T \cdot (\mathbf{y}_i - \mathbf{X}_i \widehat{\beta}_i) \\ \widehat{\beta} &= \frac{1}{N} \sum_{i=1}^N \widehat{\beta}_i \\ \widehat{\Delta} &= \frac{1}{N} \sum_{i=1}^N (\widehat{\beta}_i - \widehat{\beta}) \cdot (\widehat{\beta}_i - \widehat{\beta})^T \\ \widehat{\beta}_i &= \left[\widehat{\Delta}^{-1} + \frac{1}{\widehat{\sigma}_i^2} \mathbf{X}_i^T \mathbf{X}_i \right]^{-1} \left[\widehat{\Delta}^{-1} \widehat{\beta} + \frac{1}{\widehat{\sigma}_i^2} \mathbf{X}_i^T \mathbf{X}_i \widehat{\mathbf{b}}_i \right] \end{aligned}$$

5.2 Les estimateurs bayésiens

Il y a de fortes ressemblances entre l'approche des modèles à coefficients aléatoires et l'approche bayésienne

de modèles à coefficients fixes. Un partisan de l'approche bayésienne, analysant un modèle à coefficients fixes β , supposera que la valeur courante du vecteur de coefficients est déterminée par une expérience aléatoire. Autrement dit, le vecteur β est une variable aléatoire ayant une certaine distribution de probabilité *a priori*. On choisira alors une densité de probabilité *a priori* $p(\beta, \theta)$ associée aux paramètres d'intérêt du modèle β et $\theta (= \sigma_i^2, \Delta)$. En utilisant le théorème de Bayes, on peut alors déterminer la densité jointe *a posteriori* des paramètres:

$$p(\beta, \theta | y) \propto p(\beta, \theta) \cdot L(\beta, \theta | y) \quad (16)$$

où $p(\beta, \theta | y)$ est la densité jointe *a posteriori* et $L(\beta, \theta | y)$ est la fonction de vraisemblance. Différents auteurs ont utilisé une approche bayésienne pour estimer les modèles à coefficients aléatoires parce que la plupart des estimateurs basés sur une approche d'échantillonnage peuvent être obtenus à partir d'un cadre bayésien (Lindley et Smith (1972), Hsiao (1996)). De plus, l'approche bayésienne précise le rôle de l'information *a priori* ou bien de l'ignorance qui caractérise les paramètres θ de la distribution de y . Dans l'approche d'échantillonnage, l'économètre dispose des données mais ne connaît pas θ , il utilise $f(y | \theta)$ indirectement pour faire de l'inférence sur la valeur de θ étant donné y . A l'inverse dans le cadre bayésien, les paramètres θ sont traités comme des variables aléatoires et l'ignorance relative à θ nécessite la spécification d'un *a priori* diffus (ou matrice de variances-covariances *a priori*) du type: $p(\theta) \propto Cte$. Dans le modèle de Swamy, le fait que le vecteur β_i ait une moyenne constante β et une matrice de variances-covariances Δ revient à spécifier une information *a priori* sur le vecteur de paramètres β .

Dans les modèles à coefficients aléatoires, l'intérêt est souvent centré sur les paramètres de moyenne β et de dispersion σ_i^2 ainsi que sur la mesure de l'hétérogénéité Δ . Certains auteurs se sont intéressés à tel ou tel paramètre. Ainsi, Pesaran et Smith (1995) centrent leur étude sur l'estimation de β , Mairesse et Griliches (1990) s'intéressent à β et Δ .

Dans le cadre bayésien, on fait également l'hypothèse que le vecteur β_i est distribué selon une loi normale de moyenne β et de matrice de variances-covariances Δ :

$$\beta_i \sim N(\beta, \Delta) \quad (17)$$

Cette équation (17) définit la distribution *a priori* de β_i . Comme cette distribution dépend des paramètres inconnus β et Δ , on doit faire des hypothèses sur ces hyper-paramètres (β, Δ). On pourra ensuite en déduire la distribution *a posteriori* de β_i . Si β, σ_i^2 et Δ étaient connus, la distribution *a posteriori* de β_i serait donnée

par la loi normale d'espérance:

$$\hat{\beta}_i^* = \left[\Delta^{-1} + \frac{1}{\sigma_i^2} \mathbf{X}_i^\top \mathbf{X}_i \right]^{-1} \left[\Delta^{-1} \beta + \frac{1}{\sigma_i^2} \mathbf{X}_i^\top \mathbf{X}_i \hat{\mathbf{b}}_i \right] \quad (18)$$

et de variance:

$$V[\hat{\beta}_i^*] = \left[\Delta^{-1} + \frac{1}{\sigma_i^2} \mathbf{X}_i^\top \mathbf{X}_i \right]^{-1} \quad (19)$$

où $\hat{\mathbf{b}}_i$ est l'estimateur des MCO de β_i . La moyenne de la distribution *a posteriori* de β est définie par:

$$\hat{\beta}^* = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i^* \quad (20)$$

