

Déterminants du trafic des véhicules légers et élasticités : une approche spatiale sur données régionales françaises

Alain Pirotte * et Jean-Loup Madre **

Traditionnellement, les trafics observés sont expliqués à l'intérieur d'une zone (ici l'une des 21 régions françaises) par les caractéristiques de ses habitants : le revenu réel, le taux d'équipement en automobiles, etc. Or, les trafics, ici celui des véhicules légers, impliquent par nature des échanges entre régions plus ou moins proches. L'économétrie spatiale sur données de panel permet de mettre en évidence ces interactions et de fournir, à partir de séries régionales remontant à 1985, de nouvelles estimations des élasticités du trafic des véhicules légers exprimé en kilométrage annuel moyen par véhicule au prix des carburants, au revenu réel et au taux d'équipement.

Plusieurs spécifications spatiales sont retenues, à savoir les modèles spatiaux auto-régressifs à effets aléatoires et à effets fixes. Ils sont estimés grâce à l'estimateur du maximum de vraisemblance basé sur une procédure itérative à deux étapes. Les résultats obtenus révèlent des interdépendances spatiales significatives qui caractérisent le trafic régional des véhicules légers. L'avantage de l'introduction de la dimension spatiale est qu'elle réduit les biais de spécification potentiels qui peuvent affecter de façon importante les élasticités estimées. Par ailleurs, ce type de modèle autorise l'estimation des effets directs, indirects et totaux associés au prix des carburants, au revenu réel et au taux d'équipement.

Globalement, les bornes des intervalles de confiance à 95 % des élasticités directes du prix des carburants, du taux d'équipement et du revenu réel sont respectivement [- 0,11 ; - 0,07], [- 0,36 ; - 0,23] et [0,38 ; 0,50]. L'essor de la seconde voiture pousse à la baisse de l'utilisation de chaque véhicule, d'où une élasticité négative au taux d'équipement, qui s'est un peu renforcée au cours du temps.

Rappel :

Les jugements et opinions exprimés par les auteurs n'engagent qu'eux mêmes, et non les institutions auxquelles ils appartiennent, ni a fortiori l'Insee.

* *ERMES-TEPP (CNRS), Université Panthéon-Assas Paris II / Sorbonne universités et IFSTTAR-DEST, Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux.*

** *IFSTTAR-DEST, Institut Français des Sciences et Technologies des Transports, de l'Aménagement et des Réseaux.*
Nous tenons à remercier deux rapporteurs anonymes pour leurs commentaires et suggestions.

Après un essor quasi-continu depuis la seconde guerre mondiale, la circulation en France des véhicules légers - ceux que l'on peut conduire avec un permis B - n'évolue plus de manière significative depuis 2002 (CCTN¹, 2010) ; compte tenu d'une meilleure efficacité énergétique (la consommation unitaire est passée de 8,7 l/100 km en 1990 à 7,4 l/100 km en 2007), ceci correspond à une lente diminution du volume de carburant utilisé. Or, depuis le tournant du siècle, le parc automobile continue à croître d'un peu moins de 2 % par an, principalement tiré par le multi-équipement. Chaque véhicule est donc moins utilisé : le kilométrage annuel moyen recule. Il était de 13 758 km en 1990 contre 13 029 km en 2007 pour les voitures particulières, qui représentaient respectivement 85 % et 84 % du parc automobile (hors poids lourds)².

Même si la diffusion sociale a amoindri les écarts d'équipement et de kilométrages selon les niveaux de vie (Boucq, Collet, Madre et Hivert, 2011), l'évolution du revenu reste un facteur de croissance prépondérant de la circulation automobile. Le ralentissement de la croissance du parc et des kilométrages depuis le milieu des années 1990 interroge sur l'émergence et/ou le renforcement d'autres facteurs explicatifs pour rendre compte de l'évolution du trafic, notamment la volatilité du prix des carburants. La période couverte par cet article s'étale de 1985 à 2007. Commencant au contre-choc pétrolier, elle est située à la charnière entre une phase de stabilité du prix des carburants et une phase de volatilité à la hausse amorcée par la pointe de l'année 2000, et poursuivie par le renchérissement continu de 2004 à mi-2008.

Les flux de trafic et les interactions régionales

Les déterminants de la circulation sont nombreux et complexes ; ils sont reliés à la structure spatiale (densité, urbanisation, etc.). En effet, la nécessité de l'automobile est différente à la campagne, où services de proximité et transports en commun ont pratiquement disparu, et en centre-ville, où la plupart des destinations peuvent être reliées à pied ou en vélo, et où les transports collectifs sont disponibles pour aller plus loin, alors qu'il est difficile de se garer et de circuler aux heures de pointe. C'est donc à l'intérieur des bassins de vie que les comportements sont les plus contrastés entre habitants des villes-centre, des banlieues et des périphéries. Toutefois, les données administratives

nécessaires pour constituer des séries temporelles fiables ne sont disponibles qu'à une plus grande échelle géographique. La dimension régionale est un bon compromis, puisqu'elle oppose des régions très urbanisées (Île-de-France, Nord-Pas de Calais, Rhône-Alpes, etc.) et des régions à dominante rurale (Limousin, etc.). Du point de vue des trafics, certaines régions sont enclavées (Auvergne, Bretagne, etc.), alors que d'autres connaissent des flux importants de transit, avec un fort dynamisme des trafics à longue distance (Bourgogne entre Lyon et Paris, Aquitaine et Languedoc-Roussillon vers l'Espagne, etc. ; Madre, 1994). En matière de politiques publiques, les régions ont acquis en 2010 une certaine marge de manœuvre pour fixer le taux de la TICPE³ et, même si le rythme des constructions routières et autoroutières devrait diminuer suite au Grenelle de l'environnement, il sera toujours nécessaire d'actualiser régulièrement les prédictions de trafic à long terme, notamment à l'appui des politiques d'entretien routier des départements à qui la décentralisation a confié l'essentiel du réseau antérieurement national. Dans ce cadre, la modélisation du trafic automobile, en particulier celle des véhicules légers, et l'évaluation de ses déterminants constituent un enjeu important. L'économétrie spatiale sur données de panel permet, non seulement de prendre en compte l'hétérogénéité régionale observée et inobservable, mais également d'évaluer les interactions régionales, pour fournir de nouvelles estimations des élasticités du trafic des véhicules légers au prix des carburants, au revenu réel et au taux d'équipement.

L'intérêt de la dimension panel

Un nombre croissant de recherches visent à modéliser les trafics routiers (Trace, 1998 ; De Jong et Gun, 2001 ; Graham et Glaister, 2004) ou la demande de carburant (Dahl, 1986 ; Dahl et Sterner, 1991a et b ; Graham et Glaister, 2002). Bien qu'il y ait un large consensus sur l'ordre de grandeur des élasticités du trafic routier (Goodwin, 1992 ; Goodwin *et al.*, 2004), il reste beaucoup à apprendre sur la variation de ces élasticités entre catégories de consommateurs (distingués par exemple en fonction de leur localisation) et sur leurs évolutions au cours du temps (en particulier lors des chocs et

1. Commission des comptes des transports de la Nation.

2. Source : CCTN (2010).

3. Depuis le 1^{er} janvier 2012, la taxe intérieure de consommation sur les produits énergétiques (TICPE) remplace la TIPP (taxe intérieure sur les produits pétroliers).

contre-chocs ou de périodes plus continues de renchérissement des carburants (Kemel *et al.*, 2011). Des méta-analyses des élasticités du trafic routier aux revenus et aux prix montrent que les différences entre les résultats des différentes études sont souvent liées aux caractéristiques des données utilisées (Espey, 1997, 1998 ; Brons *et al.*, 2008). Dans ce contexte, l'économétrie des données de panel présente bien des avantages, tout en reposant sur des hypothèses particulières, notamment quant à la représentation de la forme de l'hétérogénéité. Il est dorénavant bien établi que l'un des principaux attraits de l'économétrie de panel est de prendre en compte l'hétérogénéité inobservable entre les différents individus (Baltagi, 2008 ; Pirotte, 2011). En effet, les analyses sur séries temporelles agrégées ou en coupe transversale, qui ne contrôlent pas cette hétérogénéité, risquent de fournir des résultats biaisés (Moulton, 1986, 1987). L'hétérogénéité inobservable peut être captée par des constantes (effets fixes), par des coefficients aléatoires ou par une combinaison des deux (Hsiao, 1993). Les modèles classiques à effets fixes ou aléatoires reposent sur l'hypothèse que les coefficients estimés des variables explicatives ne diffèrent significativement ni entre les individus de la population étudiée (ici les régions de la France métropolitaine), ni au cours de la période analysée. Ces hypothèses, que l'on peut tester, sont rarement valides. Des méthodes d'estimation de modèles à coefficients hétérogènes (estimateur groupé moyen empilé⁴, estimateur bayésien itératif, estimateur bayésien hiérarchique, etc.) ont été récemment proposées (cf. Robertson et Symons, 1992 ; Pesaran et Smith, 1995 ; Pesaran *et al.*, 1996). Les estimateurs traditionnels (moindres carrés ordinaires, *Within*, *Between*, doubles moindres carrés, moments généralisés, etc.) sont clairement les seules alternatives viables pour un panel dont la dimension temporelle s'étend sur moins de dix ans. En revanche, dès que l'on atteint 25 à 50 ans de données collectées après la seconde guerre mondiale, le choix est moins évident. En outre, des corrélations transversales significatives marquent la plupart des données de panel disponibles. Cet aspect constitue un enjeu important pour toute modélisation adéquate. Dans cette perspective, l'économétrie spatiale sur données de panel est une approche fertile en plein essor depuis les années 2000.

