



HAL
open science

Econométrie spatiale (1, Autocorrélation spatiale)

Julie Le Gallo

► **To cite this version:**

Julie Le Gallo. Econométrie spatiale (1, Autocorrélation spatiale). [Rapport de recherche] Laboratoire d'analyse et de techniques économiques(LATEC). 2000, 45 p., Table, ref. bib. : 5 p. hal-01527290

HAL Id: hal-01527290

<https://hal.science/hal-01527290>

Submitted on 24 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LATEC

LABORATOIRE D'ANALYSE ET DE TECHNIQUES ÉCONOMIQUES

UMR 5118 CNRS

DOCUMENT DE TRAVAIL



Pôle d'Économie et de Gestion

2, bd Gabriel - BP 26611 - F-21066 Dijon cedex - Tél. 03 80 39 54 30 - Fax 03 80 39 54 43

Courrier électronique : secretariat.latec@u-bourgogne.fr

ISSN : 1260-8556

n° 2000-05

Econométrie spatiale
1. Autocorrélation spatiale

Julie Le Gallo

juin 2000

LATEC, Université de Bourgogne
Pôle d'Economie et de Gestion, BP 26611 - 21066 Dijon cedex
e-mail : JleGallo@aol.com
<http://www.u-bourgogne.fr/LATEC>

ECONOMETRIE SPATIALE

1. Autocorrélation spatiale

Julie Le Gallo*

Résumé

Les méthodes de l'économétrie spatiale visent à traiter les deux grandes particularités des données spatiales : l'autocorrélation spatiale qui se réfère à l'absence d'indépendance entre observations géographiques et l'hétérogénéité spatiale qui est liée à la différenciation dans l'espace des variables et des comportements. Ces techniques ont connu de nombreux développements depuis une dizaine d'années et sont de plus en plus appliquées dans les études empiriques nécessitant l'utilisation de données géographiques.

L'objectif de cet article est de présenter les diverses façons permettant de modéliser l'autocorrélation et l'hétérogénéité spatiales ainsi que les procédures d'estimation et d'inférence adaptées aux modèles incorporant ces deux effets.

L'article est divisé en deux parties. Cette première partie est consacrée au problème de l'autocorrélation spatiale (document de travail n° 2000-05) alors que la seconde portera sur le problème de l'hétérogénéité spatiale.

Mots-clés : économétrie spatiale, autocorrélation spatiale, hétérogénéité spatiale

Abstract

Spatial econometric methods aim at taking into account the two special characteristics of spatial data: spatial autocorrelation, which is the lack of independence between geographical observations, and spatial heterogeneity, which is related to the differentiation of variables and behaviors in space. These techniques have been mostly developed the last ten years and are more often applied in empirical studies with geographical data.

The aim of this article is to present the way spatial autocorrelation and spatial heterogeneity can be incorporated in regression relationships and to present the estimation and inference procedures adapted to the models incorporating these two effects.

This article is divided in two parts. This first part deals with spatial autocorrelation (working paper n° 2000-05) and the second part will deal with spatial heterogeneity.

Key-words : spatial econometrics, spatial autocorrelation, spatial heterogeneity

Classification *JEL* : C51, C52, R15

* L'auteur remercie C. Baumont, M.-C. Pichery et L. Bertinelli pour leurs commentaires et suggestions. L'auteur reste seule responsable des insuffisances que pourrait comporter ce texte.

ECONOMETRIE SPATIALE

INTRODUCTION GENERALE

Depuis une vingtaine d'années, les méthodologies en géographie et en science régionale ont connu de nombreux développements destinés à traiter les particularités des données géographiques, c'est-à-dire les observations d'une variable mesurée pour des localisations particulières réparties dans l'espace. En effet, l'introduction de l'espace dans les modèles économétriques n'est ni neutre, ni immédiate, et les techniques de l'économétrie spatiale visent précisément à prendre en compte les deux grandes spécificités des données géographiques : *l'autocorrélation spatiale* qui se réfère à l'absence d'indépendance entre observations géographiques et *l'hétérogénéité spatiale* qui est liée à la différenciation des variables et des comportements dans l'espace.

C'est à Cliff et Ord qu'on doit, après une série d'articles à la fin des années 60 et au début des années 70, un ouvrage présentant de manière synthétique l'état des savoirs en statistique et en économétrie spatiales (1973). Après cette phase initiale de reconnaissance, on assiste à la fin des années 70 et au début des années 80 au raffinement du cadre original d'analyse de Cliff et Ord et plus particulièrement au développement de la théorie de l'estimation et des tests (Ord, 1975 ; Haining, 1978 ; Anselin, 1988a). Un certain nombre d'ouvrages rendent compte de ces développements : Cliff et Ord (1981), Upton et Fingleton (1985), Griffith (1988a), Haining (1990), Cressie (1993), Jayet (1993), Bailey et Gatrell (1995), Fotheringham, Brundson et Charlton (2000).

L'objectif de cet article est de présenter les diverses façons permettant de modéliser l'autocorrélation et l'hétérogénéité spatiales ainsi que les procédures d'estimation et d'inférence adaptées aux modèles incorporant ces deux effets.

L'article est divisé en deux parties. Cette première partie est consacrée au problème de l'autocorrélation spatiale alors que la seconde portera sur le problème de l'hétérogénéité spatiale.

Data of geographic units are tied together, like bunches of grapes, not separate, like balls in an urn. Of course, mere contiguity in time and space does not of itself indicate lack of independence between units in a relevant variable or attribute, but in dealing with social data, we know that by virtue of their very social character, persons, groups and their characteristics are interrelated and not independent. Sampling error formulas may yet be developed which are applicable to these data, but until then the older formulas must be used with great caution. Likewise, other statistical measures must be carefully scrutinized when applied to these data... (Stephan, 1934)

ECONOMETRIE SPATIALE

1. Autocorrélation spatiale

INTRODUCTION

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 exposées dans les revues spécialisées et appliquées à des problèmes d'économie régionale, spatiale ou urbaine. Par exemple, Griffith (1981), Anselin et Can (1986), Griffith et Can (1996), ont abordé le traitement de l'autocorrélation spatiale dans l'étude des fonctions de densité urbaines. L'intérêt de sa prise en compte dans les modèles hédoniques de prix immobiliers a été présenté par Can (1990, 1992), Can et Megboluge (1997), Pace et Gilley (1997) alors que Florax (1992), Anselin, Varga et Acs (1997, 1998) et Varga (1998) ont étudié les externalités spatiales d'information dues à la recherche universitaire et la R&D.

Plus récemment cependant, les méthodes de l'économétrie spatiale ont été appliquées à d'autres sujets tels que l'analyse de la demande (Case, 1991), l'économie internationale (Aten, 1996, 1997), l'économie publique (Case et al., 1993 ; Brueckner, 1998), l'économie rurale (Benirschka et Binkley, 1994), ou encore les phénomènes de croissance et de convergence (Rey and Montouri 1999 ; Fingleton 1999 ; Baumont et al., 2000a, 2000b). Ces méthodes sont en effet potentiellement applicables à toutes les études empiriques nécessitant l'utilisation de données spatiales.