5.2.1 L'estimateur bayésien itératif

Comme, en général, σ_i^2 et Δ sont inconnus, il faut spécifier des *a priori* sur ces hyper-paramètres. *Smith* (1973) a proposé pour Δ^{-1} une distribution conjuguée de *Wishart* et pour σ_i^2 , des inverses de distribution du χ^2 indépendantes (*Lindley et Smith* (1972)). Il utilise comme estimateur le mode de la distribution *a posteriori*:

$$\hat{\sigma}_i^2 = \frac{1}{T + \varsigma_i + 2} \left[\varsigma_i \lambda_i + (\mathbf{y}_i - \mathbf{X}_i \hat{\beta}_i^*)^\top \cdot (\mathbf{y}_i - \mathbf{X}_i \hat{\beta}_i^*) \right] \quad (21)$$

et

$$\hat{\Delta}^* = \frac{1}{N - k - 2 + \delta} \left[\mathbf{R} + \sum_{i=1}^N (\hat{\beta}_i^* - \hat{\beta}^*) \cdot (\hat{\beta}_i^* - \hat{\beta}^*)^\top \right] \quad (22)$$

où ς_i , λ_i , δ et \mathbf{R} sont des paramètres issus de la spécification des lois *a priori*. *Smith* (1973) propose d'utiliser des approximations de ces paramètres en posant $\varsigma_i = 0$, $\delta = 1$ et \mathbf{R} est une matrice diagonale dont les éléments diagonaux sont choisis très petits (= 0.001). Les estimateurs sont donc:

$$\hat{\sigma}_i^2 = \frac{1}{T + 2} (\mathbf{y}_i - \mathbf{X}_i \hat{\beta}_i^*)^\top \cdot (\mathbf{y}_i - \mathbf{X}_i \hat{\beta}_i^*) \quad (23)$$

$$\hat{\Delta}^* = \frac{1}{N - k - 1} \left[\mathbf{R} + \sum_{i=1}^N (\hat{\beta}_i^* - \hat{\beta}^*) \cdot (\hat{\beta}_i^* - \hat{\beta}^*)^\top \right] \quad (24)$$

$$\hat{\beta}_i^* = \left[\hat{\Delta}^{*-1} + \frac{1}{\hat{\sigma}_i^2} \mathbf{X}_i^\top \mathbf{X}_i \right]^{-1} \left[\hat{\Delta}^{*-1} \hat{\beta}^* + \frac{1}{\hat{\sigma}_i^2} \mathbf{X}_i^\top \mathbf{X}_i \hat{\mathbf{b}}_i \right] \quad (25)$$

et

$$\hat{\beta}^* = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i^* \quad (26)$$

Les relations (23) à (26) doivent être estimées itérativement. Pour initialiser la procédure itérative, il faut utiliser les MCO. A la première itération, on calcule tout d'abord la valeur moyenne à l'aide de l'estimateur $\hat{\mathbf{b}}_i$,

puis $\hat{\sigma}_i^2$ et $\hat{\Delta}^*$ afin de pouvoir estimer $\hat{\beta}_i^*$

$$\begin{aligned}\hat{\beta}^* &= \frac{1}{N} \sum_{i=1}^N \hat{b}_i \\ \hat{\sigma}_i^2 &= \frac{1}{T+2} (y_i - X_i \hat{b}_i)^\top \cdot (y_i - X_i \hat{b}_i) \\ \hat{\Delta}^* &= \frac{1}{N-k-1} \left[R + \sum_{i=1}^N (\hat{b}_i - \hat{\beta}^*) \cdot (\hat{b}_i - \hat{\beta}^*)^\top \right] \\ \hat{\beta}_i^* &= \left[\hat{\Delta}^{*-1} + \frac{1}{\hat{\sigma}_i^2} X_i^\top X_i \right]^{-1} \left[\hat{\Delta}^{*-1} \hat{\beta}^* + \frac{1}{\hat{\sigma}_i^2} X_i^\top X_i \hat{b}_i \right]\end{aligned}$$

De l'itération 2 à l'itération M , les estimateurs sont:

$$\begin{aligned}\hat{\beta}_{i,iter}^* &= \frac{1}{N} \sum_{i=1}^N \hat{\beta}_{i,iter-1}^* \\ \hat{\sigma}_{i,iter}^2 &= \frac{1}{T+2} (y_i - X_i \hat{\beta}_{i,iter-1}^*)^\top \cdot (y_i - X_i \hat{\beta}_{i,iter-1}^*) \\ \hat{\Delta}_{i,iter}^* &= \frac{1}{N-k-1} \left[R + \sum_{i=1}^N (\hat{\beta}_{i,iter-1}^* - \hat{\beta}^*)^\top \cdot (\hat{\beta}_{i,iter-1}^* - \hat{\beta}^*) \right] \\ \hat{\beta}_{i,iter}^* &= \left[\hat{\Delta}_{i,iter}^{*-1} + \frac{1}{\hat{\sigma}_{i,iter}^2} X_i^\top X_i \right]^{-1} \left[\hat{\Delta}_{i,iter}^{*-1} \hat{\beta}_{i,iter}^* + \frac{1}{\hat{\sigma}_{i,iter}^2} X_i^\top X_i \hat{b}_i \right]\end{aligned}$$

Cet estimateur est basé sur celui des MCO, les estimateurs $\hat{\sigma}_i^2$ et $\hat{\Delta}^*$ sont sans biais uniquement dans le cas où X_i ne contient que des variables exogènes.

