

Cet article est disponible en ligne à l'adresse :

[http://www.cairn.info/article.php?ID\\_REVUE=ECOP&ID\\_NUMPUBLIE=ECOP\\_155&ID\\_ARTICLE=ECOP\\_155\\_0139](http://www.cairn.info/article.php?ID_REVUE=ECOP&ID_NUMPUBLIE=ECOP_155&ID_ARTICLE=ECOP_155_0139)

---

## Économétrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire

par Julie LE GALLO

| La Documentation française | Économie et Prévision

2002/4 - n° 155

ISSN 0249 - 4744 | pages 139 à 157

---

Pour citer cet article :

— Le Gallo J., Économétrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire, Économie et Prévision 2002/4, n° 155, p. 139-157.

---

Distribution électronique Cairn pour La Documentation française.

© La Documentation française. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

# Économétrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire

Julie Le Gallo <sup>(\*)</sup>

*L'objectif de cet article est de présenter les outils nécessaires à la prise en compte de l'autocorrélation spatiale, définie par l'absence d'indépendance entre observations géographiques, dans le cadre des modèles de régression linéaire. Alors qu'il est souvent admis que les données spatiales observées en coupe transversale sont indépendantes, cette hypothèse est rarement justifiée et devrait être au contraire systématiquement testée. Par conséquent, après avoir exposé les diverses façons de modéliser l'autocorrélation spatiale, nous présentons les procédures d'estimation et d'inférence adaptées aux modèles économétriques intégrant explicitement cet effet. Enfin, nous proposons une démarche générale visant à tester et à prendre en compte l'autocorrélation spatiale dans les travaux empiriques.*

(\*) Latic UMR-CNRS 5118 Université de Bourgogne.  
E.mails : Julie.Le-Gallo@u-bourgogne.fr ou jlegallo@uiuc.edu

L'auteur remercie C. Baumont, L. Bertinelli, C. Ertur, J.-M. Huriot, M.-C. Pichery ainsi qu'un rapporteur anonyme pour leurs commentaires et suggestions. L'auteur reste seule responsable des insuffisances que pourrait comporter ce texte.

Dans les travaux empiriques, l'économiste est souvent confronté à l'utilisation de données localisées, c'est-à-dire au traitement des observations d'une variable mesurée en des localisations différentes réparties dans l'espace. Il est souvent admis que ces données spatiales observées en coupe transversale sont indépendantes alors que cette hypothèse est rarement justifiée et devrait être systématiquement testée. Ainsi, dès 1914, Student suspectait la présence d'une relation entre différentes observations localisées. L'introduction de l'espace dans les modèles économétriques n'est ni neutre ni immédiate et les techniques de l'économétrie spatiale visent à prendre en compte la présence de deux importants effets spatiaux : l'autocorrélation spatiale qui se réfère à l'absence d'indépendance entre des observations géographiques et l'hétérogénéité spatiale qui est liée à la différenciation des variables et des comportements dans l'espace. Cet article est consacré au traitement de l'autocorrélation spatiale dans le cadre des modèles de régression linéaire <sup>(1)</sup>.

Historiquement, c'est à Cliff et Ord qu'on doit, après une série d'articles à la fin des années soixante et au début des années soixante-dix, un ouvrage présentant de manière synthétique l'état des savoirs en statistique et en économétrie spatiales (1973). La fin des années soixante-dix et les années quatre-vingt sont marquées par le raffinement du cadre original d'analyse de Cliff et Ord, plus particulièrement par le développement de la théorie de l'estimation et des tests (Ord, 1975 ; Paelinck et Klassen, 1979 ; Anselin, 1980). Un certain nombre d'ouvrages rendent compte de ces développements : Cliff et Ord (1981), Anselin (1988a), Cressie (1993), Jayet (1993), Anselin et Florax (1995a).

Jusqu'à récemment, les techniques permettant de spécifier, d'estimer et de tester la présence de l'autocorrélation spatiale dans les modèles économétriques étaient principalement publiées dans les revues spécialisées et appliquées à des problèmes d'économie régionale, spatiale ou urbaine. Ainsi, Can (1992), Can et Megboluge (1997), Pace et Gilley (1997) ont abordé le traitement de l'autocorrélation spatiale dans les modèles hédoniques de prix immobiliers, alors qu'Anselin, Varga et Acs (1997) et Varga (1998) ont étudié les externalités spatiales d'information dues à la recherche universitaire et la R&D. Toutefois, les méthodes de l'économétrie spatiale sont maintenant appliquées à d'autres sujets tels que l'analyse de la demande (Case, 1991), l'économie internationale (Aten, 1997), l'économie publique (Case *et alii*, 1993 ; Brueckner, 1998), l'économie rurale (Benirschka et Binkley, 1996), ou encore les phénomènes de croissance et de convergence (Moreno et Treban, 1997 ; Rey et Montouri, 1999 ; Fingleton 1999 ; Baumont *et alii*, 2001). Ces méthodes sont en effet potentiellement applicables à toutes les études empiriques nécessitant l'utilisation de données spatiales.

Selon Anselin et Bera (1998) et Anselin (2000), deux raisons principales peuvent être attribuées au regain de l'attention porté à la prise en compte de l'autocorrélation spatiale dans les études empiriques en économie. La première est le développement de nouveaux courants théoriques, tels que l'économie géographique et ses ramifications en économie industrielle, économie internationale, théories de la croissance ou économie du travail..., qui ont généralisé la prise en compte des interactions et des externalités spatiales dans l'analyse des décisions économiques des agents. La modélisation de l'autocorrélation spatiale permet en effet de capter l'existence, l'ampleur et l'influence des effets de débordement géographiques. La deuxième raison, d'ordre empirique, est la disponibilité croissante des données spatialisées, le fort développement actuel des logiciels de Systèmes d'Information Géographiques (SIG) et l'apparition de logiciels, tels que SpaceStat (Anselin, 1999) ou le module *S+SpatialStat pour SPlus* (Mathsoft, 1996), proposant l'estimation des principaux modèles spatiaux.

L'objectif de cet article est de présenter les outils nécessaires à une démarche économétrique visant à prendre en compte l'effet de dépendance spatiale. En effet, sur le plan économétrique, la non vérification de l'hypothèse d'indépendance des observations conduit à remettre en cause l'inférence statistique basée sur l'estimation d'une régression par les Moindres Carrés Ordinaires (MCO). Le test de l'autocorrélation spatiale s'avère alors une précaution indispensable à toute étude empirique mobilisant des données géographiques. Par conséquent, la démarche à adopter s'articule autour de deux opérations : détecter et modéliser l'autocorrélation spatiale. Les tests de l'autocorrélation spatiale permettent à la fois de détecter la présence et la forme de l'autocorrélation dans les régressions, orientant ainsi l'économiste vers la spécification économétrique appropriée. Cependant, dans les études empiriques, on peut aussi disposer *a priori* d'informations théoriques conduisant à préférer une forme particulière de dépendance spatiale entre les observations pour capter différents types d'effets de débordement géographiques. Ainsi, les tests et la modélisation de l'autocorrélation spatiale sont des opérations interdépendantes et nous organiserons notre présentation de la façon suivante.

Il est déjà nécessaire de définir au départ la structure de la dépendance spatiale organisant les données utilisées : la mesure de l'autocorrélation spatiale et sa formalisation à partir de la notion de matrice de poids sont ainsi présentées dans un premier temps (première partie). Ensuite, les différentes modélisations de l'autocorrélation spatiale s'effectuent à l'aide de plusieurs spécifications et nous présenterons les trois spécifications les plus utilisées ainsi que leurs principales propriétés (deuxième partie). Nous abordons ensuite les méthodes d'estimation adaptées à ces spécifications spatiales. Ces modèles se caractérisent par une corrélation entre les erreurs et les variables explicatives. Par conséquent, l'inférence statistique basée sur les MCO n'est pas fiable en

présence d'autocorrélation spatiale et d'autres méthodes d'estimation doivent être utilisées. Nous présentons en particulier la méthode du maximum de vraisemblance, celle des variables instrumentales et des moments généralisés (troisième partie). Enfin, nous exposons les méthodes d'inférence permettant de tester la présence d'autocorrélation spatiale dans les modèles. Les trois principes de tests basés sur le maximum de vraisemblance (multiplicateur de Lagrange, Wald et rapport de vraisemblance) peuvent être mobilisés mais les tests du multiplicateur de Lagrange ont particulièrement été utilisés en économétrie spatiale car ils permettent, à travers diverses règles de décision, de déterminer la forme prise par l'autocorrélation spatiale. En se basant sur les propriétés des différents tests, nous proposons alors une démarche générale (quatrième partie).

---

## L'autocorrélation spatiale : de quoi s'agit-il ?

---

### Définition et sources de l'autocorrélation spatiale

Anselin et Bera (1998) proposent une définition intuitive de l'autocorrélation spatiale : "*Spatial autocorrelation can be loosely defined as the coincidence of value similarity with locational similarity*". En d'autres termes, l'autocorrélation spatiale positive se traduit par une tendance à la concentration dans l'espace de valeurs faibles ou élevées d'une variable aléatoire. En revanche, l'autocorrélation spatiale négative signifie que chaque localisation tend à être entourée par des localisations voisines pour lesquelles la variable aléatoire prend des valeurs très différentes. Enfin, l'absence d'autocorrélation spatiale indique que la répartition spatiale des valeurs de la variable est aléatoire. On notera, qu'en économie, on est généralement confronté à trois types de localisation. Il peut s'agir tout d'abord de points représentant par exemple des localisations d'unités de production ou de distribution... Ces points sont souvent mesurés par leur latitude et leur longitude. Ensuite, ces localisations peuvent être des lignes, connectées entre elles ou non, comme un réseau routier ou fluvial. Enfin, les données sont parfois fournies pour des aires géographiques comme des régions ou des pays. Dans tous les cas, le nombre de ces points, de ces lignes ou de ces zones est supposé fini. Par la suite, nous nous restreindrons, sans perte de généralité, au 3<sup>ème</sup> type de localisation.