Deux raisons principales peuvent être attribuées au regain d'attention porté à la prise en compte de l'autocorrélation spatiale. La première, d'ordre théorique, est le développement de nouveaux courants prenant en compte les interactions économiques. Ainsi la Nouvelle Economie Géographique met l'accent sur les externalités spatiales, les économies d'agglomération ou toutes autres formes d'effets de débordement. Les techniques de l'économétrie spatiale peuvent également tenir compte d'autres types d'interactions entre les agents telles que les normes sociales ou les effets de voisinage. La deuxième raison, d'ordre empirique, est la disponibilité croissante des données spatialisées et le fort développement actuel des logiciels de Systèmes d'Information Géographiques (SIG).

Plusieurs aspects sont abordés dans cet article. L'autocorrélation spatiale est formalisée à partir de la notion de matrice de poids et donne lieu à des coefficients d'autocorrélation détectant la présence de dépendance spatiale dans une série statistique (section 1). Sa modélisation économétrique s'effectue grâce à diverses formes fonctionnelles (section 2) qui nécessitent des méthodes d'estimation et d'inférence adaptées. Les méthodes d'estimation les plus utilisées sont la méthode du maximum de vraisemblance, la méthode des variables instrumentales ou celle des moments généralisés (section 3). Les méthodes d'inférence, en général basées sur le principe du maximum de vraisemblance, servent à tester la présence de l'autocorrélation spatiale dans les modèles et à déterminer la forme de cette dernière. Une illustration de ces techniques est fournie à travers l'exemple de la convergence entre les régions européennes (section 4).

1. L'AUTOCORRELATION SPATIALE : DE QUOI S'AGIT-IL ?

Dès 1914, Student suspectait la présence d'une relation entre différentes observations géographiques, une idée qui entraîne l'abandon de l'hypothèse statistique fondamentale d'observations indépendantes. Ce phénomène, appelé autocorrélation spatiale, peut être relié à différentes sources (paragraphe 11). La spécification de l'autocorrélation spatiale nécessite des outils destinés à modéliser l'interdépendance entre les régions, à savoir la matrice de poids et les opérateurs spatiaux (paragraphe 12), ces outils permettant alors de tester la présence d'autocorrélation spatiale dans une série spatiale univariée, à l'aide de la statistique globale de Moran (paragraphe 13).

11. Définition et sources de l'autocorrélation spatiale

L'autocorrélation spatiale a pour point de départ la constatation selon laquelle les observations spatialisées en coupes transversales ne sont pas indépendantes. On définit l'autocorrélation spatiale comme la corrélation, positive ou négative, d'une variable avec elle-même provenant de la disposition géographique des données. Sur une carte :

1/ *une autocorrélation spatiale positive* se traduit par le regroupement géographique d'observations de classe voisine : des lieux proches se ressemblent davantage que les lieux éloignés.

2/ *une autocorrélation spatiale négative* se traduit par le regroupement géographique d'observations dissemblables : des lieux proches sont plus différents que des lieux éloignés.

3/ *une absence d'autocorrélation spatiale* indique que la répartition spatiale des observations est aléatoire : aucune relation n'existe entre la proximité des lieux et leur degré de ressemblance.

Détecter de l'autocorrélation spatiale dans une série spatiale donne alors une information supplémentaire par rapport aux statistiques traditionnelles (telles que la moyenne ou l'écart-type) sur la façon dont les différentes valeurs sont disposées géographiquement et permet de décrire la nature et le degré de l'interdépendance spatiale de la structure (Griffith, 1992b).

Lorsqu'il y a autocorrélation spatiale pour une variable, cela 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* »

Les concepts de « proximité » et de « distance » traduisant cette idée seront pris en compte à travers l'utilisation de matrices de poids. L'autocorrélation spatiale diffère donc de l'autocorrélation temporelle. Cette dernière est en effet 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 en particulier pour conséquence de complexifier 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 section 3).

L'autocorrélation spatiale a deux sources principales.

1/ L'autocorrélation spatiale 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 particulière des activités dans l'espace.

Par exemple, la diffusion d'un phénomène (comme la diffusion technologique) à partir d'un ou de plusieurs lieux d'origine implique que la fréquence ou l'intensité de la mesure de ce phénomène dépend de la distance à l'origine. Aux localisations proches les unes des autres, et donc à des distances comparables de l'origine, seront donc associées des fréquences similaires pour le phénomène étudié. Les processus d'interactions peuvent également être à la source de l'autocorrélation spatiale : 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.

2/ L'autocorrélation spatiale peut également provenir d'une mauvaise spécification du modèle, comme des variables omises spatialement autocorrélées d'une forme fonctionnelle incorrecte, de données manquantes ou d'erreurs de mesure. Elle est alors considérée comme un outil de diagnostic et de détection d'une mauvaise spécification du modèle.

Finalement, l'autocorrélation spatiale mesure donc le degré auquel un attribut en une localisation est similaire aux attributs des localisations voisines. Il est à noter que les localisations dans l'espace se divisent en trois catégories. Il peut s'agir tout d'abord de points représentant par exemple des localisations de magasins, d'aires urbaines... 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 *fini*. Cette caractéristique permet de distinguer entre les techniques de l'économétrie spatiale et celles de la géostatistique (Cressie, 1993 ; Griffith et Layne, 1999). L'économétrie spatiale est principalement utilisée lorsqu'on est en présence d'un *ensemble fini* (régulier ou non) de points ou de zones reliés entre eux par des relations de voisinage. La géostatistique concerne essentiellement les *données issues d'un processus spatial sous-jacent continu* sur l'espace étudié et est donc d'un intérêt moindre pour l'étude de données socio-économiques. Cet article porte uniquement sur les techniques de l'économétrie spatiale.

12. Les matrices de poids et les opérateurs spatiaux

Puisque les unités spatiales sont en général interdépendantes, il faut considérer les positions relatives des observations les unes par rapport aux autres en plus de leurs dimensions et de leurs structures. Pour cela, on doit spécifier la topologie du système spatial en construisant une matrice de poids. Ces matrices sont *exogènes*, elles sont définies a priori par le modélisateur compte tenu de sa connaissance sur les relations et interactions entre les

unités spatiales. Elles ne contiennent donc pas d'éléments à estimer. Les matrices de poids se classent en deux grandes catégories : les matrices de contiguïté (paragraphe 121) et les matrices de poids généralisées (paragraphe 122). Les matrices de poids permettent alors de définir la notion d'opérateur spatial ou de variable spatiale décalée, notion principale utilisée pour la modélisation économétrique (paragraphe 123).