5.2.2 L'estimateur bayésien empirique itératif

Li, Maddala et Trost (1996) ont proposé un estimateur bayésien empirique itératif, qui diffère de celui de *Smith* (1973) uniquement par les dénominateurs apparaissant dans $\hat{\sigma}_i^2$ et $\hat{\Delta}^*$:

$$\hat{\sigma}_i^2 = \frac{1}{T-k} (y_i - X_i \hat{\beta}_i^*)^\top \cdot (y_i - X_i \hat{\beta}_i^*) \quad (27)$$

$$\hat{\Delta}^* = \frac{1}{N-1} \left[R + \sum_{i=1}^N (\hat{\beta}_i^* - \hat{\beta}^*) \cdot (\hat{\beta}_i^* - \hat{\beta}^*)^\top \right] \quad (28)$$

De plus, *Hu et Maddala* (1994) ont montré, à l'aide de simulations de type Monte-Carlo, que les procédures itératives pour estimer Δ donnent de meilleures estimations (au sens de l'erreur quadratique moyenne) à la fois pour β et pour la matrice de variances-covariances Δ traduisant l'hétérogénéité des comportements. De plus, les procédures à deux étapes des MCQG (*i.e.* estimateur de *Swamy*) ne permettent pas d'obtenir des estimations efficaces de β et de β_i en présence de variables endogènes retardées (*Maddala* (1971), (1991) et *Pagan* (1986)).

En résumé, supposer que les paramètres d'un modèle ne sont probablement pas exactement identiques —

mais qu'il existe une similitude entre eux — revient à admettre que ces paramètres sont issus d'une distribution jointe de même espérance et de matrice de variances-covariances non nulle. Les estimations qui en résultent sont une moyenne pondérée de l'estimation sur l'échantillon empilé et des estimations séparées sur séries temporelles. On dit que chaque estimation individuelle est "rétrécie" pour se rapprocher de l'estimation empilée globale. On parle alors "d'estimateurs à rétrécisseur" (*"shrinkage estimators"*). Cette notion d'estimateur à rétrécisseur¹⁶ (Copas (1983), Rac (1987) et Rubin (1980)) est souvent utilisée dans la littérature relative à la prévision parce que les prédictions effectuées, avec ce type d'estimateur, et plus particulièrement avec les estimateurs bayésiens, sont bien meilleures que celles réalisées avec d'autres estimateurs.

¹⁶ Il existe d'autres estimateurs à rétrécisseur tels que celui de Stein (Zeimer et Wetzstein (1983)).

6. Références

- Berri A., F. Gardes, J.L. Madre et Ch. Starzec (1998), "Modélisation démographique de la motorisation et de l'usage de l'automobile dans différents pays", *Journées de Microéconomie Appliquée*, Pointe-à-Pitre.
- Boulahbal M. et J.L. Madre (1999), "Long term forecasting of mobility in relation to urban sprawl and demographic evolution", *Conférence de l'Union Internationale des Transports Publics*, Toronto.
- Bresson G., J.L. Madre et A. Pirotte (1997), "Comparison of forecast performances of car traffic: Time series analysis versus panel data econometrics", International Association on Travel Behavior Research, Austin (à paraître dans *Transportation Research*, A).
- Bresson G. et A. Pirotte (1995), *Econométrie des Séries Temporelles: Théorie et Applications*, PUF, coll. Economie, Paris.
- Bresson G. et A. Pirotte (1999), "Modélisation des tendances à long terme des trafics TCU", *Rapport de convention ERMES-VIA GTI*.
- Bussière Y., J. Armoogum et J.L. Madre (1995), "Demographic dynamics of mobility in urban areas: a case study of Paris and Grenoble", *World Conference on Transportation Research Proceedings*, Sydney.
- Copas J.B. (1983), "Regression, prediction and shrinkage", *Journal of The Royal Statistical Society*, B, 45, 311-354.
- Cribier F. (1997), "Vieillesse et changement social", *Colloque INRETS Villes et Vieillesse*, 8-10 octobre, Arles.
- Dargay J. et M. Hanly (1999), "Bus fare elasticities: report to the Department of the Environment, Transport and the Regions", ESRC Transport Studies Unit, *University College London*.
- Gallez C. (1994), "Modèles de projection à long terme de la structure du parc et du marché de l'automobile", *Thèse pour le Doctorat en Sciences Economiques, Université Paris I Panthéon-Sorbonne*.
- Gallez C. et J.L. Madre (1993), "Démêler les facteurs structurels et économiques: la dynamique de l'usage de l'automobile", *Journées de Microéconomie Appliquée*, Sfax.
- Greene W.H. (1990), *Econometric Analysis*, MacMillan, New York.
- Hsiao C. (1986), *Analysis of Panel Data*, Cambridge University Press, Cambridge, MA.
- Hsiao C. (1996), "Random coefficients models", in *The Econometrics of Panel Data: A Handbook of the Theory with Applications*, Matyas L. et P. Sevestre, Kluwer Academic Publishers, Dordrecht, 77-99.
- Hu W. et G.S. Maddala (1994), "Estimation and prediction problems in dynamic heterogeneous panel data