Les modèles spatiaux - couramment utilisés en sciences régionales et en économie urbaine - gèrent les interactions (autocorrélations) et l'hétérogénéité spatiales, d'abord en coupe transversale (Anselin, 1988 ; Le Gallo, 2002, 2004 ;

Anselin *et al.*, 2008 ; LeSage et Pace, 2009 ; Elhorst, 2010 ; Pirotte, 2011). Plus précisément, l'introduction de la dimension spatiale vise à capter deux effets : l'autocorrélation spatiale qui renvoie à l'absence d'indépendance entre les observations géographiques (notamment), et l'hétérogénéité spatiale qui est liée à la différenciation des variables et des comportements dans l'espace. La structure de la dépendance spatiale peut être rattachée à la situation géographique (distance), à un réseau social, etc. Au total, l'autocorrélation spatiale reflète l'existence, l'ampleur et l'influence des effets de débordement. Si l'on considère la modélisation du trafic automobile des régions françaises, on conçoit que le trafic constaté dans une région donnée est en partie engendré par celui des régions connexes. Une telle dépendance peut être captée grâce à l'introduction de la variable endogène spatialisée dans la partie explicative du modèle (*i.e.* du modèle linéaire autorégressif spatial à erreurs composées (1)).

En économétrie des données de panel, peu de travaux empiriques ont utilisé les modèles spatiaux à erreurs composées et/ou à effets fixes, et *a priori* aucune application ne concerne le trafic automobile des véhicules légers. Il est donc tentant de tirer profit de ces développements récents pour analyser la circulation d'un point de vue géographique, sur laquelle on constate une grande variabilité entre les régions métropolitaines en France.

Une littérature riche mais peu axée sur la prise en compte de la dimension spatiale

Les travaux sur le trafic routier abondent et reposent sur différents indicateurs (le nombre de véhicules kilomètres parcourus sur le réseau routier, la consommation totale de carburant, le nombre de véhicules possédés, la distance parcourue par véhicule ou encore le stock de véhicules). Plusieurs articles reflètent cet intérêt et la richesse de cette littérature tels que ceux de Dahl (1986), Dahl et Sterner (1991a, b), Goodwin (1992), Dahl (1995), Trace (1998), De Jong et Gun (2001), Graham et Glaister (2002, 2004), Goodwin, Dargay et Hanly (2004), ou encore celui de Li, Rose et Hensher (2010). De nombreux modèles ont été retenus et estimés - allant de spécifications statiques à des modèles dynamiques, estimés à partir de méthodes économétriques plus ou moins élaborées - sur différents types de données (coupes

4. *Pooled Mean Group (PMG)*.

transversales, séries chronologiques ou données de panel) et des périodes très hétérogènes. Dans ces conditions, la palette des élasticités prix, revenu et taux d'équipement présente un large spectre. Dans un article de synthèse relativement récent, Goodwin, Dargay et Hanly - en abrégé GDH - (2004) retracent l'amplitude des élasticités prix et revenu en considérant les travaux sur le trafic routier et la consommation de carburant parus dans des revues académiques, des rapports gouvernementaux, de chercheurs et de consultants depuis 1990, soit au total 69 références. Selon GDH, les valeurs des élasticités obtenues sur données de panel présentent des écarts importants et significatifs. Plus précisément, les élasticités des véhicule-km⁵ (resp. de la consommation de carburant par véhicule) estimées sur la base d'un modèle statique varient de - 0,41 à - 0,13 (resp. de - 0,89 à - 0,004) pour le prix du carburant par litre, et de 0,05 à 1,44 (resp. de 0,07 à 1,14) pour le revenu. Ces intervalles sont approximativement identiques à ceux reportés dans des articles de synthèse antérieurs. Pour leur part, Espey (1997, 1998), Brons, Nijkamp, Pels et Rietveld (2008) utilisent des méta-analyses afin de comprendre et d'expliquer la diversité des élasticités prix et revenu entre les études. Ils cherchent à identifier les facteurs qui impactent systématiquement la valeur des élasticités prix et revenu. Espey (1997, 1998) souligne le rôle fondamental joué par les hypothèses des modèles, notamment quant à leurs structures, et les caractéristiques des données (*i.e.* séries chronologiques, coupes transversales, données de panel). Sur séries chronologiques australiennes, Li, Rose et Hensher (2010) montrent que des modèles très élaborés ne produisent pas forcément de meilleures prédictions que des spécifications moins complexes. En toute généralité, le choix du modèle et des méthodes économétriques s'avèrent d'une importance capitale. Ainsi, à un niveau désagrégé (*i.e.* inférieur à la dimension nationale), un aspect prépondérant pour la modélisation du trafic automobile consiste à ne pas négliger les dépendances spatiales entre les zones. Dans un tel contexte, la difficulté principale pour expliquer le trafic automobile dans une zone délimitée (ici chacune des 21 régions françaises) avec des variables explicatives, qui reflètent seulement les profils des habitants de cette zone (revenu et taux d'équipement), provient des trafics de transit et limitrophes (*i.e.* générés par les habitants des zones voisines et par ceux de zones plus éloignées). Les développements récents de l'économétrie spatiale sur données de panel autorisent l'introduction de ces interdépendances dans l'analyse (cf.

Anselin, Le Gallo et Jayet, 2008 ; Pirotte, 2011). Ce cadre permet en particulier d'éviter certains biais de spécification et/ou d'améliorer la précision des résultats.

La dimension spatiale du trafic automobile des véhicules légers

D'après la littérature, les principales variables explicatives qui influencent le trafic automobile (*i.e.* celui des véhicules légers (VL)) sont le prix réel du carburant, le revenu réel et le stock de véhicules. Sur données de panel, il est évident que le trafic automobile de la région i dépend de ceux des régions plus ou moins proches. Baltagi et Pirotte (2010) montrent que si la dimension spatiale est négligée, alors les estimateurs usuels appliqués sur données de panel en sont grandement affectés (*i.e.* en termes de biais et/ou précision). Ceci justifie d'autant plus de situer l'analyse dans le cadre des modèles spatiaux sur données de panel. Deux modèles spatiaux sont retenus pour modéliser le trafic régional français des VL. À partir notamment des travaux de Sweeney (1979), Griffin (1979), Baltagi et Griffin (1983, 1997), le trafic annuel moyen par véhicule (km en logarithme) dans la région i à la période t , y_{it} , est expliqué par un modèle linéaire autorégressif spatial à erreurs composées, soit :

$$y_{it} = \rho \sum_{j=1}^N w_{ij} y_{jt} + x_{it} \beta + \varepsilon_{it}$$

$$i = 1, \dots, N, \quad t = 1, \dots, T \quad (1)$$

et

$$\varepsilon_{it} = \alpha_i + u_{it} \quad (2)$$

où ρ est le coefficient autorégressif spatial, β , le vecteur des paramètres de taille ($k \times 1$), x_{it} , le vecteur de dimension ($1 \times k$) des variables explicatives (ici $k = 3$, le revenu des ménages et le nombre de véhicules par tête et le prix réel du carburant), α_i , l'effet individuel aléatoire associé à la région i , IID⁶ ($0, \sigma_\alpha^2$), u_{it} , une perturbation usuelle, IID ($0, \sigma_u^2$), et w_{ij} , la notation générique des poids de la matrice spatiale W_N . Plus précisément, W_N est une matrice spatiale dont les poids sont supposés connus, de format ($N \times N$), et dont les éléments diagonaux

5. Goodwin, Dargay et Hanly (2004) mentionnent seulement une étude qui utilise l'indicateur véhicule-km par véhicule. Nous avons préféré de ne pas reporter ces chiffres « uniques ».

6. IID : indépendamment et identiquement distribué.

sont nuls⁷. Cette matrice caractérise la façon dont les interdépendances entre les régions sont retracées.

D'après (1), le vecteur des observations de dimension $(N, 1)$ de la variable expliquée y_t à la période t est donné par :

$$y_t = \rho W_N y_t + X_t \beta + \varepsilon_t \quad (3)$$

ou

$$y_t = (I_N - \rho W_N)^{-1} X_t \beta + (I_N - \rho W_N)^{-1} \varepsilon_t \quad (4)$$

où I_N représente la matrice identité de taille N , $X_t = (x_{1t}, \dots, x_{Nt})'$, une matrice de taille $(N \times k)$, $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})'$, un vecteur de dimension $(N \times 1)$, et

$$(I_N - \rho W_N)^{-1} = I_N + \rho W_N + \rho^2 W_N^2 + \rho^3 W_N^3 + \dots \quad (5)$$

L'expression (5) caractérise un effet spatial qui décline avec l'éloignement. Ainsi, dans un modèle spatial autorégressif, l'interprétation des paramètres est plus riche mais aussi plus compliquée. Un changement dans une variable explicative d'une région peut affecter potentiellement la variable expliquée de toutes les autres régions. Plusieurs chercheurs ont noté que cette spécification nécessitait une interprétation spécifique des paramètres (cf. LeSage et Pace, 2009 ; Anselin et Le Gallo, 2006 ; Kelejian *et al.*, 2006). Ainsi, le modèle spatial autorégressif exploite l'information des régions plus ou moins limitrophes. LeSage et Pace (2009, pp. 36-42 et 114-115) suggèrent des mesures d'impact⁸ incorporant cette information spatiale sur coupes transversales. Elles sont appelées les effets moyens directs, indirects et totaux. Un aspect important est que l'interprétation de l'effet direct peut être rapprochée de celle des coefficients des régressions classiques. Le principe de ces mesures proposées par LeSage et Pace (2009) reste valable sur données de panel. La partie qui aborde plus loin les estimations des modèles spatiaux autorégressifs revient sur leur interprétation.

La structure de la matrice spatiale W_N est généralement un peu arbitraire, souvent basée sur des distances entre les zones. Anselin (1988) présente une synthèse concentrée essentiellement sur les distances géographiques. D'autres auteurs retiennent des métriques plus générales, telles que la proximité sociale (Conley et Topa,

2002), ou la proximité économique ou une variante similaire (Conley, 1999 ; Pesaran *et al.*, 2004). Dans cet article, deux matrices spatiales sont considérées. Globalement, elles reflètent le fait que les trafics des régions les plus proches de la région i accroissent son trafic. La première matrice illustre la structure spatiale basée sur les distances géographiques (cf. graphique 1). Pour les régions i et j , les poids $w_{1,ij}$ de $W_{1,N}$ sont définis par

$$w_{1,ij} = \begin{cases} d_{ij}^{-\lambda} \text{ si } d_{ij} < \bar{d} \\ 0 \text{ sinon} \end{cases} \quad (6)$$

où d_{ij} est la distance (km) entre les préfectures des régions i et j , λ , un paramètre connu, \bar{d} , un paramètre de seuil au-delà duquel les régions i et j ne sont plus connectées⁹ (cf. Anselin, 1988).