121. Les matrices de contiguïté

Définissons la contiguïté entre deux régions par le fait qu'elles ont une frontière commune. Une matrice de contiguïté d'ordre 1 est une matrice carrée, symétrique, ayant autant de lignes et de colonnes qu'il y a de zones géographiques (N) et où chaque terme w_{ij} est défini comme suit :

$$w_{ij} = \begin{cases} 1 & \text{si les régions } i \text{ et } j \text{ sont contiguës à l'ordre 1} \\ 0 & \text{sinon} \end{cases} \quad [1.1]$$

Par convention, une région n'est pas contiguë avec elle-même : $w_{ii} = 0 \quad \forall i$. Si l'on désire connaître le nombre de régions contiguës à une région i , il suffit de calculer la somme des éléments de la ligne i de la matrice de contiguïté soit :

$$L_i = \sum_{j=1}^N w_{ij} \quad [1.2]$$

Le nombre total de liens existant dans le système régional est alors égal à :

$$A = \frac{1}{2} \sum_{j=1}^N L_j \quad [1.3]$$

Cette notion de contiguïté peut être généralisée : on dira que 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 . Il est à noter que 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 produirait 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 comptabilisés à l'ordre k). Blommestein (1985), Blommestein et Koper (1992) et Anselin et Smirnov (1996) ont développé des algorithmes permettant de passer d'une matrice de contiguïté à l'ordre 1 à une matrice de contiguïté d'ordre quelconque.

122. Les matrices de poids générales

Dans les matrices de poids générales, chaque élément représente l'intensité de l'interaction entre les deux régions, intensité qui n'est plus forcément liée à la contiguïté. Dans ce cas, les matrices de poids ne sont pas nécessairement symétriques, ce qui est davantage approprié lorsque les relations étudiées sont des relations de diffusion ou des relations de type centre - périphérie. Les propriétés de certaines matrices de poids ont été systématiquement étudiées par Bavaud (1998).

Une première possibilité consiste à utiliser des matrices de distance. On suppose dans ce cas que l'intensité de l'interaction entre 2 régions i et j dépend de la distance entre les centroïdes de ces régions ou entre les capitales de ces régions¹. Plusieurs indicateurs peuvent être utilisés selon la définition de la distance : distance à vol d'oiseau, distance par routes... On peut bien sûr généraliser aux temps de transport ou à des indices d'accessibilité plus généraux. Diverses formes fonctionnelles sont également disponibles, les plus utilisées étant la fonction exponentielle inverse [1.4] ou une fonction de l'inverse de la distance [1.5]. Si d_{ij} désigne la distance entre la région i et la région j , les éléments de la matrice de distance pour ces deux différents cas sont définis par :

$$w_{ij} = e^{-\alpha d_{ij}} \quad [1.4]$$

$$w_{ij} = \begin{cases} 1 / d_{ij}^\beta & \text{si } d_{ij} < \bar{d} \\ 0 & \text{sinon} \end{cases} \quad [1.5]$$

α et β sont des paramètres déterminés a priori, \bar{d} est la valeur seuil au-delà de laquelle on suppose qu'il n'y a pas d'interaction directe entre la région i et la région j .

Pour affiner cette spécification, Cliff et Ord (1981) 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. Formellement, un élément de la matrice de poids s'écrit :

$$w_{ij} = (d_{ij})^{-a} (\beta_{ij})^b \quad [1.6]$$

β_{ij} est la proportion de la frontière intérieure de l'unité i qui est en contact avec l'unité j , a et b sont des paramètres déterminés a priori.

Enfin, certains auteurs ont utilisé des matrices de poids dans lesquelles les éléments ont perdu toute référence à la localisation géographique des données. Par exemple, Aten (1996, 1997) utilise une matrice de flux commerciaux pour étudier l'autocorrélation spatiale dans les prix internationaux. En sociométrie, les poids peuvent refléter l'appartenance de deux individus à la même catégorie sociale. Case et al. (1993) proposent ainsi d'utiliser des poids de la forme suivante : $w_{ij} = 1/|x_i - x_j|$ où x_i et x_j sont des observations sur des caractéristiques socioéconomiques pertinentes, telles que le revenu par tête ou le pourcentage de la population dans un groupe racial ou ethnique.

Les matrices de contiguïté ou les matrices de poids généralisées sont souvent standardisées de telle sorte que la somme de chaque ligne est égale à 1. Les éléments d'une matrice standardisée sont égaux à :

$$w_{ij}^s = \frac{w_{ij}}{\sum_j w_{ij}} \quad [1.7]$$

Les poids sont alors compris entre 0 et 1 et cette opération rend les paramètres spatiaux dans les processus spatiaux comparables entre les modèles (voir paragraphe 32). Pour les matrices

¹ Ce schéma est également valable lorsque les unités considérées ne sont pas des zones mais des points.

de contiguïté, cette opération facilite également l'interprétation des éléments du décalage spatial.

123. Opérateurs spatiaux ou variables spatiales décalées

L'utilisation des matrices de poids conduit à la possibilité de relier une observation de la variable en un point de l'espace avec les observations de cette variable pour d'autres lieux du système. Formellement, pour N régions, l'opérateur spatial associé avec la matrice de poids W est défini par le vecteur $(N,1) : Wy$. Wy est aussi appelé « variable spatiale décalée » associée à y . Le $i^{\text{ème}}$ élément de cette variable décalée est :

$$[Wy]_i = \sum_{j=1}^N w_{ij}y_j \quad [1.8]$$

Lorsque la matrice de poids 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 .

Le concept d'opérateur spatial qui en découle est particulièrement important car il permet d'introduire l'autocorrélation spatiale dans les modèles économétriques : l'autocorrélation spatiale est modélisée comme une relation fonctionnelle entre une variable y ou un terme d'erreur ε , et son décalage spatial associé, Wy pour une variable endogène décalée ou $W\varepsilon$ pour une erreur spatialement décalée. Les spécifications qui en résultent sont étudiées dans la section 2.

13. Le test de l'autocorrélation spatiale d'une variable quantitative

Pour mesurer l'autocorrélation spatiale présente dans des variables quantitatives, la statistique globale la plus utilisée est celle de Moran (1948). Elle s'écrit formellement de la façon suivante :

$$I = \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S_0} \bigg/ \frac{\sum_i (x_i - \bar{x})^2}{N} \quad [1.9]$$

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

Le numérateur s'interprète comme la covariance entre unités contiguës, chaque contiguïté étant pondérée par w_{ij}/S_0 . Elle est normalisée par le dénominateur qui est la variance totale observée.

On pourra trouver les expressions analytiques de l'espérance et de la variance de cette statistique sous diverses hypothèses (dont celle de normalité) dans Cliff et Ord (1973, 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.

L'utilisation de ce coefficient dans la recherche de l'autocorrélation spatiale pose parfois des problèmes rendant difficile l'interprétation des résultats.

1/ La mesure de l'autocorrélation spatiale est dépendante à la matrice de poids utilisée : accepter 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.

2/ Un autre problème est relié à la façon dont les données spatiales sont agrégées et qui peut avoir un effet sur la mesure de l'autocorrélation spatiale. Il s'agit du « MAUP » ou « Modifiable Areal Unit Problem » (Arbia, 1989) qui recouvre deux problèmes potentiels. Premièrement, l'autocorrélation spatiale peut être affectée par le niveau d'agrégation utilisé, c'est l'effet d'échelle (Chou, 1991). Par exemple, les résultats peuvent changer selon qu'on utilise les données sur les régions européennes au niveau NUTS1 ou au niveau NUTS2. Deuxièmement, il y a beaucoup de façons de découper une région en plusieurs subdivisions, ce qui donne lieu à de nombreuses configurations spatiales. L'autocorrélation spatiale est aussi sensible à ce problème de forme des unités spatiales (voir Griffith, 1992b pour un exemple).