La présence d'autocorrélation spatiale pour une variable signifie qu'il y a une relation fonctionnelle entre ce qui se passe en un point de l'espace et ce qui se passe ailleurs. Tobler (1979) l'avait déjà souligné en suggérant la première loi de la géographie suivante : "*Everything is related to everything else, but closer things more so*"<sup>(2)</sup>. L'autocorrélation spatiale diffère donc de l'autocorrélation temporelle qui est unidirectionnelle puisque seul le passé influence le futur. En revanche, l'autocorrélation spatiale est multidirectionnelle puisque "tout est relié à tout". Cette interdépendance généralisée a pour conséquence de rendre plus complexes les méthodes de traitement de l'autocorrélation spatiale. Par exemple, certaines méthodes d'estimation valables pour les séries temporelles ne sont pas directement transposables au cas spatial (voir troisième partie).

L'autocorrélation spatiale a deux sources principales :

- elle peut provenir du fait que les données sont affectées par des processus qui relient des lieux différents et qui sont à l'origine d'une organisation spatiale particulière des activités. En effet, les processus d'interactions sont source d'autocorrélation spatiale lorsque les événements ou les circonstances en un lieu donné affectent les conditions en d'autres lieux si ces derniers interagissent d'une manière ou d'une autre, par des mouvements de biens, de personnes, de capitaux, des externalités spatiales ou toutes les formes de comportements où un acteur économique réagit aux actions d'autres acteurs. Par exemple, la diffusion d'un phénomène (comme la diffusion technologique) à partir d'un ou de plusieurs lieux d'origine implique que l'intensité de la mesure de ce phénomène dépend de la distance à l'origine. Aux localisations proches les unes des autres, donc à des distances comparables de l'origine, seront donc associées des intensités similaires pour le phénomène étudié ;
- elle peut également provenir d'une mauvaise spécification du modèle, comme des variables omises spatialement autocorrélées, d'une forme fonctionnelle incorrecte ou d'erreurs de mesure (c'est en particulier le cas lorsque l'étendue spatiale du phénomène étudié ne coïncide pas avec les unités spatiales d'observation). Elle est alors considérée comme un outil de diagnostic et de détection d'une mauvaise spécification du modèle.

## Les matrices de poids et variables spatiales décalées

Pour capter l'interdépendance entre régions, il faut considérer leurs positions relatives. Pour cela, on doit spécifier de manière exogène la topologie du système spatial en construisant une matrice de poids. Cette matrice est une matrice carrée, ayant autant de lignes et de colonnes qu'il y a de zones géographiques (on note  $N$  le nombre de régions), où chaque terme  $w_{ij}$  représente la façon dont la région  $i$  et la région  $j$  sont connectées spatialement.

Les matrices les plus utilisées sont les *matrices de contiguïté*. La contiguïté entre deux régions se définit par le fait qu'elles ont une frontière commune et chaque terme de cette matrice est égal à 1 si les régions sont contiguës à l'ordre 1 et 0 sinon (par convention, une région n'est pas contiguë avec elle-même :  $w_{ii} = 0, \forall i$ ). Cette notion de contiguïté peut être généralisée : deux régions  $i$  et  $j$  sont contiguës à l'ordre  $k$  si  $k$  est le nombre *minimal* de frontières à traverser pour aller de  $i$  à  $j$ <sup>(3)</sup>.

Ces matrices de contiguïté sont souvent utilisées en raison de leur simplicité mais apparaissent restrictives pour ce qui est de leur définition de la connexion spatiale entre régions. Une autre possibilité consiste à utiliser des *matrices de distance*. On suppose dans ce cas que l'intensité de l'interaction entre deux régions  $i$  et  $j$  dépend de la distance entre les centroïdes de ces régions<sup>(4)</sup>. Plusieurs indicateurs peuvent être utilisés selon la définition de la distance : distance à vol d'oiseau, distance par routes ou généralisation aux temps de transport ou à des indices d'accessibilité. Diverses spécifications sont également disponibles, les plus utilisées étant la fonction exponentielle négative (1) ou une fonction de l'inverse de la distance (2). Si  $d_{ij}$  désigne la distance entre la région  $i$  et la région  $j$ , les éléments de la matrice de distance pour ces deux cas sont définis par :

$$(1) w_{ij} = e^{-\alpha d_{ij}}$$

$$(2) w_{ij} = \begin{cases} d_{ij}^{-\beta} & \text{si } d_{ij} < \bar{d} \\ 0 & \text{sinon} \end{cases}$$

$\alpha$  et  $\beta$  sont des paramètres déterminés *a priori*,  $\bar{d}$  est la valeur seuil au-delà de laquelle on suppose que les régions  $i$  et  $j$  ne sont pas connectées.

Ces spécifications ont été affinées par Cliff et Ord (1981), qui utilisent une combinaison d'une mesure de distance et de la longueur relative de la frontière commune entre ces deux lieux, pour tenir compte de l'irrégularité des zonages. On trouvera d'autres spécifications possibles de la matrice de poids dans Upton et Fingleton (1985) ou Anselin (1988a).

Les matrices de poids sont souvent standardisées : chaque élément  $w_{ij}$  de la matrice est divisé par la somme totale de la ligne  $\sum_j w_{ij}$ . Les poids sont alors compris entre 0 et 1 et cette opération rend les paramètres spatiaux comparables entre les modèles économétriques. L'interprétation des poids est également modifiée. Par exemple, dans le cas d'une matrice de distance, standardiser la matrice de poids signifie qu'on suppose que la connexion entre deux régions dépend de la distance relative entre elles et non plus de la distance absolue.

On définit enfin la *variable spatiale décalée* ("spatial lag"), qui pour  $N$  régions est associée à la matrice de poids  $W$  et la variable aléatoire  $x$ , qui est définie par le vecteur  $(N, 1) : Wx$ . Lorsque  $W$  est une matrice standardisée, le  $i^{\text{ème}}$  élément de la variable spatiale décalée contient la moyenne pondérée des observations des régions voisines à la région  $i$ . Cette variable joue un rôle primordial dans la spécification des modèles économétriques spatiaux (voir deuxième partie).

## Le test de l'autocorrélation spatiale d'une variable quantitative

La statistique la plus utilisée pour tester l'autocorrélation spatiale dans une série est celle de Moran (1948). Cette statistique s'écrit formellement de la façon suivante :

$$(3) I = \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{S_0} \bigg/ \frac{\sum_i (x_i - \bar{x})^2}{N}$$

$$\text{avec : } S_0 = \sum_i \sum_j w_{ij} \text{ et } \bar{x} = (1/N) \sum_i x_i.$$

Le numérateur s'interprète comme la covariance pondérée entre unités voisines alors que le dénominateur représente la variance totale observée. Les expressions analytiques de l'espérance et de la variance de cette statistique sous diverses hypothèses (comme celle de normalité) sont données dans Cliff et Ord (1981). La statistique de Moran centrée et réduite suit asymptotiquement une loi normale d'espérance nulle et de variance unitaire et sert ainsi de base au test de l'autocorrélation spatiale dans une série.

Enfin, pour clore cette partie, il est utile de préciser que d'une façon générale, la mise en œuvre de ce test, ainsi que des estimations et autres tests présentés dans les parties suivantes, nécessite quelques précautions méthodologiques liées à l'utilisation de données spatiales.

D'une part, la mesure de l'autocorrélation spatiale est dépendante de la matrice de poids utilisée : rejeter l'absence ou la présence d'autocorrélation spatiale pour une définition du voisinage n'implique pas toujours qu'on arrivera à la même conclusion avec d'autres définitions du voisinage. Il est donc nécessaire d'évaluer, dans la mesure du possible, la robustesse des résultats obtenus au choix de la matrice de poids.

D'autre part, la façon dont les données spatiales sont agrégées a parfois un effet sur la mesure de l'autocorrélation spatiale. Le "MAUP" ou "Modifiable Areal Unit Problem" (Arbia, 1989) recouvre deux aspects. Premièrement, l'autocorrélation spatiale peut être affectée par le niveau d'agrégation utilisé, c'est l'effet d'échelle (Chou, 1991). C'est le cas par exemple lorsque les résultats varient selon qu'on utilise les données sur les régions européennes au niveau NUTS1 ou au niveau NUTS2. Deuxièmement, l'autocorrélation spatiale est aussi sensible à la forme des unités spatiales : il y en a beaucoup de façons de découper une région en plusieurs subdivisions, ce qui donne lieu à de nombreuses configurations spatiales (voir par exemple Griffith, 1992). Evidemment, le mode d'agrégation des données spatiales est très souvent imposé à l'économiste à travers la disponibilité des banques de données.

---

## Autocorrélation spatiale et modèles économétriques

---

La prise en compte de l'autocorrélation spatiale dans les modèles économétriques peut s'effectuer de plusieurs manières : par des *variables spatiales décalées*, endogènes ou exogènes, ou par une *autocorrélation spatiale des erreurs*. En coupe transversale, nous prendrons comme point de départ le modèle de régression linéaire classique suivant :

$$(4) \quad y = X \beta + \varepsilon$$

Soient les conventions d'écriture suivantes :  $N$  est le nombre d'observations,  $y$  est le vecteur  $(N,1)$  des observations de la variable dépendante,  $X$  est la matrice  $(N,K)$  des observations des variables exogènes,  $K$  étant le nombre de paramètres inconnus,  $\beta$  est le vecteur  $(K,1)$  des coefficients de régression inconnus,  $\varepsilon$  est le vecteur  $(N,1)$  des erreurs. Pour les erreurs, sauf indication contraire, on supposera :

$H_1$  : Les erreurs  $\varepsilon_i$  sont identiquement et indépendamment distribuées (*iid*), d'espérance nulle et de variance finie  $\sigma^2$  ;

$H_2$  : La matrice  $X$  est non-stochastique, de rang complet  $K$  et  $\lim_{n \rightarrow \infty} (1/N) X'X = Q$  où  $Q$  est une matrice finie et non-singulière.