Ensuite, une matrice de contiguïté symétrique à l'ordre 1 est également utilisée (cf. graphique 2). Pour les régions i et j , les poids $w_{2,ij}$ de $W_{2,N}$ sont définis par

$$w_{2,ij} = \begin{cases} 1 \text{ si la région } i \text{ a une frontière commune} \\ \text{avec la région } j \\ 0 \text{ sinon} \end{cases} \quad (7)$$

En pratique, les matrices spatiales $W_{1,N}$ et $W_{2,N}$ sont habituellement normalisées. Autrement dit, chaque élément w_{ij} est divisé par la somme

des poids de la ligne $\sum_{j=1}^N w_{ij}$. Tous les poids sont

donc compris entre 0 et 1. Dans ce cas, leur interprétation s'en trouve modifiée. Si l'on prend le cas d'une matrice de distance, la normalisation signifie que le lien entre deux régions i et j est fonction maintenant de la distance relative et non plus de la distance absolue. Cette normalisation facilite l'interprétation des coefficients et leurs comparaisons¹⁰. Dans le modèle (1), le coefficient ρ mesure comment le trafic d'une région est relié à celui des autres régions conditionnellement aux variables explicatives. Une

7. Par convention, une région n'est pas contiguë à elle-même, $w_{ii} = 0, \forall i = j$

8. Se reporter à Abreu et al. (2005) pour des terminologies différentes (les effets directs, indirects et induits) et leurs interprétations.

9. Les valeurs retenues pour les paramètres λ et \bar{d} sont respectivement égales à 0,3 et 380 kms. D'autres valeurs plus élevées ont également été retenues pour le paramètre λ , les résultats ne s'en trouvant pas pour autant fondamentalement remis en cause.

10. Se reporter à Kelejian et Robinson (1995), Haining (1990), Anselin (1988).

valeur positive de ρ indique un impact significatif des autres trafics régionaux.

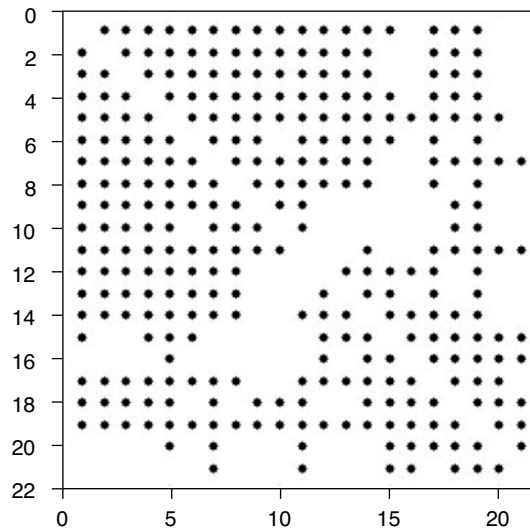
Une hypothèse forte et souvent non vérifiée en pratique dans les modèles à erreurs composées est la non corrélation entre les spécificités individuelles et les variables explicatives, *i.e.*

$$E[\alpha_i x'_{it}] = 0, \text{ pour } i = 1, \dots, N, \text{ et } t = 1, \dots, T.$$

Si cette hypothèse n'est pas vérifiée, alors les estimations du modèle à erreurs composées sont biaisées. La statistique d'Hausman¹¹ permet de tester si ce problème existe, et d'orienter l'analyse vers le choix d'un estimateur ayant des

11. Se reporter à Hausman (1978).

Graphique 1
Matrice de distance entre les régions (inverse des distances)

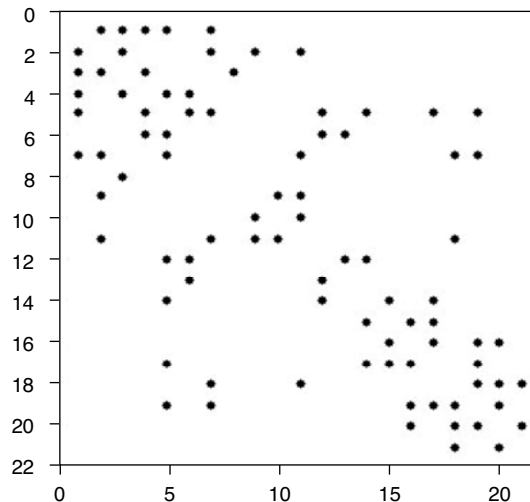


Lecture : les numéros portés en abscisse et en ordonnée désignent les 21 régions de France métropolitaine. La structure de la matrice spatiale laisse apparaître 69,4 % d'éléments non nuls.

Champ : 21 régions de France métropolitaine (la Corse est agrégée à la région PACA).

Source : calcul des auteurs.

Graphique 2
Matrice de contiguïté des régions à l'ordre 1



Lecture : les numéros portés en abscisse et en ordonnée désignent les 21 régions de France métropolitaine. La structure de la matrice spatiale laisse apparaître 19,5 % d'éléments non nuls.

Champ : 21 régions de France métropolitaine (Corse agrégée à la région PACA).

Source : calcul des auteurs.

propriétés statistiques satisfaisantes. Elle est reportée dans le tableau qui fournit les résultats d'estimation. Les spécifications qui viennent

d'être présentées sont estimées grâce à l'estimateur du maximum de vraisemblance (MV), (cf. encadré 1).

Encadré 1

LES MODÈLES SPATIAUX AUTORÉGRESSIFS UTILISÉS ET LEURS MÉTHODES D'ESTIMATION

Le modèle spatial autorégressif à erreurs composées

Matriciellement, il s'écrit :

$$y = \rho(I_T \otimes W_N)y + X\beta + \varepsilon \quad (1)$$

avec

$$\varepsilon = (e_T \otimes I_N)\alpha + u \quad (2)$$

où $y = (y_1, \dots, y_T)'$, $X = (X_1, \dots, X_T)'$, $\alpha = (\alpha_1, \dots, \alpha_N)'$ et $u = (u_1, \dots, u_T)'$. ρ est le paramètre autorégressif, et β est le vecteur des paramètres des variables explicatives X . e_T est un vecteur unitaire de dimension $(T \times 1)$. I_N et I_T sont des matrices identités respectivement de dimension N et T . W_N est la matrice spatiale de format $(N \times N)$. Sous l'hypothèse de normalité des perturbations, la log-vraisemblance est donnée par :

$$L = -\frac{NT}{2} \ln(2\pi\sigma_u^2) + \frac{N}{2} \ln \theta^2 + T \ln |B_N| - \frac{1}{2\sigma_u^2} \varepsilon' \Sigma^{-1} \varepsilon \quad (3)$$

avec

$$\varepsilon = (I_T \otimes B_N)y - X\beta, B_N = (I_N - \rho W_N),$$

$$\Omega = \sigma_u^2 \Sigma, \Sigma = Q_{0,N} + \frac{1}{\theta^2} Q_{1,N}$$

et

$$\theta^2 = \frac{\sigma_u^2}{T\sigma_\alpha^2 + \sigma_u^2}$$

$$Q_{0,N} = \left(I_T - \frac{J_T}{T} \right) \otimes I_N$$

$$Q_{1,N} = \frac{J_T}{T} \otimes I_N$$

où $J_T = e_T e_T'$. Les opérateurs matriciels $Q_{0,N}$ et $Q_{1,N}$ correspondent aux transformations habituelles (i.e. intra-individuel - *Within* - et inter-individuel - *Between*) utilisées sur données de panel. La seule différence est qu'ils sont exprimés ici par rapport à un empilement année par année de tous les individus et non par rapport à un empilement individu par individu pour toutes les années. Les paramètres β et σ_u^2 sont obtenus grâce aux conditions du premier ordre associé à la log-vraisemblance (cf. Elhorst, 2003). Les estimateurs du maximum de vraisemblance (MV) de β et σ_u^2 sont donnés par :

$$\hat{\beta} = (X'X)^{-1} X'u^* \quad (4)$$

$$\hat{\sigma}_u^2 = \frac{\varepsilon' \varepsilon}{NT} \quad (5)$$

où

$$y^* = (I_T \otimes B_N)(I_{NT} - (1-\theta)Q_{1,N})y, \quad (6)$$

$$X^* = (I_{NT} - (1-\theta)Q_{1,N})X, \quad (7)$$

$$\varepsilon^* = y^* - X^* \hat{\beta}. \quad (8)$$

Les estimateurs (4) et (5) sont tous les deux fonction des paramètres ρ et θ^2 . Grâce à (4) et (5), on concentre la log-vraisemblance (3) par rapport à ρ et θ^2 . On obtient :

$$L_C = C^* - \frac{NT}{2} \ln(\varepsilon^* \varepsilon^*) + \frac{N}{2} \ln \theta^2 + T \ln |B_N| \quad (9)$$

où $C^* = -\left(\frac{NT}{2}\right) (\ln 2\pi) - \frac{NT}{2} + \left(\frac{NT}{2}\right) (\ln(NT))$. Une procédure itérative à deux étapes est appliquée pour estimer ces paramètres. Dans une première étape, on commence par fixer des valeurs initiales aux paramètres ρ et θ^2 afin d'obtenir $\hat{\beta}$ et $\hat{\sigma}_u^2$. Dans une seconde étape, on recherche les valeurs de ρ et θ^2 qui maximisent la vraisemblance concentrée (équation 9). Il est possible d'alterner vers l'arrière et vers l'avant entre l'estimation des paramètres ρ et θ^2 conditionnellement au vecteur des perturbations ε^* , et l'estimation de β et σ_u^2 conditionnellement aux valeurs de ρ et θ^2 , ceci jusqu'à l'obtention de la convergence.