2. AUTOCORRELATION SPATIALE ET MODELES ECONOMETRIQUES

Dans les années 1970, Paelinck et Klaassen formulaient cinq principes à respecter dans la formulation de modèles économétriques spatiaux (voir par exemple Paelinck et Klaassen, 1979 ; Ancot, Kuiper et Paelinck, 1983) :

- *Le principe d'interdépendance spatiale* : « a priori les modèles spatiaux sont à formuler de façon spatialement interdépendante. Qu'il s'agisse de phénomènes de consommation, de production, d'investissement, les influences mutuelles sont innombrables (...) ».

- *Le principe d'asymétrie des relations spatiales* : « Un certain nombre de relations spatiales interdépendantes sont vraisemblablement asymétriques. En théorie spatiale, l'on a découvert depuis longtemps qu'une matrice des distances n'est pas forcément symétrique (...) ».

- *Le principe d'allotopie* : « Des phénomènes économiques localisés dans un espace donné doivent souvent être expliqués par des facteurs causaux localisés dans d'autres espaces. Un exemple classique est le triangle de Weber, fournissant parfois une solution de localisation optimale intérieure, donc différente des points où se trouvent les marchés de produits et de facteurs ; de même les phénomènes migratoires s'expliquent par la comparaison d'avantages et d'inconvénients localisés en des endroits différents de l'espace considéré ».

- *Le principe géographique* : « l'espace bidimensionnel, comme contenant possible des activités économiques, doit apparaître explicitement dans les modèles ».

- *Le principe de distinction entre interaction ex-ante et interaction ex-post* : les phénomènes d'installation et les phénomènes de fonctionnement doivent être distingués.

Ces différents principes guident la façon dont les modèles économétriques spatiaux doivent être spécifiés. Le principe d'asymétrie et le principe géographique conduisent à l'utilisation de matrices de poids explicitant la forme et l'intensité de l'interaction spatiale entre les régions. Le principe de l'interdépendance spatiale et celui de l'allotopie impliquent l'emploi de variables spatiales décalées, endogènes ou exogènes (paragraphe 21) ou d'une autocorrélation spatiale des erreurs (paragraphe 22). Des processus plus complexes dans lesquels ces deux aspects sont combinés sont également envisageables (paragraphe 23) ainsi que l'introduction de l'autocorrélation spatiale dans les modèles de panel et les modèles de choix discrets (paragraphe 24).

En coupes transversales, nous prendrons comme point de départ le modèle de régression linéaire classique suivant :

$$y = X\beta + \varepsilon \quad [2.1]$$

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, β est le vecteur $(K,1)$ des coefficients de régression inconnus, ε est le vecteur $(N,1)$ des erreurs.

Sauf indication contraire, on supposera vérifiées les deux hypothèses suivantes :

H_1 : Les erreurs ε_i sont identiquement et indépendamment distribuées (iid), d'espérance nulle et de variance finie σ^2 .

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

21. Variable(s) décalée(s) : endogène(s) et exogène(s)

211. Le modèle autorégressif

La première façon de prendre en compte l'autocorrélation spatiale est le modèle autorégressif spatial : une « variable endogène décalée » est incluse dans le modèle [2.1] :

$$y = \rho W y + X\beta + \varepsilon \quad [2.2]$$

Wy est la variable endogène décalée pour la matrice de poids W , ρ est le paramètre spatial autorégressif indiquant l'ampleur de l'interaction existant entre les observations de y . Dans ce modèle, l'observation y_i est ainsi en partie expliquée par les valeurs prises par y dans les régions voisines : $(Wy)_i = \sum_{j \neq i} w_{ij} y_j$. On rappelle que si la matrice W est standardisée, cette valeur s'interprète comme la moyenne des valeurs de y sur les observations voisines à i . De plus, cette standardisation produit une estimation du paramètre ρ inférieure à 1 en valeur absolue (voir la paragraphe 32) ce qui facilite la comparaison de l'ampleur de l'autocorrélation spatiale lorsque le modèle [2.2] est estimé pour plusieurs matrices de poids.

Le modèle [2.2] se réécrit de la façon suivante :

$$(I - \rho W)y = X\beta + \varepsilon \quad [2.3]$$

Dans ce modèle : $y^* = (I - \rho W)y$ s'interprète comme la variable dépendante « filtrée », dans laquelle les effets de l'autocorrélation spatiale ont été éliminés. Supposons la matrice $(I - \rho W)$ non-singulière². Dans ce cas, [2.3] se réécrit sous la forme suivante :

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \quad [2.4]$$

La matrice des variances-covariances de y s'écrit donc :

$$V(y) = (I - \rho W)^{-1} E(\varepsilon\varepsilon')(I - \rho W')^{-1} \quad [2.5a]$$

$$V(y) = \sigma^2 [(I - \rho W')(I - \rho W)]^{-1} \quad [2.5b]$$

² Ceci est vrai lorsque ρ est différent de 0 et lorsque l'inverse de ρ n'est pas une valeur propre de W .

Cette matrice des variances-covariances est une matrice pleine, ce qui implique que chaque localisation est corrélée avec chaque autre localisation mais cette corrélation diminue avec le degré de proximité.

Lorsqu'une variable endogène décalée est ignorée dans la spécification du modèle, alors qu'elle est présente dans le processus générateur des données, les estimateurs des MCO dans le modèle spatial [2.1] seront biaisés et non convergents.

Ce modèle autorégressif a par exemple été utilisé pour modéliser les interactions stratégiques et la concurrence fiscale entre communes (Case et al., 1993 ; Brueckner, 1998), les externalités de voisinage dans les modèles hédoniques de prix immobiliers (Can, 1990, 1992 ; Macedo, 1998), les externalités spatiales influençant la distribution de la population dans les espaces urbains (Griffith, 1981 ; Griffith et Can, 1995) ou influençant la production de firmes concurrentes (Griffith, 1999).

212. Le modèle régressif croisé

Une autre façon de traiter l'interdépendance des observations est d'inclure une ou plusieurs « variables exogènes décalées » dans [2.1] :

$$y = X\beta + WZ\delta + \varepsilon \quad [2.6]$$

Z est une matrice de dimension (N,L) de L diverses variables correspondant ou non aux variables incluses dans X . WZ est alors la variable exogène décalée pour la matrice de poids W et δ est le vecteur des paramètres spatiaux indiquant l'ampleur de la corrélation spatiale existant entre les observations de y et celles de Z . Ainsi, 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, la production d'une région peut être influencée par la disponibilité du travail ou le montant du capital public dans les régions voisines (Florax et Folmer, 1992 ; Kelejian et Robinson, 1997 ; Boarnet, 1998). Cette spécification s'avère donc utile pour estimer les effets de débordement géographiques liés aux infrastructures publiques.