### Variable endogène décalée

Dans le modèle autorégressif spatial, une "variable endogène décalée"  $Wy$  est incluse dans le modèle (4) :

$$(5) \quad y = \rho Wy + X \beta + \varepsilon$$

$Wy$  est la variable endogène décalée pour la matrice de poids  $W$ ,  $\rho$  est le paramètre spatial autorégressif indiquant l'intensité de l'interaction existant entre les observations de  $y$ .

Dans ce modèle, l'observation  $y_i$  est en partie expliquée par les valeurs prises par  $y$  dans les régions voisines :

$$(Wy)_i = \sum_{j \neq i} w_{ij} y_j$$

$(Wy)_i$  s'interprète comme la moyenne des valeurs de  $y$  sur les observations voisines à  $i$  lorsque  $W$  est standardisée. Cette standardisation facilite aussi la comparaison de l'ampleur de  $\rho$  lorsque (5) est estimé pour plusieurs matrices de poids.

L'introduction de  $Wy$  dans le modèle (4) est un moyen d'apprécier le degré de dépendance spatiale alors que les autres variables sont contrôlées. Symétriquement, il permet de contrôler la dépendance spatiale pour évaluer l'impact des autres variables explicatives. Lorsqu'une variable endogène décalée est ignorée dans la spécification du modèle, mais présente dans le processus générateur des données, les estimateurs des MCO dans le modèle aspatial (4) seront biaisés et non convergents.

Ce modèle autorégressif a été utilisé, par exemple, pour modéliser les interactions stratégiques et la concurrence fiscale entre communes (Case *et alii*, 1993 ; Brueckner, 1998) ou encore les externalités de voisinage dans les modèles hédoniques de prix immobiliers (Can, 1992 ; Macedo, 1998).

### Variable exogène décalée

Une autre façon de traiter l'interdépendance des observations est d'inclure une ou plusieurs "variables exogènes décalées" dans (4) :

$$(6) \quad y = X\beta + WZ\delta + \varepsilon$$

$Z$  est une matrice de dimension  $(N, L)$  de  $L$  diverses variables correspondant ou non aux variables incluses dans  $X$ ,  $WZ$  est l'ensemble des variables exogènes décalées pour la matrice de poids  $W$  et  $\delta$  est le vecteur des paramètres spatiaux indiquant l'intensité de la corrélation spatiale existant entre les observations de  $y$  et celles de  $Z$ .

Ainsi, dans ce modèle, l'observation  $y_i$  est expliquée par les valeurs prises par les variables de  $X$  dans la région  $i$  et par les variables de  $Z$  dans les régions voisines. Par exemple, cette spécification permet d'estimer les effets de débordement géographiques liés aux migrations et aux infrastructures publiques : la production d'une région peut être influencée par la disponibilité du travail ou le montant du capital public dans les régions voisines (Kelejian et Robinson, 1997 ; Moreno *et alii*, 1997 ; Boarnet, 1998).

Au contraire du modèle autorégressif (5) et du modèle comportant une autocorrélation spatiale des erreurs (voir *infra*), l'estimation du modèle régressif croisé peut être basée sur les MCO si  $H_1$  est vérifiée pour  $X^* = [X \ WZ]$  de dimension  $(N, K + L)$ .

### Autocorrélation des erreurs

Pour spécifier une autocorrélation spatiale des erreurs, on utilise le plus souvent un processus autorégressif sur les erreurs<sup>(5)</sup> :

$$(7) \quad \begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \end{aligned}$$

Le paramètre  $\lambda$  reflète l'intensité de l'interdépendance entre les résidus et  $u$  est le terme d'erreur tel que :  $u \sim iid(0, \sigma^2 I)$ . Omettre à tort une autocorrélation spatiale des erreurs produit des estimateurs non biaisés mais inefficients.

Le modèle (7) peut se réécrire sous la forme du *modèle de Durbin spatial contraint* où apparaissent une variable endogène décalée et l'ensemble des variables exogènes décalées :

$$(8) \quad y = \lambda Wy + X\beta - \lambda WX\beta + u$$

L'équivalence entre (8) et (7) impose alors un ensemble de contraintes non-linéaires sur les coefficients qui font l'objet d'un test (quatrième partie, test du facteur commun).

La détection de l'autocorrélation spatiale des erreurs s'interprète souvent comme un problème dans la spécification du modèle, telle que l'omission de variables pertinentes : l'effet qui n'est pas capté dans les variables explicatives se retrouve dans les erreurs sous la forme d'une autocorrélation spatiale. Ce modèle a alors été utilisé pour améliorer les estimations des modèles de prix hédoniques (Pace et Gilley, 1997 ; Dubin, 1998 ; Dubin *et alii*, 1999), ou de convergence conditionnelle (Fingleton, 1999). Il a également permis de modéliser la diffusion d'un choc aléatoire portant sur le *PIB* d'une région vers les autres régions aux États-Unis (Rey et Montouri, 1999) ou en Europe (Baumont *et alii*, 2001).

## Encadré 1 : quelques propriétés des modèles spatiaux

### Modèle autorégressif d'ordre 1

*Effet de diffusion et effet multiplicateur*

Supposons la matrice  $(I - \rho W)$  non-singulière<sup>(7)</sup>. Le modèle (5) se réécrit alors sous la forme suivante :

$$(1) y = (I - \rho W)^{-1} X \beta + (I - \rho W)^{-1} \varepsilon$$

L'espérance mathématique de  $y$  s'écrit :  $E(y) = (I - \rho W)^{-1} X \beta$ . La matrice inverse  $(I - \rho W)^{-1}$  est une matrice pleine qui implique une série infinie pour les variables explicatives et pour le terme d'erreur associé à toutes les localisations :

$$(2) y = (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) X \beta + (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) \varepsilon$$

– Concernant les variables explicatives, cette expression signifie qu'en moyenne la valeur de  $y$  dans une région  $i$  n'est pas seulement expliquée par les valeurs des variables explicatives associées à cette région, mais aussi par celles associées à toutes les régions (voisines de  $i$  ou non) à travers la transformation spatiale inverse  $(I - \rho W)^{-1}$ . *Cet effet de multiplicateur spatial décline avec l'éloignement.*

– Concernant le processus des erreurs, cette expression signifie qu'un choc aléatoire dans une région  $i$  affecte non seulement la valeur de  $y$  de cette région, mais a également un impact sur les valeurs de  $y$  dans les autres régions à travers la même transformation. C'est l'effet de diffusion, effet qui décline aussi avec l'éloignement.

*Matrice des variances-covariances*

De (1), on déduit la matrice des variances-covariances de  $y$  :

$$(3) V(y) = \sigma^2 [(I - \rho W') (I - \rho W)]^{-1}$$

Cette matrice des variances-covariances est pleine ; un processus spatial autorégressif conduit donc à une covariance non nulle entre chaque paire d'observations et décroissante avec l'ordre de proximité.

### Modèle avec autocorrélation des erreurs

*Effet de diffusion et effet multiplicateur*

Si  $(I - \lambda W)$  est non-singulière, le modèle (7) se réécrit sous la forme suivante :

$$(4) y = X \beta + (I - \lambda W)^{-1} u$$

Cette expression fait apparaître un effet de diffusion spatiale comme pour le modèle (5) mais comme  $E(y) = X \beta$ , il n'y a pas d'effet de multiplicateur spatial.

*Matrice des variances-covariances*

De la formulation (4), il s'ensuit que :

$$(5) V(y) = V(\varepsilon) = \sigma^2 [(I - \lambda W') (I - \lambda W)]^{-1}$$

De (5), on déduit que la covariance entre chaque paire d'erreurs est non nulle et décroissante avec l'ordre de proximité. Par ailleurs, la structure d'erreurs (5) induit des éléments de la diagonale de  $V(\varepsilon)$  non constants, ce qui implique l'hétéroscédasticité des erreurs  $\varepsilon$ , que  $u$  soit hétéroscédastique ou non (McMillen, 1992)

L'encadré 1 fournit quelques propriétés des modèles spatiaux présentés dans cette partie.

### Généralisation

Jusqu'à présent, les modèles présentés n'incluent qu'un seul type de dépendance spatiale (variable endogène décalée ou autocorrélation des erreurs) et qu'un seul ordre de dépendance. Différents auteurs ont proposé et étudié des processus beaucoup plus généraux comportant plusieurs décalages et/ou comportant à la fois une variable autorégressive et une autocorrélation des erreurs (Brandsma et Kelletaper, 1979 ; Huang, 1984 ; Anselin, 1988a ; Jayet, 1993). Par exemple, considérons le modèle général suivant :

$$(9) \begin{aligned} y &= \rho_1 W_1 y + \rho_2 W_2 y + \dots + \rho_p W_p y + \varepsilon \\ \varepsilon &= \lambda_1 W_1 \varepsilon + \lambda_2 W_2 \varepsilon + \dots + \lambda_q W_q \varepsilon + u \end{aligned}$$

En général, les différentes matrices de poids  $W_i$  sont associées au  $i^{\text{ème}}$  ordre de contiguïté.

Anselin (1980, 1988a) a étudié les propriétés du processus contenant à la fois une variable endogène décalée et une autocorrélation des erreurs. Formellement, le modèle s'exprime comme une combinaison de (5) et de (7), avec des poids différents :

$$(10) \quad \begin{aligned} y &= \rho W_1 y + X \beta + \varepsilon \\ \varepsilon &= \lambda W_2 \varepsilon + u \end{aligned}$$

Ce modèle peut se réécrire sous la forme :  $y = \rho W_1 y + \lambda W_2 y - \rho \lambda W_2 W_1 y + X \beta - \lambda W_2 X \beta + u$ . Il s'agit d'une forme de Durbin étendue avec des contraintes non-linéaires supplémentaires<sup>(6)</sup>.

Les méthodologies d'estimation et d'inférence adaptées aux modèles à plusieurs décalages sont basées sur les mêmes principes que les modèles à un seul décalage. On pourra consulter Brandsma et Kelletaper (1979), Huang (1984), Jayet (1993) ou Anselin et Florax (1995b) pour plus de détails sur ces modèles. Dans la suite de l'exposé, nous nous limiterons aux modèles à un seul décalage.