- models", Working paper, Department of Economics, *The Ohio State University*.
- Hurwicz L. (1950), "Systems with non-additive disturbances", in *Statistical Inference in Dynamic Economic Models*, Koopmans T.C, eds, John Wiley, New York, 330-372.
- Kessler D. et A. Masson. (1985), "Petit guide pour décomposer un phénomène en termes d'effets d'âge, de cohorte et de moment", in *Cycles de Vie et Générations*, Kessler D. et A. Masson, Economica.
- Klein L.R (1953), *A Textbook of Econometrics*, Row Peterson, Evanston, Illinois.
- Klein L.R (1989), "Developments and prospects in macroeconomic modeling", *Eastern Economic Journal*, 15, 287-304.
- Li H., Maddala G.S et R.P Trost (1996), "Estimation des élasticités de court et de long termes de la demande d'électricité sur données de panel à partir d'estimateurs à rétrécisseur", *Economie et Prévision*, 126, 127-142.
- Lindley D.V et A.F.M Smith (1972), "Bayes estimates for the linear model", *Journal of The Royal Statistical Society*, B, 34, 1-41.
- Maddala G.S (1971), "Generalized least squares with an estimated variance covariance matrix", *Econometrica*, 1, .
- Maddala G.S (1991), "To pool or not to pool: that is the question", *Journal of Quantitative Economics*, 7, 255-264.
- Madre J.L. (1999), "Les nouveaux captifs de l'automobile", *Les Cahiers de l'IAURIF*, 122, 29-36.
- Madre J.L. et J. Maffre (1997), "La mobilité des résidents français: panorama général et évolution", *Recherche, transport Sécurité*, 56.
- Madre J.L. et A. Pirotte (1997), "Regionalisation of car-fleet and traffic forecasts", in *Understanding Travel Behaviour in an Era of Change*, P. Stopher and M. Lee-Gosselin Eds., Pergamon Press.
- Mairesse J. et Z. Griliches (1990), "Heterogeneity in panel data: are there stable production functions?", *Essays in Honor of Edmond Malinvaud*, M.I.T Press, Cambridge, Mass.
- Nicholls D.F et B.G Quinn (1982), *Random Coefficients Autoregressive Models: An Introduction*, Springer-Verlag, Berlin.
- Pagan A.(1986), "Two stages and related estimators and their applications", *Review of Economic Studies*, 53, 517-538.
- Pesaran M.H et R. Smith (1995), "Estimating long-run relationships from dynamic heterogeneous panels", *Journal of Applied Econometrics*, 68, 79-113.
- Quinet E.(1971), *Analyse Economique des Transports*, R.U.F. Coll. Economie, Paris.
- Raj B. et A. Ullah (1981), *Econometrics: A Varying Coefficients Approach*, Crom-Helm, Londres.
- Rao C.R (1975), "Simultaneous estimation of parameters in different linear models and applications to bio-

metric problems", *Biometrics*, 31, 545-554.

Rao C.R (1987), "Prediction of future observations in growth curve models", *Statistical Science*, 2(4), 434-471.

Rubin D.B (1980), "Using empirical Bayes techniques in the Law School validity studies", *Journal of The American Statistical Association*, 75, 801-827.

Smith A.E.M (1973), "A general bayesian linear model", *Journal of The Royal Statistical Society*, B, 35, 67-75.

Swamy P.A.VB (1970), "Efficient inference in a random coefficient regression model", *Econometrica*, 38, 311-323.

Swamy P.A.VB (1971), *Statistical Inference in Random Coefficient Regression Models*, Springer-Verlag, Berlin.

Swamy P.A.VB et G.S Tavlas (1995), "Random coefficients models: theory and applications", *Journal of Economic Surveys*, 9(2), 165-196.

Wiel M. (1999), *La Transition Urbaine ou le Passage de la Ville Pédestre à la Ville Motorisée*, Editions Mardaga, Coll. Architecture+ Recherches, Sprimont, Belgique.

Zeimer R.F et M.E Wetzstein (1983), "An Stein-rule method for pooling data", *Economics Letters*, 1, 137-143.