Au contraire du modèle autorégressif [2.2] et des modèles comportant une autocorrélation spatiale des erreurs (voir infra, paragraphe 22), l'estimation du modèle régressif croisé peut être basée sur les Moindres Carrés Ordinaires si H_1 est vérifiée et si H_2 est vérifiée pour la matrice des variables explicatives $X^* = [X \ WZ]$ de dimension $(N,K+L)$.

22. Autocorrélation des erreurs

Une autre façon d'incorporer l'autocorrélation spatiale dans un modèle de régression est de spécifier un processus spatial pour les erreurs.

221. Les erreurs suivent un processus autorégressif

La spécification la plus utilisée est un processus spatial autorégressif dans les erreurs :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \end{aligned} \quad [2.7]$$

λ est le paramètre représentant l'intensité de l'autocorrélation spatiale entre les résidus de la régression, u est le terme d'erreur tel que : $u \sim iid(0, \sigma^2 I)$.

Cette spécification appelle plusieurs remarques.

1/ Tout d'abord, si la matrice $(I - \lambda W)$ est non-singulière, le modèle [2.7] se réécrit sous la forme suivante :

$$y = X\beta + (I - \lambda W)^{-1}u \quad [2.8]$$

Cette expression signifie qu'un choc aléatoire dans une région i spécifique affecte non seulement la valeur de y de cette région, mais a également un impact sur les valeurs de y dans les autres régions (voisines de i ou non) à travers la transformation spatiale inverse $(I - \lambda W)^{-1}$ qui est une matrice pleine.

2/ Ensuite, de la formulation [2.8], il s'ensuit que :

$$V(y) = V(\varepsilon) = E[(I - \lambda W)^{-1}uu'(I - \lambda W')^{-1}] \quad [2.9a]$$

$$V(y) = V(\varepsilon) = \sigma^2[(I - \lambda W')(I - \lambda W)^{-1}] \quad [2.9b]$$

On retrouve pour ε et pour y une structure identique à celle du modèle spatial autorégressif [2.2]. Un processus spatial autorégressif des erreurs conduit donc à une covariance non nulle entre chaque paire d'erreurs et d'observations, mais qui décroît avec l'ordre de proximité.

Par ailleurs, la structure d'erreurs [2.9b] induit des éléments de la diagonale qui ne sont pas constants dans la matrice des variances-covariances, ce qui implique donc de l'hétéroscédasticité pour les erreurs ε , que u soit hétéroscédastique ou non (McMillen, 1992).

3/ Le modèle [2.7] peut aussi être exprimé en termes de variables spatialement filtrées. Après avoir prémultiplié les deux termes de [2.8] par $(I - \lambda W)$, on obtient l'expression suivante :

$$(I - \lambda W)y = (I - \lambda W)X\beta + u \quad [2.10]$$

Ce modèle de régression est un modèle où les variables expliquées *et* les variables explicatives sont filtrées spatialement, le terme d'erreur u étant homoscedastique.

4/ Enfin, le modèle [2.7] peut se réécrire d'une façon qui fait apparaître à la fois une variable endogène décalée et l'ensemble des variables exogènes décalées. En effet, en déplaçant le terme autorégressif de [2.10] à droite, on obtient le modèle suivant, dénommé le *modèle de Durbin spatial* (suridentifié) :

$$y = \lambda Wy + X\beta - \lambda WX\beta + u \quad [2.11]$$

On interprète parfois la détection de l'autocorrélation spatiale des erreurs comme un problème dans la spécification du modèle, telle que l'omission de variables significatives (telles que des variables exogènes supplémentaires) : l'effet qui n'est pas capté dans les variables explicatives se retrouve dans les erreurs sous la forme d'une autocorrélation spatiale. Dans cette perspective, l'autocorrélation spatiale est un instrument permettant de repérer l'existence de variables significatives, mais non prises en compte dans le modèle. Si d'autres variables explicatives pertinentes ne peuvent être trouvées, alors l'autocorrélation spatiale est un substitut à ces variables omises (les estimateurs ne sont pas les meilleurs tout en étant plus précis que ceux obtenus avec les MCO). Ainsi, ce modèle a été utilisé pour améliorer les estimations des lois de Kaldor (Bernat, 1996), celles des modèles de prix hédoniques (Pace et Gilley, 1997 ; Dubin, 1998 ; Dubin et al., 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 Etats-Unis (Rey et Montouri, 1999) ou en Europe (Baumont et al., 2000b).

222. Deux alternatives

Deux alternatives au processus autorégressif ont été proposées même si leur application est beaucoup moins fréquente dans la littérature. Il s'agit du processus moyenne mobile et de la spécification proposée par Kelejian et Robinson (1993).

Formellement, le processus moyenne mobile est spécifié de la façon suivante (Cliff et Ord, 1981 ; Haining, 1978, 1990 ; Moore, 1988 ; Cressie, 1993) :

$$\varepsilon = \gamma W u + u \quad [2.12]$$

où γ est le coefficient moyenne mobile et u est un terme d'erreur homoscedastique. Par conséquent, la matrice des variances-covariances qui résulte de ce processus est la suivante :

$$E[\varepsilon\varepsilon'] = \sigma^2(I + \gamma W)(I + \gamma W') = \sigma^2[I + \gamma(W + W') + \gamma^2 WW'] \quad [2.13]$$

Contrairement à la matrice des variances-covariances associée au processus autorégressif, [2.13] n'est pas une matrice pleine. Les covariances non nulles n'existent que pour les voisins de premier ordre ($W + W'$) et les voisins de second ordre (WW'). Ce processus suppose donc une interaction moins forte que pour le processus autorégressif. Enfin, sauf cas particulier, les éléments de la diagonale de [2.13] ne seront pas constants, ce qui provoque une hétéroscedasticité induite dans ε , quelle que soit la nature de u .

Kelejian et Robinson (1993, 1995) ont proposé une autre spécification dans laquelle l'erreur est la somme de deux termes indépendants, l'un étant associé à la région (une moyenne d'erreurs voisines) et l'autre étant spécifique à la localisation :

$$\varepsilon = W u + v \quad [2.14]$$

où u et v sont des erreurs homoscedastiques et indépendantes.

La matrice de variances-covariances de ε est alors :

$$E[\varepsilon\varepsilon'] = \sigma_u^2 I + \sigma_v^2 WW' = \sigma^2(I + \varphi WW') \quad [2.15]$$

où σ_u^2 et σ_v^2 sont les variances respectivement associées à u et v , $\sigma^2 = \sigma_u^2 + \sigma_v^2 > 0$ et $\varphi = \sigma_u^2 / \sigma_v^2$. L'interaction spatiale impliquée par [2.15] est encore plus limitée que dans [2.14] car elle ne concerne que les voisins de premier et de second ordre contenus dans les éléments non nuls de WW' . L'hétéroscedasticité est toujours une conséquence sauf si toutes les localisations ont le même nombre de voisins et des poids identiques. Toutefois, il s'agit d'une situation exclue par les hypothèses nécessitées pour une étude asymptotique du modèle (Kelejian et Robinson, 1993).