Le modèle spatial autorégressif à effets fixes

Sous forme matricielle, cette spécification s'écrit:

$$y = \rho(I_T \otimes W_N)y + X\beta + (e_T \otimes I_N)\alpha + u \quad (10)$$

Afin d'éliminer les effets spécifiques individuels, on utilise l'opérateur intra-individuel $Q_{0,N}$ précédemment défini. L'application de cette transformation à (10) donne :

$$Q_{0,N}y = \rho(I_T \otimes W_N)Q_{0,N}y + Q_{0,N}X\beta + Q_{0,N}u \quad (11)$$

Sous l'hypothèse de normalité des perturbations, la log-vraisemblance s'exprime en fonction des écarts aux moyennes individuelles en tenant compte de l'autocorrélation des perturbations, soit :

$$L = -\frac{NT}{2} \ln(2\pi) - \frac{NT}{2} \ln \sigma_u^2 + T \ln |B_N| - \frac{1}{2\sigma_u^2} u^{**} u^* \quad (12)$$

avec

$$u^* = (I_T \otimes B_N)Q_{0,N}y - Q_{0,N}X\beta. \quad (13)$$

De façon similaire à la procédure décrite pour le modèle spatial autorégressif à erreurs composées, une procédure itérative à deux-étapes est utilisée pour maximiser la log-vraisemblance (équation 12).

L'hétérogénéité des données régionales

La base de données régionales a été élaborée à partir de différentes sources administratives (cf. encadré 2). Globalement, les chiffres du trafic des véhicules légers, du parc automobile, de la population, du revenu réel et du prix des carburants sur l'ensemble de la période 1985-2007 traduisent des écarts importants entre les minima et maxima (cf. tableau 1). Les écarts-types de ces variables montrent une dispersion significative de leurs observations. Ceci est le

reflet d'une hétérogénéité régionale prononcée, qui est valable pour toutes les variables de l'échantillon sauf pour le prix des carburants, homogène pour l'ensemble des régions (*i.e.* la variabilité n'est que temporelle, cf. encadré 2). Les kilométrages annuels moyens régionaux des véhicules légers sur la période 1985-2007 sont hétérogènes (cf. carte).

On constate une relative stabilité des valeurs des quartiles depuis les années 2000 pour le trafic des véhicules légers, le parc automobile, la

Encadré 2

LES DONNÉES

Les informations utilisées combinent plusieurs sources administratives de données régionales. Plus précisément, il s'agit de séries annuelles, remontant à 1985, couvrant 21 régions françaises de métropole. La Corse a été agrégée à la région Provence-Alpes-Côte d'Azur (PACA).

Le revenu disponible brut, la population et l'indice général des prix à la consommation

Le revenu disponible brut des ménages (*RDB*), la population (*Pop*) et l'indice général des prix (P_{IPC} , base 100 en 2007) sont fournis par l'Insee. Le revenu réel des ménages (*Rev*) est calculé en déflatant le revenu disponible brut des ménages par l'indice général des prix à la consommation.

Les véhicules légers sont ceux que l'on peut conduire avec un permis B (*i.e.* les voitures particulières et les véhicules utilitaires légers de moins de 3,5 tonnes). Leur parc, leur kilométrage et leur consommation de carburant sont décrits par les Bilans de la circulation élaborés par la Commission des comptes des transports de la Nation (CCTN). Sur cette base, il a été possible d'obtenir des séries homogènes remontant à 1990. Pour les années antérieures, les séries ont été rétropolées sur la base des travaux de Pirotte et Madre (1992, 1997).

Le trafic des véhicules légers

L'évolution des trafics (Tra_{it} dans la région i l'année t) se calcule à partir :

- des ventes de carburant ($Carb_{jt}$, $j=e$ pour les essences, et d pour le gazole consommé par les véhicules diesel), source : Comité professionnel des pétroles (CPDP) ;
- de l'évolution de la consommation unitaire des véhicules légers (litres au 100 km en distinguant diesel et essences, MS_{jt}), et de la part de chaque carburant consommée par les véhicules légers (LV_{jt}).

Au final, puisque les variables MS_{jt} et LV_{jt} ne sont disponibles qu'au niveau national, le trafic dans la région i l'année t est obtenu par :

$$Tra_{it} = Carb_{et} \times MS_{et} \times LV_{et} + Carb_{dt} \times MS_{dt} \times LV_{dt}$$

Le prix relatif des carburants

Pour mesurer l'évolution du prix relatif TTC des carburants, deux indicateurs sont retenus :

- le prix moyen au litre (P_L) calculé à partir des données de prix par catégories de carburants pondérés par les volumes consommés (source : CPDP) ;
- un coût moyen du carburant au kilomètre (P_C) estimé sur la base de l'efficacité énergétique des véhicules en litres pour 100 km (source : CCTN).

Les prix des carburants ont été administrés jusqu'en 1985. Ce n'est que depuis leur libéralisation que leur évolution est susceptible de différer d'une région à l'autre. La base de données de TNS WorldPanel pourrait fournir des informations au niveau régional, mais son échantillon est trop limité (3 000 véhicules France entière) pour permettre des estimations suffisamment précises sur les évolutions annuelles. En conséquence, nous retenons l'hypothèse classique en micro-économie selon laquelle le prix évolue de la même manière pour tous les consommateurs (*i.e.* dans chaque région).

Le parc des véhicules légers

La ventilation régionale du parc de véhicules légers (*Parc*) n'a pas posé de problème jusqu'en 2001 grâce aux statistiques de vignette, malgré un champ un peu plus large (jusqu'à 16 tonnes au lieu de 3,5 tonnes pour les véhicules légers). Après cette date, on a utilisé le Fichier central des automobiles qui dénombre les cartes grises en se limitant aux véhicules de moins de 15 ans (car seulement la moitié des cartes grises des voitures ne circulant plus en France sont renvoyées aux préfetures). Là encore, notamment pour assurer une transition cohérente entre les statistiques de vignette et de carte grise, le total national a été calé sur le parc automobile fourni par le Bilan de la circulation élaboré par la CCTN.

population et le revenu réel (cf. tableau 2). En revanche, les quartiles par région (cf. tableau 3) montrent une forte hétérogénéité. C'est pour les

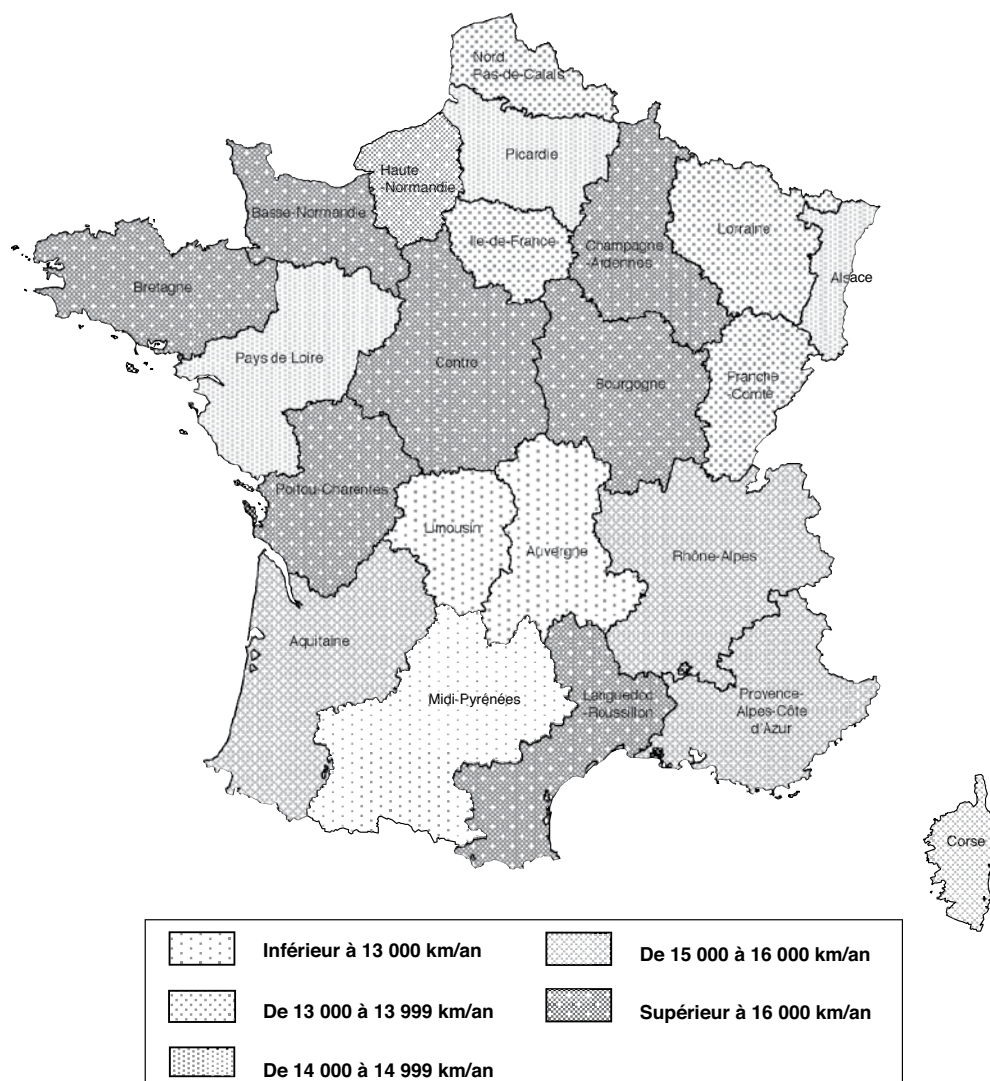
régions Île-de-France, Rhône-Alpes, Provence-Alpes-Côte d'Azur et Nord-Pas-de-Calais que les niveaux médians sont globalement les plus

Tableau 1
Statistiques descriptives des variables du trafic des véhicules légers, du parc automobile, de la population, du revenu réel et du prix des carburants (1985-2007)

	Moyenne	Écart-type	Minimum	Maximum	Total (France)
Trafic (<i>Tra</i>) en 100 millions de véh/km	210,6	139,5	41,5	705,7	4 422,5
Parc automobile (<i>Parc</i>) en milliers	1 418,0	974,1	356,0	5 112,1	29 779,2
Population (<i>Pop</i>) en milliers	2 768,3	2 173,1	711,5	1 1616,5	58 134,6
Revenu réel (<i>Rev</i>) en milliards	49,43	47,60	9,95	313,03	1 038,10
Prix/litre $\frac{P_L}{(P_{IPC})}$ en euros	0,73	0,19	0,54	1,16	0,73
Coût/km $\frac{P_C}{(P_{IPC})}$ en euros	0,06	0,01	0,05	0,08	0,06

Champ : 21 régions de France métropolitaine (Corse agrégée à la région PACA).
Source : CPDP, Ministère des Finances, Insee.