## Estimation des modèles spatiaux

En présence d'autocorrélation spatiale, les estimateurs MCO ne sont pas convergents (variable autorégressive) et ils sont inefficients (autocorrélation spatiale des erreurs). Des estimateurs convergents et asymptotiquement efficaces sont alors obtenus grâce à la méthode du maximum de vraisemblance, celle des variables instrumentales ou des moments généralisés.

### La non-convergence des MCO et ses conséquences

Prenons le modèle spatial le plus général (10). Il se réécrit de la façon suivante :

$$y = (I - \rho W_1)^{-1} X \beta + (I - \rho W_1)^{-1} \varepsilon$$

$$\text{et } \varepsilon = (I - \lambda W_2)^{-1} u.$$

Par conséquent, la matrice des variances-covariances de l'erreur  $\varepsilon$  est :  $V(\varepsilon) = E(\varepsilon \varepsilon') = \sigma^2 (I - \lambda W_2)^{-1} (I - \lambda W_2')^{-1}$ .

La corrélation entre la variable explicative  $W_1 y$  et l'erreur  $\varepsilon$  s'écrit donc :

$$(11) \quad E(W_1 y \varepsilon') = E[W_1 (I - \rho W_1)^{-1} X \beta \varepsilon' + W_1 (I - \rho W_1)^{-1} \varepsilon \varepsilon']$$

$$(12) \quad E(W_1 y \varepsilon') = \sigma^2 W_1 (I - \rho W_1)^{-1} (I - \lambda W_2)^{-1} (I - \lambda W_2')^{-1}$$

Puisque l'expression (12) est non nulle en général, les éléments de la variable endogène décalée sont corrélés avec ceux des erreurs et les MCO ne fournissent donc pas d'estimateurs convergents des paramètres du modèle (10) (Kelejian et Prucha, 1998).

– Pour le modèle spatial autorégressif (5), les estimateurs des MCO ne sont pas convergents car la variable endogène décalée  $W y$  est corrélée avec l'erreur  $\varepsilon$ , quelle que soit la distribution de cette erreur : si  $\lambda = 0$  et  $W_1 = W$  dans (12), alors  $E(W y \varepsilon') = \sigma^2 W (I - \rho W)^{-1}$ . Ce résultat contraste avec une propriété de séries temporelles où les estimateurs des MCO restent convergents en présence d'une ou de plusieurs variables retardées tant que les erreurs ne sont pas corrélées.

– Pour le modèle à erreurs autorégressives (7), l'estimateur de  $\beta$  par les MCO est sans biais mais inefficace puisque les erreurs sont non sphériques. La solution théorique consistant à employer les moindres carrés quasi-généralisés (MCQG) n'est pas applicable au cas spatial puisque l'estimateur de  $\lambda$  par les MCO n'est pas convergent (Anselin, 1988a, chap. 6). Sous des hypothèses similaires à celles utilisées ici, Kelejian et Prucha (1997) montrent de plus que les estimateurs obtenus par la méthode des variables instrumentales non linéaire appliquée au modèle spatial de Durbin (8) n'est pas non plus convergente, l'une des conditions données par Amemiya (1985) n'étant pas vérifiée.

### Estimation par le maximum de vraisemblance

Les conditions pour la convergence, l'efficacité et la normalité asymptotique des estimateurs du maximum de vraisemblance (MV) sont dérivées du cadre général de Heijmans et Magnus (1986a, 1986b) pour le modèle spatial autorégressif et de celui de Magnus (1978) pour le modèle avec autocorrélation spatiale des erreurs. En plus des restrictions habituelles sur la variance et les moments d'ordre supérieur des variables du modèle, ces conditions se traduisent par des contraintes sur les poids et sur l'espace des paramètres des coefficients spatiaux (Anselin et Kelejian, 1997 ; Kelejian et Prucha, 1998, 1999)<sup>(8)</sup>.

Pour le modèle général (10), le point de départ est l'hypothèse de normalité des termes d'erreur :  $u \sim Nid(0, \sigma^2 I)$ . Puisque dans (10),

$u = (I - \lambda W_2)(y - \rho W_1 y - X\beta)$ , le jacobien de la transformation est :

$$J = \det(\partial u / \partial y) = |I - \rho W_1| |I - \lambda W_2|$$

La fonction de log-vraisemblance de  $y$  s'écrit donc :

$$(13) \ln L(y | \rho, \lambda, \beta) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) + \ln |I - \rho W_1| + \ln |I - \lambda W_2| - \frac{1}{2\sigma^2} u'u$$

De cette expression, on déduit l'espace des paramètres de  $\rho$  et  $\lambda$ . Soient  $\omega_{\max}$  la valeur propre positive maximale et  $\omega_{\min}$  la valeur propre négative la plus grande en valeur absolue de  $W$ . Alors on démontre que  $|I - \rho W_1|$  et  $|I - \lambda W_2|$  sont positifs si l'inégalité  $1/\omega_{\min} < \psi < 1/\omega_{\max}$  (avec  $\psi = \rho$  ou  $\lambda$ ) est respectée. Pour une matrice standardisée en ligne,  $\omega_{\max} = 1$  et  $\omega_{\min} > -1$ .

On trouvera dans Anselin (1988a, 1988b), les expressions du score et de la matrice d'information pour le modèle général (10). Le système d'équations résultant des conditions du premier ordre n'admet pas dans ce cas de solution analytique. En revanche, les systèmes correspondant aux modèles (5) et (7) admettent des solutions issues des conditions du premier ordre permettant de construire une fonction de log-vraisemblance concentrée. On pourra se référer à Anselin (1988a) et Le Gallo (2000) pour plus de détails sur les formes prises par les estimateurs dans ces deux cas particuliers. Le problème posé par le jacobien est abordé dans l'encadré 2.

#### Encadré 2 : le jacobien

Le problème principal de l'estimation par le maximum de vraisemblance est la présence du jacobien puisque la maximisation de la fonction de log-vraisemblance nécessite son évaluation pour chaque nouvelle valeur de  $\rho$  ou de  $\lambda$ . Il s'agit d'une opération lourde, même pour de petits échantillons, dans la mesure où il faut calculer le déterminant d'une matrice carrée de dimension  $N$ .

Plusieurs solutions ont été proposées dans la littérature (Le Gallo, 2000). La plus ancienne et encore la plus utilisée a été proposée par Ord (1975). Elle consiste à exploiter la décomposition du jacobien en termes des  $N$  valeurs propres de la matrice de poids  $W$ :

$$(1) |I - \rho W| = \prod_{i=1}^N (1 - \rho \omega_i) \Rightarrow \ln |I - \rho W| = \sum_{i=1}^N \ln(1 - \rho \omega_i)$$

Ainsi, au lieu de calculer le déterminant de  $I - \rho W$  à chaque étape, il n'est besoin de calculer qu'une fois pour toutes les valeurs propres de  $W$  si l'on utilise cette formulation simplifiée. La fonction de vraisemblance globale se réduit alors à la somme des vraisemblances individuelles. Certains auteurs fournissent des codes pour quelques logiciels économétriques basés sur cette propriété (Anselin et Hudak, 1992 ; Griffith, 1993 ; Li, 1996).

## Autres méthodes d'estimation

Dans le modèle autorégressif, le principal problème est la corrélation entre la variable endogène décalée et le terme d'erreur. Dans ce cas, la méthode des variables instrumentales a été proposée par Anselin (1988a), Kelejian et Robinson (1993) ou Kelejian et Prucha (1998). Les estimateurs sont convergents et efficaces compte tenu d'un choix approprié d'instruments. En pratique, les variables explicatives  $X$ , stochastiques ou non, indépendantes des erreurs, doivent nécessairement figurer dans les instruments puisqu'elles ne sont pas corrélées avec les erreurs. Pour la variable endogène décalée, on pourra prendre  $WX$  comme instrument ou des décalages d'ordre supérieur. En notant  $Z$  la matrice  $(N, P)$  des instruments ( $P \geq K + 1$ ), l'estimateur des variables instrumentales s'écrit :

$$(14) \hat{\beta}_{IV} = [\tilde{X}'Z(Z'Z)^{-1}Z'\tilde{X}]^{-1}\tilde{X}'Z(Z'Z)^{-1}Z'y$$

avec :

$$\tilde{X} = [Wy \ X], \quad \text{Asy } V(\hat{\beta}_{IV}) = \sigma^2 [\tilde{X}'Z(Z'Z)^{-1}Z'\tilde{X}]^{-1} \quad \text{et} \quad \sigma^2 = (y - \tilde{X}\beta_{IV})'(y - \tilde{X}\beta_{IV})/N.$$

Cette approche a été étendue à des structures d'erreurs plus complexes (Anselin, 1988a). Sous des hypothèses raisonnables satisfaites lorsque les poids sont basés sur la contiguïté, les estimateurs des variables instrumentales sont convergents et asymptotiquement normaux.

Comme la méthode des variables instrumentales ne fournit pas d'estimateurs convergents pour le coefficient spatial dans le modèle à erreurs autocorrélées, Kelejian et Prucha (1998, 1999) ont récemment développé une approche basée sur la méthode des moments généralisés (GMM). Ils développent un ensemble de conditions sur les moments permettant l'estimation des équations pour les paramètres dans le modèle à erreurs autocorrélées (7).

Si  $u \sim \text{iid}(0, \sigma^2 I)$ , les trois conditions sont les suivantes :

$$(15) \begin{aligned} E[u'u/N] &= \sigma^2 \\ [u'W'Wu/N] &= \sigma^2 (1/N) \text{tr}(W'W) \\ E[u'Wu/N] &= 0 \end{aligned}$$

Remplacer  $u$  par  $\hat{\varepsilon} - \lambda W \hat{\varepsilon}$  ( $\hat{\varepsilon}$  étant le vecteur des résidus des MCO), donne un système de trois équations pour les paramètres  $\lambda$ ,  $\lambda^2$  et  $\sigma^2$ . Dans cette approche,  $\lambda$  est considéré comme un paramètre de nuisance et on ne peut pas tester sa significativité<sup>(9)</sup>.