23. Autres processus

Les modèles présentés dans les paragraphes précédentes 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 des processus beaucoup plus généraux combinant tous ces aspects. La forme la plus générale est la forme autorégressive moyenne mobile (Spatial AutoRegressive Moving Average ou SARMA(p,q)) proposée par Huang (1984) :

$$\begin{aligned} y &= \rho_1 W_1 y + \rho_2 W_2 y + \dots + \rho_p W_p y + \varepsilon \\ \varepsilon &= \gamma_1 W_1 u + \gamma_2 W_2 u + \dots + \gamma_q W_q u + u \end{aligned} \quad [2.16]$$

W_i est la matrice de poids associée au $i^{\text{ème}}$ ordre de contiguïté. On pourrait envisager de la même un processus où les erreurs suivent un processus spatial autorégressif d'ordre q (Spatial AutoRegressive AutoRegressive) ou SARAR(p,q).

Les matrices W_i ne se réfèrent pas forcément à différents ordres de contiguïté. Ainsi, pour les régions européennes, Rietveld et Wintershoven (1999) cherchent à savoir si l'interdépendance entre régions d'un même pays est différente de l'interdépendance entre régions de pays différents. Ils estiment donc un SARAR (2,2) où les éléments de la matrice W_1 sont égaux à 1 si les régions d'un même pays sont contiguës et 0 sinon et où les éléments de la matrice W_2 sont égaux à 1 si les régions de pays différents sont contiguës et 0 sinon.

Anselin (1980, 1988a) a étudié les propriétés du processus SARAR(1,1) contenant une variable endogène décalée et une autocorrélation des erreurs. Formellement, le modèle s'exprime comme une combinaison de [2.2] et de [2.7], avec des poids différents :

$$\begin{aligned} y &= \rho W_1 y + X\beta + \varepsilon \\ \varepsilon &= \lambda W_2 \varepsilon + u \end{aligned} \quad [2.17]$$

Le premier terme multiplié par λW_2 donne : $\lambda W_2 y = \rho \lambda W_2 W_1 y + \lambda W_2 X\beta + \lambda W_2 \varepsilon$. Il reste à soustraire l'équation obtenue de l'équation initiale :

$$y = \rho W_1 y + \lambda W_2 y - \rho \lambda W_2 W_1 y + X\beta - \lambda W_2 X\beta + u \quad [2.18]$$

Il s'agit d'un modèle de Durbin spatial étendu avec des contraintes linéaires supplémentaires. Lorsque $W_1 = W_2 = W$, le modèle devient :

$$y = (\rho + \lambda) W y - \rho \lambda W^2 y + X\beta - \lambda W X\beta + u \quad [2.19]$$

Comme le modèle est suridentifié pour ρ , Anselin (1980) a donné les conditions et restrictions permettant une identification de tous les paramètres³.

Théoriquement, ce problème d'identification est important mais pratiquement, il est d'une portée moindre car en général, on retient rarement un modèle avec les deux types d'effets spatiaux, on cherche à modéliser la dépendance spatiale, soit par l'autocorrélation des erreurs, soit par une variable spatiale autorégressive, et non les deux. Anselin et Bera (1998) considèrent d'ailleurs que ces processus d'ordre supérieur sont le résultat d'une matrice de

³ Le modèle SARMA (1, 1) ne souffre pas de ce problème.

pois mal spécifiée et non de processus générateurs des données réalistes. Par exemple, 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 processus d'ordre supérieur alors que seule une matrice de poids bien spécifiée serait nécessaire.

24. Extensions : spécification de l'autocorrélation spatiale dans les modèles de panel et dans les modèles à variables qualitatives

Récemment, les développements en économétrie spatiale ont particulièrement porté sur l'incorporation de l'autocorrélation spatiale dans les modèles de panel et les modèles à variables qualitatives.

24.1. Les modèles spatio-temporels

L'introduction de la dimension temporelle accroît considérablement le nombre de cas qui peuvent être pris en compte dans les modèles économétriques spatiaux. Supposons qu'il y ait N régions, T périodes de temps et donc NT observations y_{it} . Considérons tout d'abord le modèle le plus simple (régression ordinaire) où l'on suppose des comportements uniformes pour tous les individus :

$$y_{it} = \sum_{k=1}^K x_{k,it} \beta_k + \varepsilon_{it} = x'_{it} \beta + \varepsilon_{it} \quad [2.20a]$$

$$\varepsilon_{it} \sim \text{iid}(0, \sigma^2) \quad \forall i, \forall t \quad [2.20b]$$

x'_{it} est le vecteur-ligne des observations pour une unité i au temps t de dimension $(1, K)$ et β est le vecteur des K paramètres inconnus. Les régresseurs exogènes peuvent être retardés dans le temps ou décalés dans l'espace.

Ces observations peuvent être regroupées par période temporelle. Notons y_t le vecteur de la variable expliquée, de dimension N qui regroupe les N observations de y_{it} pour la période t . X_t désigne la matrice (N, K) des observations des variables explicatives au temps t . ε_t est le vecteur des N erreurs au temps t . Le modèle se réécrit alors de la façon suivante :

$$y_t = X_t \beta + \varepsilon_t \quad [2.21]$$

Un modèle général incorporant les dépendances spatiale et temporelle dans la variable expliquée pourrait s'écrire de la façon suivante :

$$y_t = \lambda y_{t-1} + \rho W y_t + \gamma W y_{t-1} + X_t \beta + \varepsilon_t \quad [2.22]$$

Dans cette spécification, pour une région i à une période t , la variable expliquée dépend de sa valeur à la période précédente (y_{t-1}), de la valeur des variables expliquées des régions voisines à la même période ($W y_t$) et de la valeur des variables expliquées des régions voisines à la période précédente ($W y_{t-1}$). Des structures similaires peuvent être spécifiées pour l'erreur ε_t .

Des modèles plus simples peuvent être dérivés de [2.22]. Un cas particulier intéressant est le modèle où $\lambda = \rho = 0$ (Upton et Fingleton, 1985) :

$$y_t = \gamma W y_{t-1} + X_t \beta + \varepsilon_t \quad [2.23]$$

Au contraire des modèles précédents, l'estimation des paramètres de [2.23] peut être basée sur les moindres carrés ordinaires et c'est un modèle qui a été utilisé pour modéliser la diffusion d'un phénomène. Par exemple, Dubin (1995, 1997) étudie la diffusion d'une innovation dans le cadre d'un modèle LOGIT en données de panel : la probabilité d'adoption d'une innovation par une firme dépend entre autres de la proximité de firmes qui ont adopté cette innovation à la période précédente.

La dépendance spatiale sous la forme d'une variable endogène décalée ou sous la forme d'une autocorrélation spatiale des erreurs peut aussi être introduite dans d'autres modèles de panel, tels que le modèle SUR ou le modèle à erreurs composées. En revanche, la combinaison des effets fixes et de l'autocorrélation spatiale n'est pas possible. En effet, l'estimation de modèles spatiaux nécessite le caractère asymptotique dans le domaine spatial ($N \rightarrow \infty$) alors que les effets fixes (c'est-à-dire une variable muette pour chaque localisation) nécessitent l'estimation de N paramètres. Il n'y a donc pas d'estimateurs convergents et les effets fixes sont incompatibles avec les processus spatiaux ; une spécification à erreurs composées doit être considérée.