Zellner A. (1969), "On the aggregation problem", in *Economic Models, Estimation and Risk Programming: Essays in Honor of Gerhard Tintner*, Fox K.A, Sengupta J.K et G.V.L Narasimham, eds, Springer-Verlag, New York, 330-372.

Moyennes Individuelles (1975-1995)

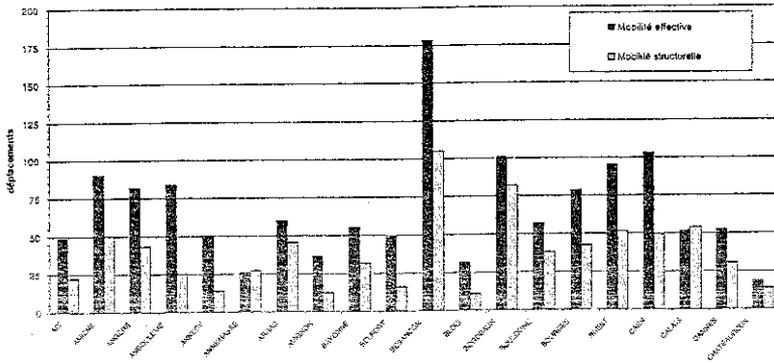


Figure 1a.

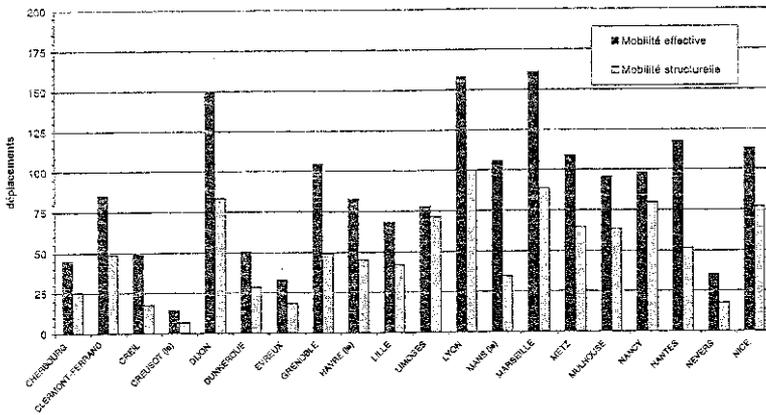
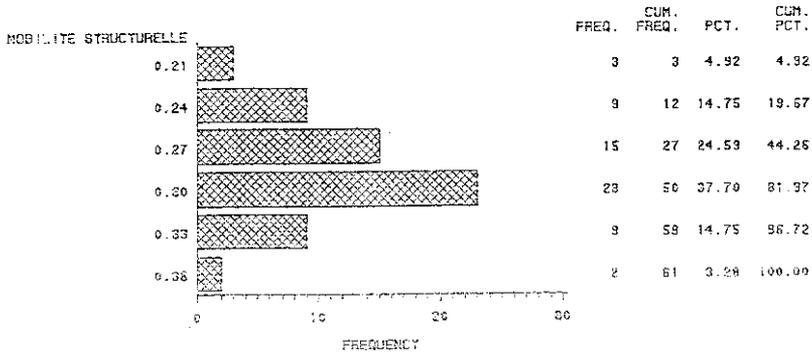


Figure 1b.

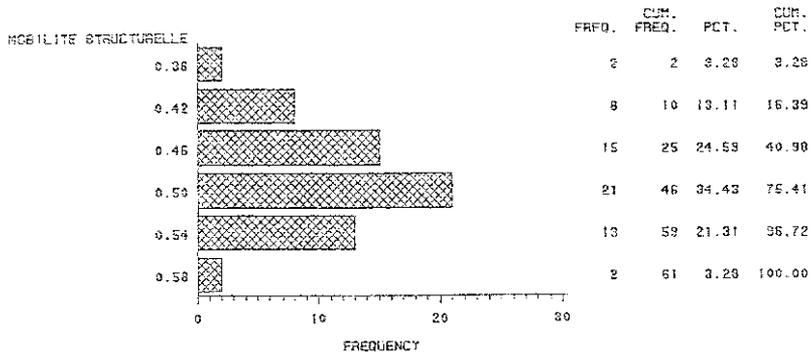
ELASTICITES DE COURT TERME (A-5)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 3a.

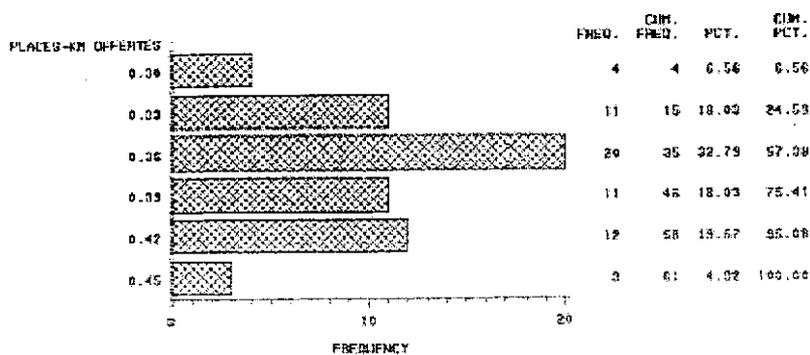
ELASTICITES DE LONG TERME (A-5)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 3b.