---

## Les tests en économétrie spatiale

---

### Principes de tests et économétrie spatiale

Pour tester la présence d'autocorrélation spatiale, les trois grands principes de tests basés sur les propriétés asymptotiques optimales de l'estimation par le maximum de vraisemblance peuvent être utilisés : le test de Wald, le test du rapport de vraisemblance et le test du multiplicateur de Lagrange. Par la suite, si l'on cherche à tester un ensemble de restrictions sur un vecteur de paramètres, nous appellerons modèle contraint le modèle où ces restrictions sont respectées par opposition au modèle non-contraint (ou modèle libre) où elles ne le sont pas. Dans ces conditions, le test de Wald exige l'estimation du modèle non-contraint, le test du ratio de vraisemblance nécessite à la fois l'estimation du modèle contraint et du modèle non-contraint et le test du multiplicateur de Lagrange (ou test du score) nécessite uniquement l'estimation du modèle contraint.

En économétrie spatiale, le test de Wald est plus naturellement mobilisé pour effectuer des tests quand le modèle spatial choisi a été estimé par la méthode appropriée. Par exemple, dans le cas de tests simples de significativité sur un coefficient, il s'agit alors de calculer le ratio du coefficient estimé à son écart-type asymptotique estimé. Les tests du ratio de vraisemblance sont quant à eux relativement simples à évaluer puisqu'ils sont obtenus comme le double de la différence entre la fonction de log-vraisemblance évaluée sous l'hypothèse alternative et la fonction de log-vraisemblance évaluée sous l'hypothèse nulle. Enfin, les tests du multiplicateur de Lagrange s'avèrent particulièrement importants. D'une part, leurs formulations analytiques sont moins complexes que celles des tests de Wald correspondants. D'autre part, ces tests nécessitent uniquement l'estimation du modèle

contraint qui, dans le cas d'hypothèses simples sur les paramètres, est moins complexe à estimer que le modèle non-contraint : il s'agit souvent du modèle de régression simple estimé par les MCO. Ces tests s'avèrent donc particulièrement utiles pour rechercher la spécification du modèle. En effet, ils permettent d'abord de vérifier si l'estimation de modèles plus complexes est nécessaire sans avoir à mettre en œuvre immédiatement l'estimation de modèles spatiaux qui soulève d'importantes difficultés. Ensuite, si un modèle spatial s'avère effectivement nécessaire, ils permettent d'orienter le choix d'une spécification lorsqu'on n'a pas d'*a priori* sur la forme prise par l'autocorrélation spatiale ou de vérifier que le modèle choisi sur une base théorique est correctement spécifié (par exemple, pour vérifier qu'il ne reste pas une autocorrélation des erreurs si on a choisi un modèle avec variable endogène décalée). Dans les paragraphes suivants, nous nous concentrerons donc plus particulièrement sur les tests du multiplicateur de Lagrange.

### Les tests de l'autocorrélation spatiale

Mis à part les tests du multiplicateur de Lagrange, la statistique de Moran (1950a, 1950b) vise également à tester l'autocorrélation spatiale. En revanche, nous n'évoquerons pas le test de Kelejian-Robinson (1992) en raison de sa faible puissance (Anselin et Florax, 1995b).

#### Le test de Moran

Le test de Moran présenté dans la première partie n'était pas lié à un modèle statistique spécifique. Néanmoins, ce test a été adapté pour tester l'autocorrélation spatiale dans les résidus de la régression (4) estimée par les MCO et se présente formellement dans ce cas de la façon suivante en notation matricielle :

$$(16) \quad I = \frac{N}{S_0} \left( \frac{\tilde{\varepsilon}' W \tilde{\varepsilon}}{\tilde{\varepsilon}' \tilde{\varepsilon}} \right)$$

où  $\tilde{\varepsilon} = y - X\tilde{\beta}$  est le vecteur des résidus estimés des MCO et  $S_0$  est un facteur de standardisation égal à la somme de tous les éléments de  $W$ . Cette statistique se simplifie pour une matrice standardisée puisque dans ce cas  $S_0 = N$ .

Sous l'hypothèse nulle d'indépendance spatiale, le test  $I$  de Moran est un test localement meilleur invariant (King, 1981). Cliff et Ord (1973) ont dérivé les deux premiers moments de la statistique  $I$  sous l'hypothèse nulle :

$$(17) \quad E(I) = \frac{tr(MW)}{N-K} \text{ et } V(I) = \frac{tr(MW) + \{tr(MW)\}^2}{(N-K)(N-K+2)} - [E(I)]^2$$

où  $M$  est la matrice symétrique et idempotente habituelle :  $M = I - (X'X)^{-1}X'$ .

Le test se base alors sur la statistique de Moran centrée et réduite :  $Z(I) = [I - E(I)] / \sqrt{V(I)}$ . Pour des résidus normalement distribués et une matrice de poids "qui se comporte bien",  $Z(I)$  suit asymptotiquement une loi normale centrée et réduite. Anselin et Kelejian (1997) et Pinkse (1998) donnent ainsi des conditions formelles pour la normalité asymptotique du test de Moran dans plusieurs autres types de modèles<sup>(10)</sup>.

#### Tests d'une hypothèse simple

– On considère tout d'abord le cas où les erreurs suivent un processus spatial autorégressif (7) :  $\varepsilon = \lambda W \varepsilon + u$  pour lequel on teste  $H_0 : \lambda = 0$ . Sous l'hypothèse nulle, on retrouve le modèle linéaire classique (4).

La statistique de test du multiplicateur de Lagrange s'écrit de la façon suivante :

$$(18) \quad LM_{ERR} = \frac{[\tilde{\varepsilon}' W \tilde{\varepsilon} / \tilde{\sigma}^2]^2}{T}$$

où  $T = tr[(W' + W)W]$ ,  $\tilde{\varepsilon}$  et  $\tilde{\sigma}^2$  sont les estimations de  $\varepsilon$  et  $\sigma^2$  sous  $H_0$ . Puisqu'il n'y a qu'une seule contrainte,  $LM_{ERR}$  converge asymptotiquement vers une loi du chi-deux à un degré de liberté :  $LM_{ERR} \rightarrow \chi_1^2$ .

– Le test d'une variable endogène décalée a été proposé par Anselin (1988a). Soit l'hypothèse nulle  $H_0 : \rho = 0$  dans (5). La statistique de test du multiplicateur de Lagrange est la suivante :

$$(19) LM_{LAG} = \frac{[\tilde{\varepsilon}'W\tilde{y} / \tilde{\sigma}^2]^2}{\tilde{T}_1}$$

$$\text{où } \tilde{T}_1 = \left[ (WX\tilde{\beta})(I - X(X'X)^{-1}X')(WX\tilde{\beta}) + T\tilde{\sigma}^2 \right] / \tilde{\sigma}^2$$

$\tilde{\beta}$  et  $\tilde{\sigma}^2$  sont les estimations de  $\beta$  et  $\sigma^2$  sous  $H_0$ . Ce test converge également vers une loi du chi-deux à un degré de liberté.

Les formulations analytiques des tests de Wald et du rapport de vraisemblance pour une variable endogène décalée et pour une autocorrélation spatiale des erreurs et leur dérivation à partir des fonctions de log-vraisemblance associées sont détaillées dans Anselin (1988a) et Anselin et Bera (1998).

#### *Les tests en présence d'une autocorrélation des erreurs et d'une variable décalée*

Il est utile de savoir si le modèle correct contient à la fois une autocorrélation des erreurs et une variable autorégressive. Anselin et Bera (1998) remarquent en effet que  $LM_{ERR}$  est la statistique de test correspondant à  $H_0 : \lambda = 0$  en supposant une spécification correcte pour le reste du modèle, c'est-à-dire  $\rho = 0$ . En revanche, si  $\rho \neq 0$ , ce test n'est plus valide, même asymptotiquement, et il n'est plus distribué selon une loi du chi-deux centré à un degré de liberté. Pour une inférence statistique valide, il faut donc prendre en compte la possible variable endogène décalée lorsqu'on teste l'autocorrélation spatiale des erreurs et vice-versa. Plusieurs stratégies sont possibles :

– la première approche consiste à tester l'hypothèse nulle jointe  $H_0 : \lambda = \rho = 0$  dans le modèle (10) grâce au principe du multiplicateur de Lagrange. Ainsi, le test, noté *SARMA*, peut être effectué à partir des résidus des MCO dans le modèle simple (4). Sous  $H_0$ , *SARMA* converge asymptotiquement vers une loi du chi-deux à deux degrés de liberté. Si l'hypothèse nulle est rejetée, on ne connaît pas la nature exacte de la dépendance spatiale ;

– la deuxième approche consiste à faire un test du multiplicateur de Lagrange pour une forme de dépendance spatiale lorsque l'autre forme n'est pas contrainte. Par exemple, on teste  $H_0 : \lambda = 0$  en présence de  $\rho$ . Sous l'hypothèse nulle, on retrouve le modèle (5) alors que sous l'hypothèse alternative, on retrouve le modèle général (10). Le test, noté  $LM_{ERR}^*$ , est alors basé sur les résidus de l'estimation *MV* dans le modèle (5). Sous  $H_0$ ,  $LM_{ERR}^* \rightarrow \chi_1^2$ . On peut également tester l'hypothèse nulle  $H_0 : \rho = 0$  en présence de  $\lambda$ , le test, noté  $LM_{LAG}^*$  est alors basé sur les résidus de l'estimation *MV* dans le modèle avec autocorrélation des erreurs (7). Sous  $H_0$ ,  $LM_{LAG}^* \rightarrow \chi_1^2$  ;

– la dernière approche est celle de Bera et Yoon (1993) qui a été reprise par Anselin *et alii* (1996). Elle consiste à utiliser des tests robustes à une mauvaise spécification locale. Par exemple, il s'agit d'ajuster  $LM_{ERR}$  pour que sa distribution asymptotique reste un  $\chi^2$  centré, même en présence locale de  $\rho$ . Ce test s'effectue à partir des résidus des MCO du modèle simple (4). De la même façon, les résidus du modèle (4) servent au test de  $H_0 : \rho = 0$  dans la présence locale de  $\lambda$ . Ces deux tests convergent asymptotiquement vers une loi du chi-deux à un degré de liberté.