Considérons par exemple le modèle SUR spatial dans lequel $N > T$ (Arora et Brown, 1977 ; Hordijk, 1979, Anselin, 1988a, 1988c). Les paramètres sont constants dans l'espace mais variables pour chaque période. Le modèle s'écrit alors de la façon suivante :

$$y_{it} = x'_{it} \beta_t + \varepsilon_{it} \quad [2.24]$$

Dans ce modèle, x'_{it} est le vecteur-ligne des observations pour une unité i au temps t de dimension $(1, K_t)$ et β_t est le vecteur des K_t paramètres. Si l'on regroupe les N individus pour la période t , le modèle devient :

$$y_t = X_t \beta_t + \varepsilon_t \quad [2.25]$$

Regroupons à présent les T périodes, le nombre total de paramètres devient $K^* = \sum K_t$ et le modèle s'écrit :

$$Y = X\beta + \varepsilon \quad [2.26]$$

où Y est le vecteur $(NT, 1)$ des observations de la variable dépendante, X est la matrice bloc-diagonale de dimension (NT, K^*) contenant les X_t , β est le vecteur $(K^*, 1)$ des paramètres et ε est le vecteur $(NT, 1)$ des erreurs. Pour chaque équation, il est possible d'être confronté à un problème d'autocorrélation spatiale des erreurs :

$$\varepsilon_t = \lambda_t W \varepsilon_t + u_t \quad [2.27]$$

Les hypothèses suivantes sont posées sur les erreurs résiduelles u_t :

$$E(u_t) = 0 \quad \forall t \quad [2.28]$$

$$E(u_t u_t') = \sigma_t^2 I_N \quad [2.29a]$$

$$E(u_{it} u_{js'}) = 0 \quad \forall t, s \quad \forall i \neq j \quad (\text{absence d'autocorrélation croisée}) \quad [2.29b]$$

$$E(u_{it} u_{is}) = \sigma_{ts} \quad \forall i \quad \Rightarrow \quad E(u_t u_s') = \sigma_{ts} I_N \quad \forall t, s \quad (\text{covariance temporelle}) \quad [2.29c]$$

Les erreurs complètes u ont donc les propriétés suivantes : $E(u) = 0$ et $E(uu') = \Sigma \otimes I_N$, où Σ est la matrice d'élément générique : $[\sigma_{is}]$. Une application récente de ce modèle est celle de Rey et Montouri (1999) qui ont estimé un modèle SUR spatial pour étudier la convergence entre les régions américaines pour 2 sous-périodes.

242. Les modèles à variables qualitatives

La prise en compte de l'autocorrélation spatiale dans les modèles de choix discrets, dans les modèles censurés ou à sélection d'échantillon constitue aujourd'hui un thème de recherche actif, après avoir longtemps été ignoré, compte tenu de la complexité des procédures d'estimation de ces modèles. Les problèmes impliqués par la prise en compte de l'autocorrélation spatiale dans les modèles à variables qualitatives peuvent être illustrés à partir du modèle de choix discret avec la variable latente y_i^* :

$$y_i^* = x_i' \beta + \varepsilon_i \quad [2.30]$$

où ε_i est une variable aléatoire pour laquelle une distribution donnée est supposée (normale pour le modèle PROBIT, logistique pour un modèle LOGIT). y_i^* n'est pas observée mais on observe :

$$y_i = 1 \text{ lorsque } y_i^* \geq 0 \quad [2.31a]$$

$$y_i = 0 \text{ lorsque } y_i^* < 0 \quad [2.31b]$$

L'autocorrélation spatiale peut être introduite dans ce modèle sous la forme d'une variable endogène décalée ou d'une autocorrélation spatiale pour le terme d'erreur ε_i . Des problèmes similaires apparaissent pour ces deux cas. Considérons par exemple une autocorrélation spatiale des erreurs :

$$\varepsilon = (I - \lambda W)^{-1} u \quad [2.32a]$$

$$\Rightarrow \varepsilon_i = \sum_j \delta_{ij} u_j \quad [2.32b]$$

où δ_{ij} est un élément de $(I - \lambda W)^{-1}$. Pour pouvoir calculer la probabilité que $y_i = 1$ pour chaque observation : $P(y_i = 1) = P(\varepsilon_i < x_i' \beta)$, il faut déterminer la loi de probabilité de ε . Dans le cas du modèle PROBIT, on sait qu'une combinaison linéaire de variables normales est encore une variable normale. Si u est identiquement et indépendamment distribué d'espérance nulle et de matrice des variances-covariances I , alors ε suit une loi normale d'espérance nulle et de matrice des variances-covariances :

$$E(\varepsilon \varepsilon') = [(I - \lambda W')(I - \lambda W)]^{-1} \quad [2.33]$$

Cette expression n'est pas bloc-diagonale et conduit à des variances hétéroscédastiques. L'autocorrélation spatiale introduit donc des interdépendances entre observations qui interdisent le calcul des probabilités séparément pour chaque localisation : la probabilité $P(\varepsilon_i < x_i' \beta)$ ne peut plus être dérivée de la distribution normale standard univariée mais doit être explicitement dérivée comme la distribution marginale du vecteur multivarié à N

dimensions, dont la matrice des variances-covariances contient le paramètre autorégressif λ (McMillen, 1992 ; Jayet, 1993). Des problèmes similaires apparaissent dès lors que l'autocorrélation spatiale est introduite dans un modèle TOBIT 2 (McMillen, 1995a).

3. ESTIMATION DES MODELES SPATIAUX

Lorsque l'autocorrélation spatiale est modélisée, la méthode des moindres carrés ordinaires (MCO) n'est plus adaptée : les estimateurs obtenus par cette méthode ne sont pas convergents lorsqu'il y a une variable endogène décalée et ils sont inefficients en présence d'une autocorrélation spatiale des erreurs (paragraphe 31). D'autres méthodes d'estimation sont alors nécessaires pour trouver des estimateurs convergents et efficaces. La méthode la plus couramment utilisée est celle du maximum de vraisemblance à information complète (paragraphe 32) mais il est également possible d'avoir recours à la méthode des variables instrumentales ou à celle des moments généralisés (paragraphe 33).

31. La non-convergence des MCO et ses conséquences

Prenons le modèle spatial le plus général, incluant à la fois une variable endogène décalée et une autocorrélation spatiale des erreurs (modèle [2.17]) :

$$\begin{aligned} y &= X\beta + \rho W_1 y + \varepsilon \\ \varepsilon &= \lambda W_2 \varepsilon + u \\ u &\sim \text{iid}(0, \sigma^2 I) \end{aligned}$$

De cette écriture, il s'ensuit que $y = (I - \rho W_1)^{-1} X\beta + (I - \rho W_1)^{-1} \varepsilon$ et que $\varepsilon = (I - \lambda W_2)^{-1} u$. Par conséquent, la matrice des variances-covariances de l'erreur ε est :

$$V(\varepsilon) = E(\varepsilon\varepsilon') = \sigma^2(I - \lambda W_2)^{-1}(I - \lambda W_2')^{-1} \quad [3.1]$$

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

$$E(W_1 y \varepsilon') = E[W_1(I - \rho W_1)^{-1} X\beta \varepsilon' + W_1(I - \rho W_1)^{-1} \varepsilon \varepsilon'] \quad [3.2a]$$

$$E(W_1 y \varepsilon') = W_1(I - \rho W_1)^{-1} E(\varepsilon \varepsilon') \quad [3.2b]$$

$$E(W_1 y \varepsilon') = \sigma^2 W_1(I - \rho W_1)^{-1}(I - \lambda W_2)^{-1}(I - \lambda W_2')^{-1} \quad [3.2c]$$

Cette dernière expression est non nulle en général, les éléments de la variable endogène décalée sont corrélés avec ceux des erreurs et les paramètres du modèle [2.17] ne peuvent donc pas être estimés d'une façon convergente par les MCO (Kelejian et Prucha, 1998). 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.