Carte : Kilométrages annuels moyens des véhicules légers par région sur la période 1985-2007



Source : CPDP-Ministère de l'Economie et des Finances-CCTN.

Tableau 2
Quartiles par année du trafic, du parc automobile, de la population et du revenu réel

	Trafic (en 100 millions de véh/km)				Parc automobile (en milliers)				Population (en milliers)				Revenu réel (en milliards d'euros)			
	Total	Q1	Médiane	Q3	Total	Q1	Médiane	Q3	Total	Q1	Médiane	Q3	Total	Q1	Médiane	Q3
1985	318,9	8,10	11,19	16,84	23 767	638	844	1 291	55 160	1 423	1 884	2 732	791,0	19,1	25,4	38,7
1986	336,6	8,42	12,07	17,60	24 306	649	859	1 323	55 415	1 426	1 895	2 745	811,2	19,9	26,2	39,8
1987	352,0	8,77	12,72	18,55	24 896	662	878	1 360	55 684	1 427	1 909	2 755	816,2	19,8	26,4	39,3
1988	369,8	9,12	13,45	19,22	25 541	678	898	1 397	55 960	1 432	1 924	2 767	844,2	20,4	26,9	40,5
1989	385,2	9,59	14,30	19,96	26 221	694	920	1 432	56 263	1 437	1 941	2 780	875,8	21,3	27,9	42,1
1990	398,6	10,02	14,96	20,98	26 925	711	942	1 465	56 577	1 442	1 962	2 795	904,1	22,1	28,9	43,4
1991	405,3	10,40	15,02	21,05	27 538	724	961	1 500	56 840	1 446	1 976	2 808	920,0	22,1	29,6	44,3
1992	415,7	11,03	14,41	22,52	28 075	736	978	1 532	57 110	1 449	1 990	2 820	942,8	22,6	30,3	45,3
1993	419,9	11,30	15,21	22,55	28 599	750	997	1 563	57 369	1 454	2 005	2 833	929,1	22,3	29,9	44,7
1994	428,5	11,51	15,34	23,06	29 160	766	1 021	1 599	57 565	1 457	2 016	2 842	960,7	23,1	30,9	46,2
1995	439,2	11,77	15,44	23,27	29 565	778	1 036	1 627	57 752	1 460	2 027	2 852	986,0	23,7	31,7	47,4
1996	445,1	11,74	15,66	23,82	29 851	786	1 047	1 650	57 936	1 462	2 040	2 861	989,5	23,7	31,8	47,6
1997	453,8	11,90	15,63	24,68	30 344	799	1 066	1 687	58 116	1 465	2 052	2 875	1 013,3	24,3	32,6	48,7
1998	468,0	12,51	16,35	26,64	31 023	818	1 091	1 734	58 299	1 467	2 065	2 888	1 042,5	25,0	33,6	50,1
1999	481,6	12,83	17,03	27,75	31 834	841	1 120	1 782	58 496	1 469	2 075	2 906	1 067,6	25,6	34,4	51,4
2000	483,5	12,63	16,98	27,88	32 204	849	1 144	1 797	58 858	1 473	2 088	2 931	1 073,0	25,0	35,1	51,4
2001	502,4	13,35	17,68	29,26	32 811	860	1 165	1 840	59 266	1 478	2 093	2 960	1 127,6	26,5	37,8	55,2
2002	508,9	13,54	18,47	30,48	33 392	871	1 184	1 876	59 685	1 483	2 098	2 990	1 179,7	27,7	39,5	57,4
2003	513,4	13,93	18,81	30,37	33 693	882	1 192	1 890	60 101	1 487	2 102	3 020	1 211,8	28,6	40,7	59,0
2004	514,3	14,71	18,38	28,08	33 699	878	1 190	1 887	60 505	1 491	2 105	3 050	1 266,1	29,7	42,5	61,8
2005	510,1	14,50	18,09	27,82	33 660	876	1 200	1 887	60 963	1 495	2 110	3 082	1 309,6	30,9	43,5	64,9
2006	508,8	14,64	17,93	27,83	33 623	871	1 218	1 886	61 399	1 499	2 115	3 113	1 372,2	32,7	45,3	69,2
2007	512,1	14,84	18,42	27,98	34 191	881	1 247	1 923	61 771	1 502	2 117	3 139	1 442,5	34,0	46,8	72,5

Champ : 21 régions de France métropolitaine (Corse agrégée à la région PACA).

Source : CPDP, Ministère des Finances, Insee.

Tableau 3

Quartiles par région du trafic, du parc automobile, de la population et du revenu réel

	Trafic (en 100 millions de véh/km)				Parc automobile (en milliers)				Population (en milliers)				Revenu réel (en milliards d'euros)			
	Total	Q1	Médiane	Q3	Total	Q1	Médiane	Q3	Total	Q1	Médiane	Q3	Total	Q1	Médiane	Q3
	Île-de-France	14 396,8	600,6	631,5	652,8	107 027,5	4 366,5	4 712,5	4 912,9	250 719,8	10 624,5	10 871,4	11 123,5	5 389,7	208,6	227,3
Champagne-Ardenne	2 493,1	89,9	110,7	123,9	15 375,6	637,8	661,0	696,0	30 944,2	1 341,4	1 347,0	1 348,3	524,6	20,8	22,9	24,1
Picardie	3 029,9	114,3	129,1	149,0	20 892,9	787,3	884,0	995,9	42 360,0	1 806,5	1 849,1	1 870,1	667,7	24,8	27,6	32,1
Haute-Normandie	3 224,3	124,8	140,6	156,2	20 396,7	782,3	869,0	951,3	40 567,1	1 733,5	1 772,7	1 790,6	677,1	26,0	28,7	31,5
Centre	5 025,4	201,4	215,4	243,9	31 210,9	1 220,5	1 352,5	1 506,7	55 681,1	2 366,1	2 421,3	2 464,6	981,6	36,6	40,5	47,5
Basse-Normandie	2 669,9	99,6	114,6	130,7	16 328,5	679,0	716,0	730,5	32 514,8	1 389,7	1 409,9	1 433,8	545,2	20,4	22,8	25,6
Bourgogne	3 591,3	137,0	157,3	176,3	20 466,2	815,8	889,0	974,5	37 126,7	1 608,8	1 613,8	1 616,3	644,5	24,3	26,8	30,7
Nord-Pas-de-Calais	5 476,2	195,9	233,6	277,6	39 002,1	1 505,3	1 681,5	1 900,9	91 684,7	3 962,7	3 988,9	4 002,3	1 407,9	54,7	59,7	65,3
Lorraine	3 392,7	123,0	149,9	167,9	25 849,5	997,0	1 120,0	1 256,6	53 316,3	2 311,4	2 315,6	2 323,5	897,9	34,2	37,7	42,6
Alsace	2 891,7	99,9	127,2	144,0	20 366,3	775,0	885,5	1 007,9	39 075,8	1 620,9	1 693,4	1 759,1	710,5	26,2	29,2	34,9
Franche-Comté	1 874,3	66,7	81,3	94,3	13 499,7	517,8	587,5	657,9	25 662,5	1 096,3	1 112,7	1 127,9	427,3	15,7	17,4	21,2
Pays de la Loire	5 863,0	205,9	237,3	311,5	40 260,1	1 519,0	1 719,5	1 994,5	73 357,8	3 051,6	3 154,9	3 289,9	1 226,0	44,3	49,5	60,4
Bretagne	5 725,4	211,6	240,5	295,4	34 812,8	1 388,3	1 550,3	1 619,3	66 415,2	2 791,2	2 847,2	2 961,2	1 129,5	41,8	46,4	54,3
Poitou-Charentes	3 652,2	149,1	157,9	171,1	22 392,0	876,0	963,0	1 083,7	37 627,4	1 594,2	1 621,5	1 665,2	626,5	23,1	25,3	30,4
Aquitaine	5 896,5	231,3	253,3	279,3	38 825,3	1 480,0	1 661,5	1 901,3	66 533,3	2 790,5	2 860,1	2 970,5	1 171,6	43,5	47,8	56,2
Midi-Pyrénées	4 266,5	159,3	177,7	211,0	34 673,6	1 319,0	1 482,0	1 696,0	58 393,4	2 426,8	2 500,9	2 617,4	1 006,1	37,2	40,9	48,5
Limousin	1 267,4	46,5	53,1	62,1	9 770,0	390,0	424,0	463,1	16 607,3	715,7	721,1	726,4	290,0	10,9	12,2	13,6
Rhône-Alpes	1 0947,0	443,7	481,5	529,6	6 9352,9	2 676,5	2 978,5	3 403,4	128 175,5	5 334,8	5 538,4	5 755,4	2 276,3	82,9	90,8	113,9
Auvergne	2 273,2	82,8	94,9	116,8	17 358,8	684,5	749,0	832,3	30 381,4	1 312,2	1 319,8	1 326,3	521,3	19,5	21,6	24,8
Languedoc-Roussillon	4 657,4	188,3	194,5	230,0	28 862,6	1 087,5	1 218,5	1 420,2	51 765,9	2 107,4	2 219,2	2 362,4	852,7	31,1	34,3	41,5
PACA+Corse	9 103,0	360,3	385,9	449,5	58 198,6	2 243,3	2 495,5	2 854,3	108 186,3	4 495,2	4 674,5	4 866,1	1 902,5	69,8	77,5	93,0

Champ : 21 régions de France métropolitaine (Corse agrégée à la région PACA).