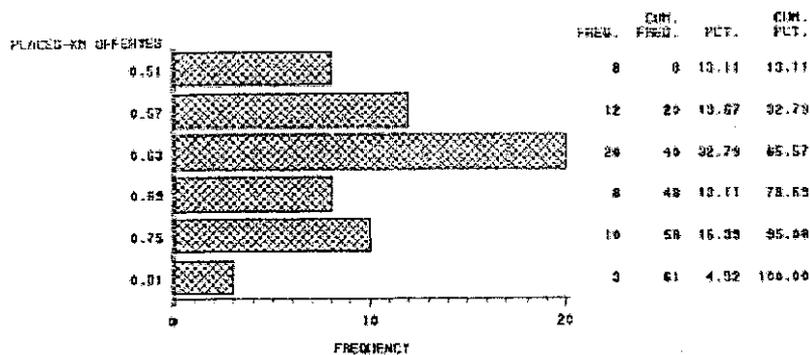
ELASTICITES DE COURT TERME (A--C)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 4a.

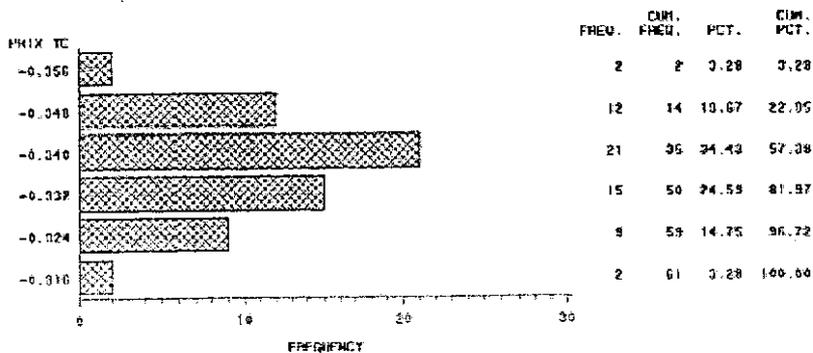
ELASTICITES DE LONG TERME (A--C)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 4b.

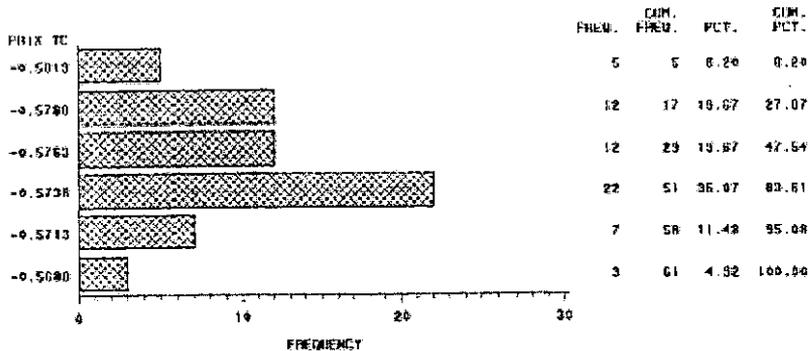
ELASTICITES DE COURT TERME (A-D)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 5a.

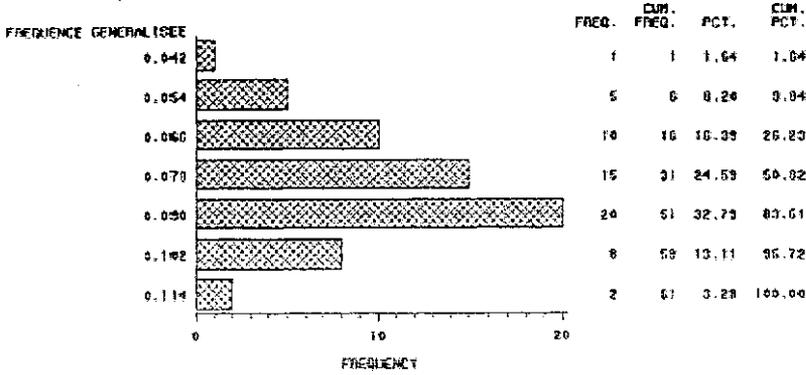
ELASTICITES DE LONG TERME (A-D)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 5b.