La formulation analytique de ces tests est fournie dans l'encadré 3 (page suivante).

#### *Puissance et robustesse des tests*

Des simulations récentes de Monte-Carlo effectuées en particulier par Anselin et Rey (1991), Florax et Rey (1995) et Anselin et Florax (1995b) fournissent quelques indications quant aux performances de ces tests asymptotiques en échantillon fini.

Anselin et Rey (1991) et Anselin et Florax (1995b) ont comparé les performances des tests de Moran,  $LM_{ERR}$ ,  $LM_{LAG}$ ,  $RLM_{ERR}$  et  $RLM_{LAG}$  pour différentes matrices de poids (sur zonage régulier ou non), différentes distributions des erreurs et différentes tailles d'échantillon. Plusieurs résultats ressortent. Premièrement,  $LM_{LAG}$  est le test le plus puissant et robuste à la non-normalité des erreurs. Deuxièmement, le test *I* de Moran est puissant pour les deux alternatives, variable endogène décalée ou autocorrélation des erreurs. Ce test devient ainsi un indicateur général d'une mauvaise spécification du modèle, quelle que soit la forme de la dépendance spatiale omise mais ne peut discriminer entre les deux formes possibles. Troisièmement, les tests  $LM_{ERR}$  et  $LM_{LAG}$  ont les plus grandes puissances pour leur alternative respective alors que les tests ajustés  $RLM_{ERR}$  et  $RLM_{LAG}$  ont également de bonnes performances en termes de puissance et de taille nominale. Ces tests du score sont donc à privilégier dans la recherche de la spécification du modèle. Enfin, Florax et Rey (1995) ont étudié les conséquences d'une mauvaise spécification de la matrice de poids sur la puissance des tests

### Encadré 3 : tests d'autocorrélation spatiale - formulation analytique

#### SARMA

$$SARMA = \tilde{E}^{-1} \left[ (\tilde{d}_\lambda)^2 \frac{\tilde{D}}{\tilde{\sigma}^2} + (\tilde{d}_\rho)^2 T_{22} - 2\tilde{d}_\lambda \tilde{d}_\rho T_{12} \right]$$

où  $\tilde{d}_\lambda$  et  $\tilde{d}_\rho$  sont respectivement les scores par rapport à  $\lambda$  et  $\rho$  évalués sous  $H_0$ ,

$$\begin{aligned} T_{ij} &= tr [W_i W_j + W_i' W_j'] \\ \tilde{D} &= (W_1 X \tilde{\beta})' M (W_1 X \tilde{\beta}) + T_{11} \tilde{\sigma}^2, \\ E &= (\tilde{D} / \tilde{\sigma}^2) T_{22} - (T_{12})^2 \end{aligned}$$

#### $LM_{ERR}^*$

$$LM_{ERR}^* = \frac{\hat{d}_\rho^2}{T_{22} - (T_{21A})^2 \hat{V}(\hat{\rho})}$$

où  $T_{21A} = tr [W_2 W_1 A^{-1} + W_2' W_1 A^{-1}]$ ,  $A = I - \hat{\rho} W_1$ , le "chapeau" désigne les estimateurs MV dans le modèle (5) obtenus par optimisation non linéaire,  $\hat{V}(\hat{\rho})$  est la variance de  $\hat{\rho}$  dans le modèle (5).

#### $LM_{LAG}^*$

$$LM_{LAG}^* = \frac{[\hat{\varepsilon}' B' B W_1 y]^2}{H_\rho - H_{\theta\rho} \hat{V}(\hat{\theta}) H'_{\theta\rho}}$$

où  $\hat{\varepsilon}$  est le vecteur des résidus estimés par le MV dans le modèle avec erreurs autorégressives (7),

$\theta = (\beta', \lambda, \sigma^2)$ ,  $B = I - \hat{\lambda} W_2$ ,  $\hat{V}(\hat{\theta})$  est la matrice des variances-covariances estimée de  $\hat{\theta}$  dans le modèle (7). Les autres termes sont :

$$H_\rho = tr W_1^2 + tr (B W_1 B^{-1})' (B W_1 B^{-1}) + \frac{(B W_1 X \beta)' (B W_1 X \beta)}{\sigma^2}$$

et

$$H'_{\theta\rho} = \begin{bmatrix} \frac{(B X)' B W_1 X \beta}{\sigma^2} \\ tr (W_2 B^{-1})' B W_1 B^{-1} + tr W_2 W_1 B^{-1} \\ 0 \end{bmatrix}$$

#### $RLM_{ERR}$

$$RLM_{ERR} = \frac{[\tilde{d}_\lambda - T_{12} \tilde{\sigma}^2 \tilde{D}^{-1} \tilde{d}_\rho]^2}{T_{22} - (T_{12})^2 \tilde{\sigma}^2 \tilde{D}}$$

#### $RLM_{LAG}$

$$RLM_{LAG} = \frac{[\tilde{d}_\rho - T_{12} T_{22}^{-1} \tilde{d}_\lambda]^2}{\tilde{\sigma}^2 \tilde{D} - (T_{12})^2 T_{22}^{-1}}$$

précédents. Une sur-spécification cause une baisse de puissance alors qu'une sous-spécification augmente la puissance des tests en présence d'autocorrélation spatiale positive et la diminue en cas d'autocorrélation spatiale négative<sup>(11)</sup>. Le coefficient de Moran est moins affecté que les autres tests par une mauvaise spécification de la matrice de poids.

### À la recherche de la spécification du modèle

Les tests de spécification présentés permettent de détecter une omission de l'autocorrélation spatiale et la forme prise par cette dernière dans le modèle. Combinés au test du facteur commun, ils fournissent des règles de décision permettant de rechercher la meilleure spécification du modèle.

#### Le test du facteur commun

Le test du facteur commun permet de choisir entre un modèle avec autocorrélation des erreurs et un modèle avec variable endogène décalée et l'ensemble des variables explicatives décalées (Burrige, 1980 ; Bivand, 1984). Le point de départ est la formulation de Durbin : les deux modèles suivants (7) et (8) sont équivalents :

$$(7) y = X\beta + \varepsilon \quad \text{avec} \quad \varepsilon = \lambda W\varepsilon + u$$

$$\Leftrightarrow (8) y = \lambda Wy + X\beta - \lambda WX\beta + u$$

Le modèle (8) est estimé par :

$$(20) y = \lambda Wy + X\beta + WX\delta + u$$

Par conséquent, pour savoir si le modèle (20) peut se réduire au modèle (7), il faut tester l'hypothèse suivante :  $H_0 : \lambda\beta + \delta = 0$ <sup>(12)</sup>. Ce test s'effectue avec l'un des trois tests traditionnels : Wald, rapport de vraisemblance ou multiplicateur de Lagrange. Tous trois convergent vers une loi du chi-deux à  $K - 1$  degrés de liberté (en ignorant le terme constant).

L'encadré 4 résume les principaux tests présentés dans cette partie.

#### Les règles de décision

Si l'on ne dispose pas d'une spécification spatiale *a priori*, différents tests peuvent être combinés pour choisir la meilleure spécification du modèle :

– la première étape consiste à estimer le modèle simple (4) par les MCO et effectuer le test de Moran (16) et le test SARMA. Le rejet de l'hypothèse nulle dans ces deux cas indique une mauvaise spécification du modèle et une

#### Encadré 4 : résumé des différents tests

Modèle 1: $y = X\beta + \varepsilon$	$\varepsilon \sim Nid(0, \sigma^2 I)$
Modèle 2: $y = \rho Wy + X\beta + \varepsilon$	$\varepsilon \sim Nid(0, \sigma^2 I)$
Modèle 3: $y = X\beta + \varepsilon$	$\varepsilon = \lambda W\varepsilon + u \quad u \sim Nid(0, \sigma^2 I)$
Modèle 4: $y = \lambda Wy + X\beta + WX\delta + u$	$u \sim Nid(0, \sigma^2 I)$
Modèle 5: $y = \rho Wy + X\beta + \varepsilon$	$\varepsilon = \lambda W_2\varepsilon + u \quad u \sim Nid(0, \sigma^2 I)$

Test	Hypothèse nulle	Modèle non-contraint	Modèle contraint	Estimation
$LM_{LAG}$	$\rho = 0$	Modèle 2	Modèle 1	MCO
$LM_{ERR}$	$\lambda = 0$	Modèle 3	Modèle 1	MCO
SARMA	$\rho = 0$ et $\lambda = 0$	Modèle 5	Modèle 1	MCO
$LM_{LAG}^*$	$\rho = 0$	Modèle 5	Modèle 3	MV
$LM_{ERR}^*$	$\lambda = 0$	Modèle 5	Modèle 2	MV
FC	$\lambda\beta + \delta = 0$	Modèle 4	Modèle 3	MV

Note : les équivalents robustes de  $LM_{LAG}$  et  $LM_{ERR}$  se calculent à partir des résidus MCO du modèle 1

omission à tort de l'autocorrélation spatiale mais ne permet pas de connaître la forme prise par l'autocorrélation spatiale ;

– si les tests indiquent la présence de dépendance spatiale, il est souvent utile de commencer par inclure dans le modèle, si possible, des variables supplémentaires. Il peut s'agir de variables exogènes supplémentaires qui sont susceptibles d'éliminer la dépendance spatiale ou des variables exogènes décalées spatialement (Florax et Folmer, 1992) ;

– si l'ajout de variables exogènes supplémentaires n'a pas éliminé l'autocorrélation spatiale, il faut alors estimer un modèle incorporant une variable autorégressive ou une autocorrélation des erreurs. Le choix entre ces deux formes de la dépendance spatiale s'effectue en comparant les niveaux de significativité des tests du multiplicateur de Lagrange  $LM_{ERR}$  (18),  $LM_{LAG}$  (19) et leurs versions robustes  $RLM_{ERR}$  et  $RLM_{LAG}$ .