Source : CDP, Ministère des Finances, Insee.

élevés pour le trafic des véhicules légers, le parc automobile, la population et le revenu réel. On remarque également que les inégalités de revenus y sont fortes (*i.e.* l'amplitude des quartiles est importante). En revanche, les régions Auvergne et Limousin sont marquées par de faibles niveaux médians et des inégalités peu importantes. Ce constat se vérifie également pour la Franche Comté et Champagne-Ardenne, et dans une moindre mesure pour la Haute-Normandie, la Basse-Normandie, l'Alsace et la Bourgogne.

L'hétérogénéité entre régions est prédominante (cf. tableau 4). En effet, la variabilité inter-régionale représente plus de 90 % de la variabilité totale pour le trafic des véhicules légers, le parc automobile, la population et le revenu réel. Néanmoins, lorsque les variables sont exprimées par tête et prises en logarithme, la répartition entre les deux variabilités est plus homogène, résultat d'un effet taille atténué.

Enfin, le nombre moyen de liens s'élève à 14 pour la matrice de distance et à 4 pour la matrice de contiguïté à l'ordre un (cf. tableau 5).

Autrement dit, chaque région possède en moyenne 14 voisins (qui peuvent être des voisins de voisins donc plus ou moins éloignés dans l'espace) si l'on considère la matrice de distance et 4 voisins par rapport à la matrice de contiguïté à l'ordre un. La valeur moyenne du poids de chaque matrice spatiale¹² est respectivement de 0,068 et 0,244. Enfin, 80,5 % des poids de la matrice de contiguïté à l'ordre un sont nuls contre 30,6 % pour ceux associés à la matrice de distance (cf. graphiques 1 et 2).

Les estimations des modèles spatiaux

Avant de présenter les résultats d'estimation, les tests du multiplicateur de Lagrange joint et conditionnels développés par Baltagi et Liu (2008) ont été appliqués. L'ensemble de ces tests nous amènent à rejeter dans tous les cas l'hypothèse nulle (*i.e.* absence de corrélation spatiale et/ou d'hétérogénéité régionale). Par

12. Le raisonnement est basé sur des matrices spatiales normalisées.

Tableau 4
Variabilités inter-régionale, intra-régionale et totale des variables liées au trafic des véhicules légers (1985-2007)

	Inter-régionale	Intra-régionale	Total
Trafic (Tra)	94,35	5,65	100
Parc automobile ($Parc$)	96,21	3,79	100
Population (Pop)	99,62	0,38	100
Revenu réel (Rev)	94,22	5,78	100
Prix/litre $\frac{P_L}{(P_{IPC})}$	0	100	100
Coût/km $\frac{P_C}{(P_{IPC})}$	0	100	100
$\ln\left(\frac{Tra}{Parc}\right)$	64,01	35,99	100
$\ln\left(\frac{Rev}{Parc}\right)$	16,37	83,53	100
$\ln\left(\frac{Parc}{Pop}\right)$	49,21	50,79	100
$\ln\left(\frac{P_L}{P_{IPC}}\right)$	0	100	100
$\ln\left(\frac{P_C}{P_{IPC}}\right)$	0	100	100

Lecture : la variabilité inter-régionale traduit les différences permanentes entre les régions ; la seconde composante retrace les écarts entre la situation de la région, à chaque date, et sa position moyenne sur la période. Par exemple, si l'on s'intéresse au trafic des véhicules légers (Tra) de la région i à la date t . Dans ce cas, la composante inter-régionale de la variabilité représente la partie des différences de trafics entre les régions qui est permanente. La composante intra-régionale rend compte, quant à elle, des écarts, à chaque date, entre le trafic de la région et son trafic moyen sur la période.

Champ : 21 régions de France métropolitaine (Corse agrégée à la région PACA).

Source : CPDP, Ministère des Finances, Insee.

exemple, en considérant la matrice spatiale de distance $W_{1,N}$ et le test de l'hypothèse jointe $H_0 : \rho = \sigma_u^2 = 0$ la statistique calculée du multiplicateur de Lagrange s'élève à 2 501,06. Cette valeur est donc strictement supérieure à celle du χ^2 à deux degrés de liberté au seuil de 5 % (*i.e.* 5,99)¹³.

Globalement, les signes des coefficients des différents modèles estimés sont cohérents avec la littérature sur le sujet (cf. tableau 6). Plus précisément, le revenu par tête a un impact positif sur le trafic par véhicule, tandis que les signes des coefficients associés aux variables prix des carburants¹⁴ et taux d'équipement des ménages sont négatifs. Ce dernier résultat signifie que l'augmentation du taux d'équipement des ménages réduit le kilométrage moyen des véhicules légers (Robin, 2010). Pour les modèles à effets fixes et à erreurs composées (*i.e.* colonne 2 à 4), les élasticités ponctuelles du revenu varient de 0,51 à 0,56. Pour le taux d'équipement (resp. pour le prix des carburants), les valeurs se situent entre - 0,25 et - 0,21 (resp. entre - 0,14 et - 0,12). Ces ordres de grandeur restent comparables à ceux figurant dans la littérature (cf. Goodwin *et al.*, 2004). Ils se rapprochent en valeur absolue plutôt des valeurs basses. Néanmoins, ces spécifications négligent en tant que telles les interdépendances spatiales. Or, Pirotte et Madre (2011) ont montré qu'elles existaient, et qu'ils convenaient de les modéliser explicitement. Le fait d'ignorer la dépendance spatiale qui affecte le trafic des véhicules légers (*i.e.* l'aspect spatial autorégressif du modèle) a pour conséquence d'engendrer des biais qui peuvent être importants (cf. LeSage et Pace, 2009). De plus, Baltagi et Pirotte (2010) ont montré que l'omission de la dimension spatiale sur les estimateurs habituels utilisés sur données de panel peut conduire à des résultats inexacts en termes de biais et de précision.

Les coefficients autorégressifs spatiaux $\hat{\rho}$ apparaissent statistiquement significatifs (cf. les quatre

dernières colonnes du tableau 6). Par exemple, si l'on prend celui du modèle spatial autorégressif à erreurs composées (matrice de distance, colonne 5), il s'élève à 0,44. Cela signifie que si le trafic des véhicules légers augmente de 1 % dans les régions limitrophes, la répercussion sur le trafic de la région concernée sera de 0,44 %. L'intervalle de confiance à 95 % conduit pour l'ensemble des régions à dire que le coefficient autorégressif spatial se situe entre 0,28 et 0,61.

Par ailleurs, on constate que les résultats des modèles spatiaux autorégressifs sont relativement proches, et plus éloignés de ceux des modèles à erreurs composées et à effets fixes individuels. Afin de pouvoir discriminer entre eux, on utilise les valeurs de la log-vraisemblance (cf. dernière ligne du tableau 6). Si l'on retient la valeur de la log-vraisemblance du modèle à erreurs composées (617,9) et celle du modèle spatial autorégressif à erreurs composées - matrice de distance - (629,2), cette dernière spécification domine le modèle à effets aléatoires souvent utilisé sur données de panel. Plus précisément, on peut en conclure que cette différence est significative en la comparant à un khi-deux à un degré de liberté (cf. Greene, 2008). Un raisonnement similaire est possible pour discriminer les modèles spatiaux à erreurs composées entre eux en se référant à LeSage et Pace (2009, p. 164). En outre, les statistiques calculées du test d'Hausman (1978) nous amènent à chaque fois à rejeter l'hypothèse de corrélation des effets spécifiques individuels avec les variables explicatives (cf. avant-dernière ligne du tableau 6). Au final, on privilégie les résultats associés aux modèles à erreurs composées.

13. Les résultats sont strictement de même nature dans le cas de la matrice spatiale de contiguïté $W_{2,N}$.

14. Les résultats du coût moyen du carburant au kilomètre (PC) ne sont pas significativement différents de ceux du prix moyen au litre (PL). La présentation et les commentaires des résultats se font donc uniquement par rapport au prix moyen au litre.

Tableau 5
Caractéristiques de connectivité des matrices spatiales

	Matrice de distance	Matrice de contiguïté
Dimension	21	21
Part d'éléments non-nuls (%)	69,4	19,5
Nombre de liens non-nuls	306	86
Nombre maximum de liens	19	8
Nombre moyen de liens	14,5	4,09
Poids moyen	0,068	0,244

Lecture : chaque région possède en moyenne 14,5 voisins (qui peuvent être des voisins de voisins donc plus ou moins éloignés dans l'espace) si l'on considère la matrice de distance et 4,09 voisins par rapport à la matrice de contiguïté à l'ordre un.

Champ : 21 régions de France métropolitaine (Corse agrégée à la région PACA).

Source : calcul des auteurs.