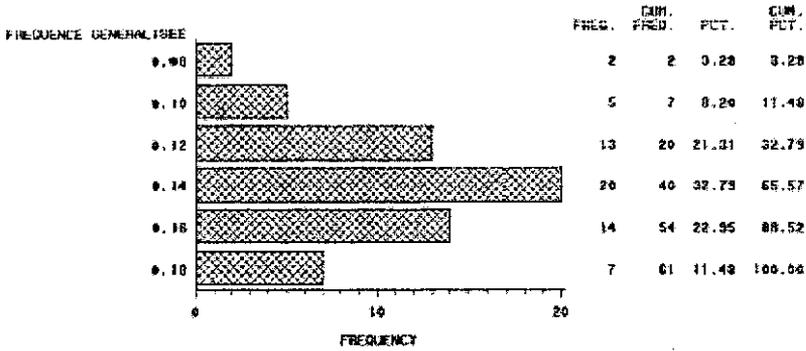
ELASTICITES DE COURT TERME (A-D)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 6a.

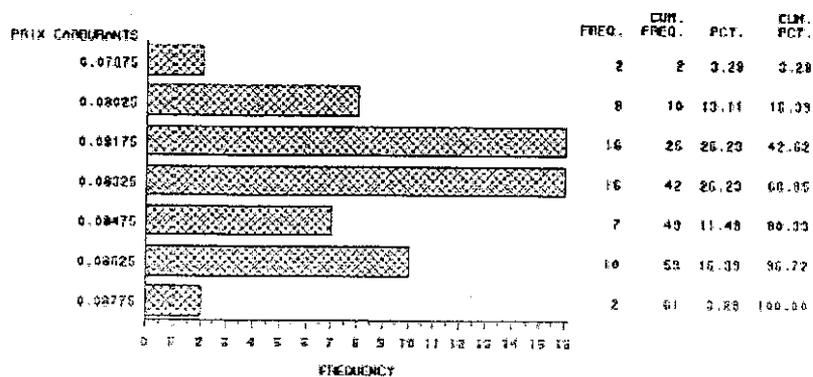
ELASTICITES DE LONG TERME (A-D)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 6b.

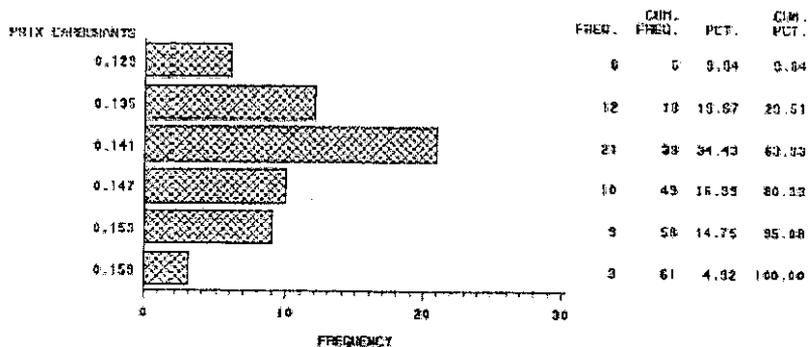
ELASTICITES DE COURT TERME (A-3)



REPARTITION DES PIV SELON LES ELASTICITES

Figure 7a.

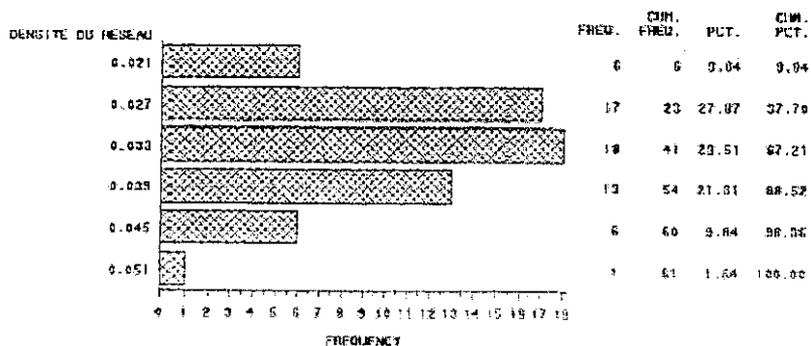
ELASTICITES DE LONG TERME (A-5)



REPARTITION DES PIV SELON LES ELASTICITES

Figure 7b.

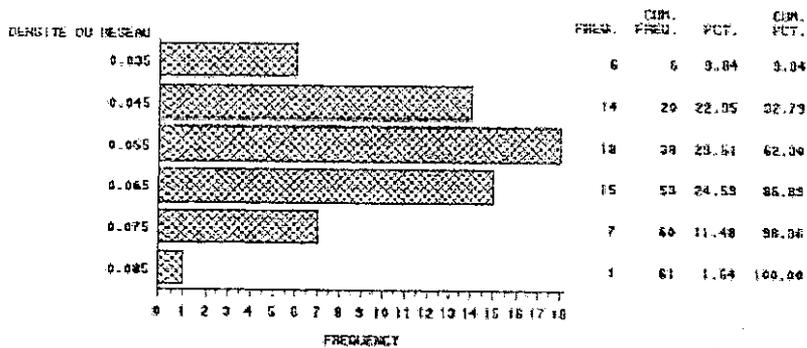
ELASTICITES DE COURT TERME (A-D)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 8a.