Anselin et Rey (1991), Florax et Folmer (1992) et Anselin et Florax (1995) ont proposé de choisir l'un ou l'autre modèle (5) ou (7) en appliquant la règle de décision suivante. Si  $LM_{LAG}$  est plus significatif que  $LM_{ERR}$  et  $RLM_{LAG}$  est significatif mais pas  $RLM_{ERR}$ , on inclut une variable endogène décalée. Inversement, si  $LM_{ERR}$  est plus significatif que  $LM_{LAG}$  et  $RLM_{ERR}$  est significatif mais pas  $RLM_{LAG}$ , alors on choisit le modèle avec autocorrélation des erreurs ;

– une fois que le modèle spatial adéquat a été estimé ((5) ou (7)), trois tests supplémentaires peuvent encore être mobilisés ;

(a) pour un modèle autorégressif (5), le test  $LM_{ERR}^*$  permet de savoir si une autocorrélation spatiale des erreurs est encore nécessaire ;

(B) pour un modèle avec autocorrélation des erreurs (7), le test  $LM_{LAG}^*$  permet de savoir si une variable endogène décalée est encore nécessaire. Le test du facteur commun indique si la restriction  $\lambda\beta + \delta = 0$  peut être rejetée ou non. Si elle ne l'est pas, le modèle (20) se réduit au modèle avec autocorrélation des erreurs (7) ;

– si plusieurs modèles restent encore en compétition, le choix peut se faire avec les critères traditionnels tels que les critères d'information :

$$(21) INF = -2 \ln L + q(K)$$

$\ln L$  est la valeur de la fonction de log-vraisemblance à l'optimum,  $K$  le nombre de paramètres inconnus et  $q$  un facteur de correction qui varie selon les formulations :  $q = 2K$  pour le critère d'Akaike ou  $q = K \ln N$  pour le critère de Schwartz. Lorsqu'on compare deux modèles selon leur critère d'information, on choisit celui qui minimise ce coefficient.

---

## Conclusion

---

L'objectif de cet article était d'examiner pourquoi l'autocorrélation spatiale doit être modélisée et comment elle est introduite, estimée et testée dans les modèles économétriques. L'autocorrélation spatiale se modélise grâce aux matrices de poids. En sa présence, l'utilisation des moindres carrés ordinaires produit des estimateurs inefficients et l'inférence statistique n'est pas fiable. Lorsqu'elle est détectée, différents modèles économétriques permettent d'en tenir compte : introduction d'une variable endogène décalée et/ou d'une autocorrélation spatiale des erreurs. La caractéristique principale de ces modèles, qui détermine l'ensemble des développements suivants, est la corrélation des erreurs et des variables explicatives, et ce, quelle que soit la forme et la distribution des erreurs. Par conséquent, les moindres carrés ordinaires et les moindres carrés quasi-généralisés ne sont pas des méthodes adaptées et il faut se tourner vers d'autres méthodes d'estimation telle la méthode du maximum de vraisemblance, celle des variables instrumentales ou des moments généralisés. À travers les règles de décision, les tests de spécification permettent enfin de déterminer la forme prise par l'autocorrélation spatiale et son interprétation.

Longtemps ignorées dans les articles contenant des études empiriques, les techniques de l'économétrie spatiale sont aujourd'hui de plus en plus appliquées et permettent d'étudier la présence et l'impact des effets de débordement géographiques sur les comportements économiques. Cet article a été restreint au cadre des modèles linéaires en coupe transversale mais il faut souligner que la recherche en économétrie spatiale s'oriente actuellement vers l'incorporation des effets spatiaux dans les modèles à variables qualitatives et les modèles spatio-temporels (voir McMillen, 1992, 1995 ; Jayet, 1993 ; Anselin, 1988a, 2000 pour plus de détails sur ces extensions).

---

## Notes

---

- (1) Le traitement de l'hétérogénéité spatiale fait l'objet d'une revue dans Le Gallo (2001).
- (2) Les concepts de "proximité" et de "distance", traduisant cette idée, seront pris en compte à travers l'utilisation de matrices de poids.
- (3) La matrice de contiguïté à l'ordre  $k$  n'est pas égale à la matrice de contiguïté à l'ordre 1 élevée à la puissance  $k$ . Cette opération peut produire en effet des "routes circulaires" (des routes qui passent plusieurs fois par une même région) et des "chemins redondants" (des régions qui sont déjà contiguës à l'ordre  $k - 1$  sont encore comptabilisées à l'ordre  $k$ ). Blommestein et Koper (1992) et Anselin et Smirnov (1996) ont donc développé des algorithmes permettant de passer d'une matrice de contiguïté d'ordre 1 à une matrice d'ordre quelconque.
- (4) Ce schéma est également valable lorsque les unités considérées ne sont pas des zones mais des points.
- (5) Des alternatives au processus autorégressif ont été proposées (forme moyenne mobile ou spécification de Kelejian et Robinson, 1993) qui impliquent une interaction plus limitée entre les erreurs. Cependant, leur application est rare et elles ne seront pas traitées ici. On pourra consulter Anselin et Bera (1998) et Le Gallo (2000) pour un exposé détaillé de ces formulations.
- (6) Lorsque  $W_1 = W_2 = W$ , le modèle est suridentifié pour  $\rho$ . De plus, Anselin et Bera (1998) considèrent que ce processus est le résultat d'une matrice de poids mal spécifiée et non d'un processus générateur des données réaliste : si la matrice de poids du modèle autorégressif sous-estime la vraie interaction spatiale dans les données, il y aura une autocorrélation spatiale résiduelle des erreurs. Cela peut mener à estimer un modèle de type (10) alors que seule une matrice de poids bien spécifiée serait nécessaire. Le plus souvent, on cherche à modéliser l'autocorrélation spatiale, soit par une autocorrélation des erreurs, soit par une variable autorégressive et non les deux.
- (7) Ceci est vrai lorsque  $\rho$  est différent de 0 et lorsque l'inverse de  $\rho$  n'est pas une valeur propre de  $W$ .
- (8) En pratique, ces conditions sont largement satisfaites pour les poids basés sur la contiguïté mais pas nécessairement pour les poids généraux.
- (9) La méthode des variables instrumentales et celle des GMM peuvent être combinées pour estimer les paramètres du modèle général (10) contenant à la fois une variable autorégressive et une autocorrélation spatiale des erreurs et obtenir des estimateurs convergents (Kelejian et Prucha, 1998).
- (10) L'inférence statistique peut également être basée sur la distribution exacte de  $I$  en échantillon fini. Hepple (1998) et Tiefelsdorf (2000) ont dérivé un test exact en utilisant les résultats sur les ratios de formes quadratiques de variables normales.
- (11) Dans le cas d'une sur-spécification, des liens spatiaux sont rajoutés à tort (par exemple lorsqu'on utilise une matrice de distance au lieu d'une matrice de contiguïté) alors que dans le cas d'une sous-spécification, des liens spatiaux sont omis à tort.
- (12) Au contraire des séries temporelles, le modèle (7) n'est pas plus facile à estimer qu'un modèle avec une variable endogène décalée puisqu'il faut utiliser dans les deux cas la méthode *MV*. Le seul avantage de la forme (7) sur la forme (20) est qu'il y a moins de paramètres à estimer :  $K + 1$  au lieu de  $2K + 1$ .

---

## Bibliographie

---

- Amemiya T. (1985).** *Advanced Econometrics*, Harvard University Press.
- Anselin L. (1980).** *Estimation Methods for Spatial Autoregressive Structures*, Cornell University, Regional Science Dissertation and Monograph Series #8, Ithaca, NY.
- Anselin L. (1988a).** *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers.
- Anselin L. (1988b).** “Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity”, *Geographical Analysis*, vol. 20, pp. 1-17.
- Anselin L. (1999).** “SpaceStat, a Software Package for the Analysis of Spatial Data”, Version 1.90, Ann Arbor, BioMedware.
- Anselin L. (2000).** “Spatial Econometrics”, in Baltagi B. (ed.), *Companion to Econometrics*, Basil Blackwell.
- Anselin L. et Bera A.K. (1998).** “Spatial Dependence in Linear Regression Models with an Application to Spatial Econometrics”, in Ullah A. et Giles D.E.A. (eds.), *Handbook of Applied Economics Statistics*, Springer-Verlag, pp. 21-74.
- Anselin L., Bera A.K., Florax R. et Yoon M.J. (1996).** “Simple Diagnostic Tests for Spatial Dependence”, *Regional Science and Urban Economics*, vol. 26, pp. 77-104.
- Anselin L. et Florax R. (1995a).** *New Directions in Spatial Econometrics*, Springer.
- Anselin L. et Florax R. (1995b).** “Small Sample properties of Tests for Spatial Dependence in Regression Models”, in Anselin L. et Florax R. (eds.), *New Directions in Spatial Econometrics*, Springer, pp. 21-74.
- Anselin L. et Hudak S. (1992).** “Spatial Econometrics in Practice, a Review of Software Options”, *Regional Science and Urban Economics*, vol. 22, pp. 509-536.
- Anselin L. et Kelejian H. (1997).** “Testing for Spatial Error Autocorrelation in the Presence of Endogenous Regressors”, *International Regional Science Review*, vol. 20, pp. 153-182.
- Anselin L. et Rey S. (1991).** “Properties of Tests for Spatial Dependence in Linear Regression Models”, *Geographical Analysis*, vol. 23, pp. 112-131.
- Anselin L. et Smirnov O. (1996).** “Efficient Algorithms for Constructing Proper Higher Order Spatial Lag Operators”, *Journal of Regional Science*, vol. 36, pp. 67-89.
- Anselin L., Varga A. et Acs Z. (1997).** “Local Geographic Spillovers between University Research and High Technology Innovations”, *Journal of Urban Economics*, vol. 42, pp. 422-448.
- Arbia G. (1989).** *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems*, Kluwer Academic Publishers.
- Aten B. (1997).** “Does Space Matter? International Comparisons of the Prices of Tradables and Nontradables”, *International Regional Science Review*, vol. 20, pp. 35-52.
- Baumont C., Ertur C. et Le Gallo J. (2001).** “A Spatial Econometric Analysis of Geographic Spillovers and Growth for European Regions, 1980-1995”, *Document de travail*, n° 2001-04, LATEC, Université de Bourgogne.
- Benirschka A. et Binkley J. (1996).** “Land Price Volatility in a Geographically Dispersed Market”, *American Journal of Agricultural Economics*, vol. 76, pp. 185-195.
- Bera A.K. et Yoon M. (1993).** “Specification Testing with Locally Misspecified Alternatives”, *Econometric Theory*, vol. 9, pp. 649-658.
- Bivand R.S. (1984).** “Regression Modeling with Spatial Dependence: an Application of Some Class Selection and Estimation Methods”, *Geographical Analysis*, vol. 16, pp. 25-37.
- Blommestein H.J. et Koper N.A.M. (1992).** “Recursive Algorithms for the Elimination of Redundant Paths in Spatial Lag Operators”, *Journal of Regional Science*, vol. 32, pp. 91-111.
- Boarnet M.G. (1998).** “Spillovers and the Locational Effects of Public Infrastructure”, *Journal of Regional Science*, vol. 38, pp. 381-400.
- Brandsma A.S. et Kelletaper R.H. (1979).** “A Biparametric Approach to Spatial Autocorrelation”, *Environment and Planning A*, vol.11, pp. 51-58.
- Brueckner J.K. (1998).** “Testing for Strategic Interaction Among Local Governments: the Case of Growth Controls”, *Journal of Urban Economics*, vol. 44, pp. 438-467.
- Burridge P. (1981).** “Testing for a Common Factor in a Spatial Autoregressive Model”, *Environment and Planning A*, vol. 13, pp. 795-800.
- Can A. (1992).** “Specification and Estimation of Hedonic Housing Price Models”, *Regional Science and Urban Economics*, vol. 22, pp. 453-474.