Tableau 6
Résultats des estimations des modèles non spatiaux et spatiaux autorégressifs (1985-2007)
Variable dépendante : Ln(Tra/Parc)

	Modèle à effets fixes		Modèle à erreurs composées		Matrice de distance		Matrice de contiguïté	
	Within		MCQG		Modèle spatial autorégressif à erreurs composées		Modèle spatial autorégressif à effets fixes	
	β_j	β_j	β_j	β_j	MV	MV	MV	MV
Ln(Rev/Pop)	0,6037 (8,7634) [0,4717;0,7387]	0,5625 (8,4900) [0,4326;0,6924]	0,5607 (8,3189) [0,4286;0,6928]	0,4331 (6,2440) [0,2972;0,5691]	0,4562 (3,8997) [0,2269;0,6855]	0,4577 (6,6157) [0,3221;0,5933]	0,4595 (3,9837) [0,2334;0,6856]	
Ln(Parc/Pop)	-0,2552 (-3,7477) [-0,3886;-0,1218]	-0,2147 (-3,3026) [-0,3421;-0,0873]	-0,2130 (-3,2223) [-0,3426;-0,0834]	-0,2934 (-4,4633) [-0,4223;-0,1646]	-0,2109 (-1,8803) [-0,4308;0,0089]	-0,2302 (-3,5689) [-0,3566;-0,1038]	-0,2047 (-1,8523) [-0,4213;0,0119]	
Ln(P _i /P _{ind})	-0,1389 (-4,6523) [-0,1976;0,0802]	-0,1270 (-4,3370) [-0,1844;-0,0696]	-0,1265 (-4,2881) [-0,1844;-0,0686]	-0,0896 (-3,0600) [-0,1470;-0,0322]	-0,0921 (-1,8591) [-0,1891;0,0049]	-0,0977 (-3,3374) [-0,1551;-0,0403]	-0,1011 (-1,8439) [-0,2086;0,0064]	
ρ	-	-	-	0,4467 (5,3466) [0,2830;0,6104]	0,1946 (1,6464) [-0,0370;0,4264]	0,2670 (4,6228) [0,1538;0,3802]	0,1555 (2,1289) [0,0123;0,2987]	
θ^2	-	0,0122	0,0128	0,0113	-	0,0123	-	
Test d'Hausman	-	4,7835	617,9	629,18	9,4440	627,8	6,485	
Log-Vraisemblance	-	-	-	-	583,7	-	583,4	

Lecture : ce tableau présente les résultats des différents modèles estimés. Tous les modèles étant log-log, les coefficients s'interprètent comme des élasticités. Les colonnes deux, trois et quatre présentent les estimations associées aux modèles à effets fixes individuels et à erreurs composées. Les estimateurs associés correspondent respectivement à l'estimateur Within et à celui des Moindres Carrés Quasi-Généralisés (MCQG) proposé par Amemiya (1971) (cf. Piroette, 2011). L'estimateur du maximum de vraisemblance (MV) a également été utilisé pour estimer le modèle à erreurs composées (cf. colonne 4). Les quatre dernières colonnes concernent l'estimation des modèles spatiaux autorégressifs à effets fixes et à erreurs composées en différenciant les résultats selon la matrice spatiale retenue.

Les nombres entre parenthèses sont les t de Student calculés. Les nombres entre crochets correspondent aux intervalles de confiance à 95 %. Le coefficient de la constante et ceux de certaines variables de contrôle n'ont pas été reportés, notamment ceux affectés à des variables indicatrices spécifiques à des évolutions particulières. Par exemple, une variable indicatrice a été introduite pour tenir compte de la forte évolution du prix des carburants sur la période 2004-2007.

Champ : 21 régions de France métropolitaine (Corse agrégée à la région PACA).

Source : calcul des auteurs.

Tableau 7
Effets directs, indirects et totaux estimés associés au modèle spatial autorégressif à erreurs composées (matrice de distance)

A - Effets cumulés											
	Ln(Rev/Pop)			Ln(Parc/Pop)			Ln(P _i /P _{PC})			t de Student	
	Moyenne	Ecart-type	t de Student	Moyenne	Ecart-type	t de Student	Moyenne	Ecart-type	t de Student		
Effet direct	0,4440	0,0310	14,330	- 0,3008	0,0348	- 8,6321	- 0,0919	0,0116	- 7,9062		
Effet indirect	0,3388	0,0868	3,9007	- 0,2295	0,0960	- 2,3909	- 0,0701	0,0398	- 1,7588		
Effet total	0,78286	0,1178	6,6433	- 0,5304	0,1308	- 4,0532	- 0,1620	0,0515	- 3,1467		
B - Répartition des effets spatiaux											
W _N à l'ordre j	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Direct	Indirect
W _N ⁰	0,43315	0,43315	0,00000	- 0,29349	- 0,29349	0,00000	- 0,08966	- 0,0896	0,00000	- 0,0896	0,00000
W _N ¹	0,19349	0,00000	0,19349	- 0,13110	0,00000	- 0,13110	- 0,04005	0,00000	- 0,04005	0,00000	- 0,04005
W _N ²	0,08643	0,00729	0,07915	- 0,05856	- 0,00494	- 0,05363	- 0,01789	- 0,00151	- 0,01638	- 0,00151	- 0,01638
W _N ³	0,03861	0,00205	0,03656	- 0,02616	- 0,00139	- 0,02477	- 0,00799	- 0,00042	- 0,00757	- 0,00042	- 0,00757
W _N ⁴	0,01725	0,00090	0,01635	- 0,01169	- 0,00061	- 0,01108	- 0,00357	- 0,00019	- 0,00338	- 0,00019	- 0,00338
W _N ⁵	0,00770	0,00038	0,00732	- 0,00522	- 0,00026	- 0,00496	- 0,00159	- 0,00008	- 0,00152	- 0,00008	- 0,00152
W _N ⁶	0,00344	0,00017	0,00327	- 0,00233	- 0,00011	- 0,00222	- 0,00071	- 0,00003	- 0,00068	- 0,00003	- 0,00068
W _N ⁷	0,00154	0,00007	0,00146	- 0,00104	- 0,00005	- 0,00099	- 0,00032	- 0,00002	- 0,00030	- 0,00002	- 0,00030
W _N ⁸	0,00069	0,00003	0,00065	- 0,00047	- 0,00002	- 0,00044	- 0,00014	- 0,00001	- 0,00014	- 0,00001	- 0,00014
W _N ⁹	0,00031	0,00001	0,00029	- 0,00021	- 0,00001	- 0,00020	- 0,00006	- 0,00000	- 0,00006	- 0,00000	- 0,00006
W _N ¹⁰	0,00014	0,00001	0,00013	- 0,00009	- 0,00000	- 0,00009	- 0,00003	- 0,00000	- 0,00003	- 0,00000	- 0,00003
W _N ¹¹	0,00006	0,00000	0,00006	- 0,00004	- 0,00000	- 0,00004	- 0,00001	- 0,00000	- 0,00001	- 0,00000	- 0,00001
W _N ¹²	0,00003	0,00000	0,00001	- 0,00002	- 0,00000	- 0,00002	- 0,00001	- 0,00000	- 0,00001	- 0,00000	- 0,00001
$\sum_{q=0}^{12} W_N^q$	0,78286	0,44407	0,33879	- 0,53043	- 0,30088	- 0,22955	- 0,16205	- 0,09192	- 0,07013		

Champ : 21 régions de France métropolitaine (Corse agrégée à la région PACA).

Source : calcul des auteurs.

Au sein de ces modèles, les résultats du modèle spatial autorégressif à erreurs composées basés sur la matrice de distance (cf. colonne 5 du tableau 6) sont les plus satisfaisants.

L'avantage du modèle autorégressif spatial à erreurs composées basées sur la matrice de distance¹⁵ est qu'il exploite l'information des régions plus ou moins limitrophes (cf. tableau 7). Les élasticités directes prix, revenu et parc automobile sont respectivement égales à - 0,09, 0,44 et - 0,30 et statistiquement significatives. Si l'on se concentre sur l'effet prix, la différence entre le coefficient estimé - 0,0896 (cf. tableau 6, colonne 4) et l'élasticité prix directe estimée - 0,0919 (cf. tableau 7, colonne 8) est égale à - 0,0023. Cet écart s'interprète comme un effet de *feedback* des habitants de la région qui se déplacent dans des régions limitrophes mais qui reviennent dans la région d'origine. Cet effet est également assez faible pour le taux d'équipement et le revenu réel respectivement de - 0,0074 et 0,0109. Globalement, les bornes des intervalles de confiance à 95 % des élasticités directes du prix des carburants, du taux d'équipement et du revenu réel sont respectivement [- 0,11 ; - 0,07], [- 0,36 ; - 0,23] et [0,38 ; 0,50]. En comparant ces résultats à ceux de Piroette et Madre (2011) (*i.e.* la dimension temporelle se terminait à la fin des années 1990), l'élasticité au taux d'équipement semble diminuée significativement, l'élasticité au prix des carburants reste faible, alors que l'élasticité au revenu ne varie que faiblement. Pour les effets indirects, on peut en déduire qu'une augmentation de 1 % du prix des carburants se traduirait par une diminution du trafic des véhicules légers de - 0,04 % pour les régions voisines, de

15. Les résultats relatifs à la matrice de contiguïté peuvent être obtenus sur demande auprès des auteurs.

- 0,0163 % pour les régions voisines des régions voisines, etc. Des interprétations similaires peuvent être faites pour le taux d'équipement et le revenu réel.

* *
*

La principale difficulté pour expliquer le trafic des véhicules légers à l'intérieur d'un périmètre bien délimité (ici la région), avec des variables explicatives caractérisant les comportements de ses habitants, provient essentiellement des trafics de transit et d'échanges. Autrement dit, des trafics qui sont engendrés par des habitants extérieurs à la région. L'approche spatiale est particulièrement adaptée pour capter ces interdépendances régionales et en évaluer l'impact sur le trafic interne à une région. Ici, la modélisation spatiale des trafics des véhicules légers a révélé des interdépendances significatives entre les régions. L'influence négative du taux d'équipement sur le trafic des véhicules légers montre que les ménages continuent d'augmenter leur équipement de 1 à 2 % par an. Néanmoins, chaque véhicule est de moins en moins utilisé en moyenne. Ce phénomène a été observé en Allemagne depuis les années 1990, et dans les agglomérations françaises les plus importantes. Ceci constitue les premiers signes d'une saturation dont les effets peuvent être plus ou moins amplifiés par une faible croissance économique et une forte volatilité du prix des carburants (Collet, 2012). Ces modèles spatiaux ont également permis d'obtenir les effets directs, indirects et totaux. L'obtention des effets directs rend possible la comparaison des élasticités avec celles obtenues par des modèles plus classiques, c'est-à-dire qui ne comprennent pas de partie spatiale autorégressive. □

BIBLIOGRAPHIE

Abreu M., de Groot H.L.F. et Florax R.J.G.M. (2005), « Space and Growth: a survey of empirical evidence and methods », *Région et Développement*, n°21, pp. 13-43.