ELASTICITES DE LONG TERME (A-D)



REPARTITION DES PTU SELON LES ELASTICITES

Figure 8b.

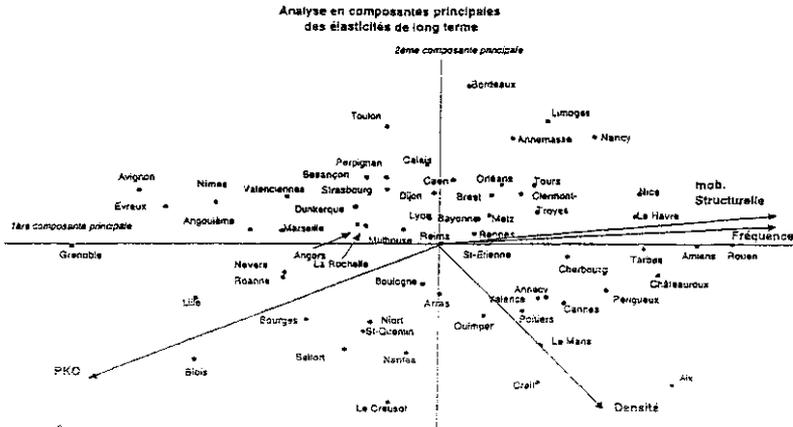


Figure 9.

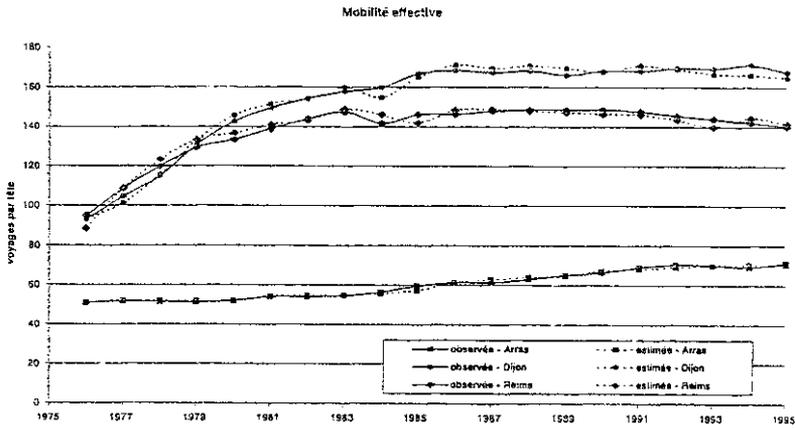


Figure 10a.

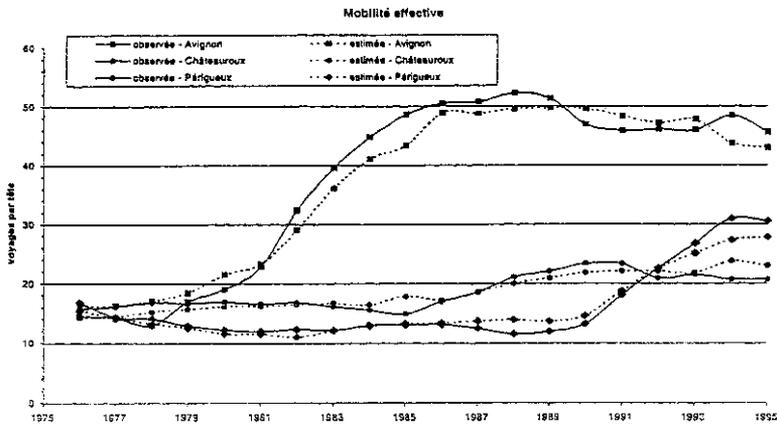


Figure 10b.

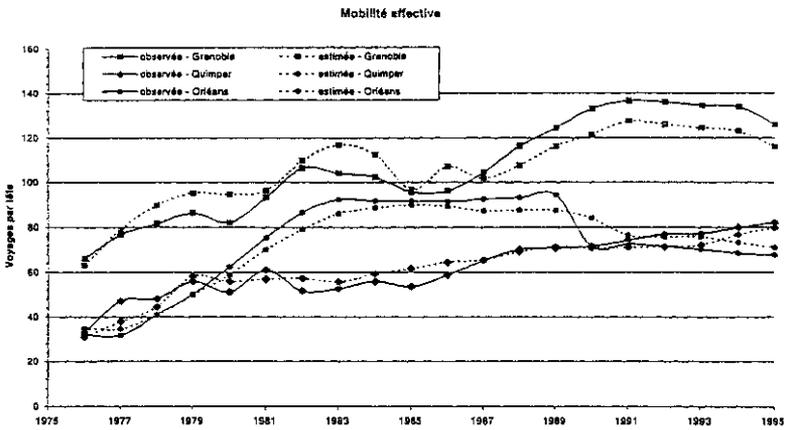


Figure 10c.