- Can A. et Megboluge I. (1997).** “Spatial Dependence and House Price Index Construction”, *Journal of Real Estate Finance and Economics*, vol. 14, pp. 203-222.
- Case A. (1991).** “Spatial Patterns in Household Demand”, *Econometrica*, vol. 59, pp. 953-965.
- Case A.C., Rosen H.S. et Hines J.R. (1993).** “Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States”, *Journal of Public Economics*, vol. 52, pp. 285-307.
- Chou Y.H. (1991).** “Map Resolution and Spatial Autocorrelation”, *Geographical Analysis*, vol. 23, pp. 228-228-246.
- Cliff A.D. et Ord J.K. (1973).** *Spatial Autocorrelation*, Pion.
- Cliff A.D. et Ord J.K. (1981).** *Spatial Processes: Models and Applications*, Pion.
- Cressie N. (1993).** *Statistics for Spatial Data*, John Wiley.
- Dubin R.A. (1998).** “Spatial Autocorrelation: a Primer”, *Journal of Housing Economics*, vol. 7, pp. 304-327.
- Dubin R., Pace R.K. et Thibodeau T.G. (1999).** “Spatial Autoregression Techniques for Real Estate Data”, *Journal of Real Estate Techniques*, vol. 7, pp. 79-95.
- Fingleton B. (1999).** “Estimates of Time to Convergence: An Analysis of Regions of European Union”, *International Regional Science Review*, vol. 22, pp. 5-34.
- Florax R. et Folmer H. (1992).** “Specification and Estimation of Spatial Linear Regression Models”, *Regional Science and Urban Economics*, vol. 22, pp. 405-432.
- Florax R. et Rey S. (1995).** “The Impacts of Misspecified Spatial Interaction in Linear Regression Models”, in Anselin L et Florax R. (eds.), *New Directions in Spatial Econometrics*, Springer, pp. 111-135.
- Griffith D.A. (1992).** “What is Spatial Autocorrelation? Reflections on the Past 25 Years of Spatial Statistics”, *L’Espace Géographique*, vol. 21, pp. 265-280.
- Griffith D.A. (1993).** *Spatial Regression on the PC: Spatial Statistics Using SAS*, Association of the American Geographers, Washington, DC.
- Heijmans R.D.H. et Magnus J.R. (1986a).** “Asymptotic Normality of Maximum Likelihood Estimators Obtained from Normally Distributed but Dependent Observations”, *Econometric Theory*, vol. 12, pp. 374-412.
- Heijmans R.D.H. et Magnus J.R. (1986b).** “Consistent Maximum-likelihood Estimation with Dependent Observations: the General (non-normal) Case and the Normal Case”, *Journal of Econometrics*, vol. 32, pp. 253-285.
- Hepple L. (1998).** “Exact Testing for Spatial Autocorrelation Among Regression Residuals”, *Environment and Planning A.*, vol. 30, pp. 85-108.
- Huang J.S. (1984).** “The Autoregressive Moving Average Model for Spatial Analysis”, *Australian Journal of Statistics*, vol. 26, pp. 169-178.
- Jayet H. (1993).** *Analyse spatiale quantitative*, Economica.
- Kelejian H.H. et Prucha I.R. (1997).** “Estimation of Spatial Regression Models with Autoregressive Errors by Two-stage Least Squares Procedures: a Serious Problem”, *International Regional Science Review*, vol. 20, pp. 103-111.
- Kelejian H.H. et Prucha I.R. (1998).** “A Generalized Spatial Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances”, *Journal of Real Estate Finance and Economics*, vol. 17, pp. 99-121.
- Kelejian H.H. et Prucha I.R. (1999).** “A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model”, *International Economic Review*, vol. 40, pp. 509-534.
- Kelejian H.H. et Robinson D.P. (1992).** “Spatial Autocorrelation: a New Computationally Simple Test with an Application to per Capita Country Police Expenditures”, *Regional Science and Urban Economics*, vol. 22, pp. 317-331.
- Kelejian H.H. et Robinson D.P. (1993).** “A Suggested Method of Estimation for Spatial Interdependent Models with Autocorrelated Errors, and an Application to a Country Expenditure Country Police Expenditure”, *Papers in Regional Science*, vol. 72, pp. 297-312.
- Kelejian H.H. et Robinson D.P. (1997).** “Infrastructure Productivity Estimation and its Underlying Econometric Specifications: a Sensitivity Analysis”, *Papers in Regional Science*, vol. 76, pp. 115-131.
- King M.L. (1981).** “A Small Sample Property of the Cliff-Ord Test for Spatial Correlation”, *Journal of the Royal Statistical Society B*, vol. 43, pp. 263-264.
- Le Gallo J. (2000).** “Économétrie spatiale, 1. Autocorrélation spatiale”, *document de travail*, n°2000-05, LATEC, Université de Bourgogne.
- Le Gallo J. (2001).** “Économétrie spatiale, 2. Hétérogénéité spatiale”, *document de travail*, n°2001-01, LATEC, Université de Bourgogne.

- Li B. (1996).** “Implementing Spatial Statistics on Parallel Computers”, in S. Arlinghaus (ed.), *Practical handbook of Spatial Statistics*, CRC Press, pp. 107-148.
- Macedo P.B.R. (1998).** “Hedonic Price Models with Spatial Effects: an Application to the Housing Market of Belo Horizonte, Brazil”, *Revista Brasileira de Economia*, vol. 52, pp. 63-81.
- Magnus, J. (1978).** “Maximum Likelihood Estimation of the GLS Model with Unknown Parameters in the Disturbance Covariance Matrix”, *Journal of Econometrics*, vol. 7, pp. 281-312 (Corrigenda, *Journal of econometrics*, vol. 10, pp. 261).
- Mathsoft (1996).** *S+SpatialStats. User’s Manual for Windows and Unix*, Mathsoft, Inc, Seattle, WA.
- McMillen D.P. (1992).** “Probit with Spatial Autocorrelation”, *Journal of Regional Science*, vol. 32, pp. 335-348.
- McMillen D.P. (1995).** “Selection Bias in Spatial Econometric Models”, *Journal of Regional Science*, vol. 35, pp. 417-436.
- Moran P. (1948).** “The Interpretation of Statistical Maps”, *Journal of the Royal Statistical Society*, vol. 10B, pp. 243-251.
- Moran P. (1950a).** “Notes on Continuous Stochastic Phenomena”, *Biometrika*, vol. 37, pp. 17-23.
- Moran P. (1950b).** “A Test for Serial Independence of Residuals”, *Biometrika*, vol. 37, pp. 178-181.
- Moreno R. et Trehan B. (1997).** “Location and the Growth of Nations”, *Journal of Economic Growth*, vol. 2, pp. 399-418.
- Moreno R., Artís M., López-Bazo J. et Suriñach J. (1998).** “Evidence on the Complex Link between Infrastructure and Regional Growth”, *International Journal of Development Planning Literature*, vol. 12, pp. 81-108.
- Ord K. (1975).** “Estimation Methods for models of Spatial Interaction”, *Journal of the American Statistical Association*, vol. 70, pp. 120-126.
- Pace R. K. et Gilley O.W. (1997).** “Using the Spatial Configuration of the Data to Improve Estimation”, *Journal of Real Estate Finance and Economics*, vol. 14, pp. 333-340.
- Paelinck J.H.P. et Klaassen L.H. (1979).** *Spatial Econometrics*, Saxon House.
- Pinkse J. (1998).** “Asymptotic Properties of the Moran and Related Tests and a Test for Spatial Correlation in Probit Models”, *Document de travail*, Department of Economics, University of British Columbia, Vancouver, BC.
- Rey S.J. et Montouri B.D. (1999).** “U.S. Regional Income Convergence: a Spatial Econometric Perspective”, *Regional Studies*, vol. 33, pp. 145-156.
- Student (1914).** “The Elimination of Spurious Correlation due to Position in Time or Space”, *Biometrika*, vol. 5, pp. 351-360.
- Tiefelsdorf M. (2000).** *Modelling Spatial Processes*, Lecture Notes in Earth Sciences, n° 87, Springer.
- Tobler W. (1979).** “Cellular geography”, in Gale S. et Olsson G. (eds), *Philosophy in geography*, Reidel, pp. 379-386.
- Upton G.J.G. et Fingleton B. (1985).** *Spatial Data Analysis by Example*, John Wiley.
- Varga A. (1998).** *University Research and Regional Innovation: a Spatial Econometric Analysis of Academic Technology Transfers*, Kluwer Academic Publishers.