Amemiya T. (1971), « The estimation of variances in a variance-components model », *International Economic Review*, n°12, pp. 1-13.

Anselin L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.

Anselin L. et Le Gallo J. (2006), « Interpolation of air quality measures in hedonic house price models: spatial aspects », *Spatial Economic Analysis*, n°1, pp. 31-52.

- Anselin, L., Le Gallo J. et Jayet H. (2008)**, « Spatial panel econometrics », Chapter 19, in L. Mátyás and P. Sevestre, eds., *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice*, Springer-Verlag, pp. 625-660.
- Baltagi, B.H. (2008)**, *Econometric Analysis of Panel Data*, Wiley and Sons, Chichester.
- Baltagi, B.H. et Griffin J.M. (1983)**, « Gasoline demand in the OECD: an application of pooling and testing procedures », *European Economic Review*, n°22, pp. 117-137.
- Baltagi, B.H. et Griffin J.M. (1997)**, « Pooled estimators v.s. their heterogeneous counterparts in the context of dynamic demand for gasoline », *Journal of Econometrics*, n°77, pp. 303-327.
- Baltagi, B.H. et Liu L. (2008)**, « Testing for random effects and spatial lag dependence in panel data models », *Statistics and Probability Letters*, n°78, pp. 3304-3306.
- Baltagi, B.H. et Pirotte A. (2010)**, « Panel data inference under spatial dependence », *Economic Modelling*, n°27, pp. 1368-1381.
- Brons, M., Nijkamp P., Pels E. et Rietveld P. (2008)**, « A meta-analysis of the price elasticity of gasoline demand. A SUR approach », *Energy Economics*, n°30, pp. 2105-2122.
- Conley, T.G. (1999)**, « GMM estimation with cross-sectional dependence », *Journal of Econometrics*, n°92, pp. 1-45.
- Boucq, E., Collet R., Madre J.-L. et Hivert L. (2011)**, « Diffusion de l'automobile en France : vers quels plafonds pour la motorisation et l'usage », *Économie et Statistique, ce numéro*.
- Collet, R. (2012)**, « Household car use in France: a demographic and economic analysis », *Economics Bulletin*, n°32, pp. 475-485.
- Conley, T.G. et Topa G. (2002)**, « Socio-economic distance and spatial patterns in unemployment », *Journal of Applied Econometrics*, n°17, pp. 303-327.
- CCTN (2010)**, *Les Transports en 2009*, 47ème rapport de la Commission des Comptes Transport de la Nation.
- Dahl, C. (1986)**, « Gasoline demand survey », *Energy Journal*, n°7, pp. 67-82.
- Dahl, C. et Sterner T. (1991a)**, « A survey of econometric gasoline demand elasticities », *International Journal of Energy Systems*, n°11, pp. 53-76.
- Dahl, C. et Sterner T. (1991b)**, « Analyzing gasoline demand elasticities: A survey », *Energy Economics*, n°11, pp. 203-210.
- Dahl, C. (1995)**, « Demand for transportation fuels: a survey of demand elasticities and their components », *Journal of Energy Literature*, n°1, pp. 3-27.
- Dargay, J., Madre J.-L. et Berri A. (2000)**, « Car ownership dynamics seen through the follow-up of cohorts: a comparison of France and the United Kingdom », *Transportation Research Record*, n°1733, pp. 31-38.
- De Jong, G. et Gunn H. (2001)**, « Recent evidence on car cost and time elasticities of travel demand in Europe », *Journal of Transport Economics and Policy*, n°35, pp. 137-160.
- Elhorst, J.P. (2003)**, « Specification and estimation of spatial panel data models », *International Regional Science Review*, n°26, pp. 244-268.
- Elhorst, J.P. (2010)**, « Applied spatial econometrics: Raising the bar », *Spatial Economic Analysis*, n°5, pp. 9-28.
- Epsey, M. (1997)**, « Explaining the variation in elasticity estimates of gasoline demand in the United States: a meta-analysis », *Energy Economics*, n°17, pp. 49-60.
- Epsey, M. (1998)**, « Gasoline demand revisited: an international meta-analysis of elasticities », *Energy Economics*, n°20, pp. 273-295.
- Goodwin, P. (1992)**, « A review of new demand elasticities with special reference to short and long run effects of price changes », *Journal of Transport Economics and Policy*, n°26, pp. 155-169.
- Goodwin, P., Dargay J. et Hanly P. (2004)**, « Elasticities of road traffic and fuel consumption with respect to price and income: a review », *Transport Reviews*, n°24, pp. 275-292.
- Graham D.J. et Glaister S. (2002)**, « The demand for automobile fuel: a survey of elasticities », *Journal of Transport Economics and Policy*, n°36, pp. 1-26.

- Graham D.J. et Glaister S. (2004)**, « Road traffic demand elasticity estimates: a review », *Transport Reviews*, n°24, pp. 261-274.
- Greene, W.H. (2008)**, *Econometric Analysis*, Pearson International Edition, Upper Saddle River, New Jersey.
- Griffin, J.M. (1979)**, *Energy conservation in the OECD: 1980 to 2000*, Ballinger, Cambridge, Massachusetts.
- Haining, R. (1990)**, *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge University, Cambridge.
- Hausman, J. (1978)**, « Specification tests in econometrics », *Econometrica*, n°46, pp. 1251-1271.
- Hensher D.A. et Stanley J.K. (2009)**, « What if Petrol increased to \$10 per liter in Melbourne? Implications on travel behaviour and public transport demand », *Roads and Transport Research*, n°18, pp. 86-93.
- Hsiao, C. (2003)**, *Analysis of Panel Data*, Cambridge University Press, Cambridge.
- Hsiao, C., Appelbe T.W. and Dineen C.R. (1993)**, « A general framework for panel data analysis - with an application to canadian customer dialed long distance service », *Journal of Econometrics*, n°59, pp. 63-86.
- Johansson, O. et Schipper L. (1997)**, « Measuring the long-run fuel demand of cars, separate estimations of vehicle stock, mean fuel intensity, and mean annual driving distance », *Journal of Transportation Economics and Policy*, n°31, pp. 277-292.
- Kelejian, H.H. et Robinson D.P. (1995)**, « Spatial correlation: A suggested alternative to the autoregressive model », in *New Directions in Spatial Econometrics*, edited by L. Anselin and R. Florax, Springer Verlag, Berlin, pp. 75-95.
- Kelejian, H.H., Tavlak G.S. et Hondronyiannis G. (2006)**, « A spatial modelling approach to contagion among emerging economies », *Open Economies Review*, n°17, pp. 423-442.
- Kemel, E., Collet R. et Hivert L. (2011)**, « Evidence for an endogenous rebound effect impacting long-run car use elasticity to fuel price », *Economics Bulletin*, n°31, pp. 2777-2786.
- Le Gallo, J. (2002)**, « Économétrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire », *Économie et Prévision*, n°155, pp. 139-171.
- Le Gallo, J. (2004)**, « Hétérogénéité spatiale : principes et méthodes », *Économie et Prévision*, n°162, pp. 151-172.
- LeSage, J.P. et Pace R.K. (2009)**, *Introduction to Spatial Econometrics*, Chapman & Hall/CRC Press, Boca Raton.
- Li Z., Rose J.M. et Hensher D.A. (2010)**, « Forecasting automobile petrol demand in Australia: a critical evaluation of empirical models », *Transportation Research Part A*, n°44, pp. 16-38.
- Madre, J.-L. (1994)**, « Disparités géographiques dans la circulation automobile en France », *Recherche Transports Sécurité*, n°43, pp. 23-34.
- Madre, J.-L. et Pirotte A. (1992)**, *Régionalisation des projections à long terme de la circulation automobile*, Rapport INRETS pour le SETRA.
- Madre, J.-L. et Pirotte A. (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.
- Moulton, B.R. (1986)**, « Random group effects and the precision of regression estimates », *Journal of Econometrics*, n°32, pp. 385-397.
- Moulton, B.R. (1987)**, « Diagnostics for group effects in regression analysis », *Journal of Business and Economic Statistics*, n°5, pp. 275-282.
- Pesaran, M.H. et Smith R. (1995)**, « Estimating long-run relationships from dynamic heterogeneous panels », *Journal of Econometrics*, n°68, pp. 79-113.
- Pesaran, M.H., Smith R. et Im K.S. (1996)**, « Dynamic linear models for heterogeneous panels », Chapter 8 in L. Mátyás and P. Sevestre, eds., *The Econometrics of Panel Data: A Handbook of the Theory With Applications*, Kluwer Academic Publishers, Dordrecht, pp. 145-195.
- Pesaran, M.H., Schuermann T. et Weiner S. (2004)**, « Modeling regional interdependencies using a global error-correcting macroeconomic model », *Journal of Business and Economic Statistics*, n°22, pp. 129-162.

Pirotte, A. (2011), *Économétrie des Données de Panel : Théories et Applications*, Collection Corpus Économie dirigée par H. Kempf, Economica.

Pirotte A. et Madre J.-L. (2012), « Car traffic elasticities: a spatial panel data analysis of french regions », *Journal of Transport Economics and Policy*, n° 45, pp. 341-365.

Robertson, D. et Symons J. (1992), « Some strange properties of panel data estimators », *Journal of Applied Econometrics*, n°7, pp. 175-189.

Robin, M. (2010), « La motorisation des ménages continue de s'accroître au prix d'un vieillissement

du parc automobile », Service de l'Observatoire et des Statistiques, CGDD, pp. 99-121.

Sétra (2009), *Sondage de circulation 2004-2005*, Rapport d'Étude.

Sweeney, J.L. (1979), « The demand for gasoline in the United States: a vintage capital model », *Review of Economics and Statistics*, n°61, pp. 169-179.

Trace (1998), *Review of existing evidence on time and cost elasticities of travel demand and on the value of travel time*, The Hague: Trace consortium.