En revanche, pour le modèle spatial autorégressif [2.2], les estimateurs des MCO ne sont pas convergents car la variable endogène décalée $W_1 y$ est corrélée avec l'erreur ε , quelle que soit la distribution de cette erreur : si $\lambda = 0$ et $W_1 = W$ dans [3.2c], alors $E(W_1 y \varepsilon') = \sigma^2 W(I - \rho W)^{-1}$. La matrice $(I - \rho W)^{-1}$ étant une matrice pleine et non triangulaire, le décalage spatial pour une observation donnée i , $(W_1 y)_i$, n'est pas seulement

corrélé avec le terme d'erreur en i , mais aussi avec tous les termes d'erreurs de toutes les localisations.

Si l'on considère à présent le modèle à erreurs autorégressives, l'estimateur de β par les MCO est sans biais mais inefficace puisque les erreurs ne sont pas homoscédastiques. Il faut donc utiliser une autre méthode d'estimation que les MCO. La solution théorique dans ce cas consiste dans l'application des moindres carrés quasi-généralisés (MCQG). Cette procédure n'est pas applicable au cas spatial puisque l'estimateur de λ 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 [2.11] n'est pas non plus convergente, l'une des conditions données par Amemiya (1985) n'étant pas vérifiée.

Finalement, pour aboutir à des estimateurs convergents et efficaces des paramètres des modèles spatiaux, il faut donc utiliser d'autres méthodes d'estimation, la plus utilisée aujourd'hui restant la méthode d'estimation par le maximum de vraisemblance.

32. Estimation par le maximum de vraisemblance

La première étude de l'estimation des modèles spatiaux par le maximum de vraisemblance a été donnée par Ord (1975). Les conditions pour la convergence, l'efficacité et la normalité asymptotique des estimateurs du maximum de vraisemblance peuvent être dérivées du cadre général de Heijmans et Magnus (1986a, 1986b) pour le modèle spatial autorégressif et de celui de Magnus (1978) pour le modèle à erreurs autorégressives. Mises à part les 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 spatiaux et sur l'espace des paramètres des coefficients spatiaux (Anselin et Kelejian, 1997 ; Kelejian et Prucha, 1998, 1999a, 1999b ; Pinkse et Slade, 1998 ; Pinkse, 1998, 1999). En pratique, ces conditions sont largement satisfaites par les poids basés sur la contiguïté mais pas nécessairement pour les poids généraux.

Sous l'hypothèse de normalité des résidus, on peut dériver la fonction de log-vraisemblance dans le cas général [2.17] ce qui permet de déterminer l'espace des paramètres de ρ et de λ . Les mêmes principes permettent de déterminer la forme de la fonction de log-vraisemblance pour les différents modèles de panel (paragraphe 321). Les procédures d'estimation faisant appel à la fonction de vraisemblance concentrée peuvent être illustrées à partir de deux cas particuliers : modèle autorégressif et erreurs autocorrélées (paragraphe 322). Dans tous les cas cependant, l'évaluation du jacobien pose problème (paragraphe 323)

321. La fonction de vraisemblance dans le cas général

Pour le modèle général [2.17], le point de départ est l'hypothèse de normalité des termes d'erreur. La fonction de vraisemblance pour le vecteur normal multivarié $u \sim \text{Nid}(0, \sigma^2 I)$ est :

$$L(u) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2} u' u\right\} \quad [3.3]$$

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

$$J = \det(\partial u / \partial y) = |I - \rho W_1| |I - \lambda W_2| \quad [3.4]$$

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

$$\ln L(y|\rho, \lambda, \beta) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) + \ln|I - \rho W_1| + \ln|I - \lambda W_2| - \frac{1}{2\sigma^2} u'u \quad [3.5]$$

De cette expression, on déduit l'espace des paramètres de ρ et λ . Soient ω_{\max} la valeur propre positive maximale et ω_{\min} la valeur propre négative la plus grande en valeur absolue de W . Alors $|I - \rho W_1|$ et $|I - \lambda W_2|$ sont positifs si l'inégalité suivante est respectée :

$$1/\omega_{\min} < \psi < 1/\omega_{\max} \quad [3.6]$$

avec $\psi = \rho$ ou λ . Pour une matrice standardisée en ligne, $\omega_{\max} = 1$ et $\omega_{\min} > -1$ (Anselin, 1982 ; Anselin, 1988a, chap.6). Si les erreurs suivent un processus moyenne mobile [2.12], δ doit être compris dans l'intervalle : $]-1/\omega_{\max}; -1/\omega_{\min}[$.

On trouvera dans Anselin (1988a, 1988b), les expressions du score et de la matrice d'information pour le modèle général comportant une variable endogène décalée, une autocorrélation spatiale des erreurs et des résidus hétéroscédastiques. 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 plus simples [2.2] et [2.7] admettent des solutions issues des conditions du premier ordre permettant de construire une fonction de log-vraisemblance concentrée.

Dans le cadre des modèles de panel, les mêmes principes peuvent être mobilisés pour déterminer la forme de la fonction de vraisemblance. Considérons ainsi le modèle SUR spatial avec autocorrélation spatiale des erreurs [2.25] :

$$\begin{aligned} y_t &= X_t \beta_t + \varepsilon_t \\ \varepsilon_t &= \lambda_t W \varepsilon_t + u_t \end{aligned}$$

avec : $E(u_t u_t') = \sigma_t^2 I_N$ et $E(u_t u_s') = \sigma_{ts} I_N$. Notons $B_t = (I - \lambda_t W)^{-1}$ pour simplifier les écritures. On a alors : $\varepsilon_t = B_t u_t$. Par conséquent :

$$E(\varepsilon_t \varepsilon_s') = E(B_t u_t u_s' B_s') = \sigma_{ts} B_t B_s' \quad [3.7]$$

$$\Rightarrow \Omega = E(\varepsilon \varepsilon') = B(\Sigma \otimes I) B' \quad [3.8]$$

Dans ces dernières expressions, ε est le vecteur d'erreur complet de dimension $(NT, 1)$ et B est la matrice bloc-diagonale carrée de dimension (NT, NT) contenant les B_t .

L'estimation de ce modèle ne peut se faire à partir de la méthode des MCQG car cette méthode n'est pas efficace. En effet, il n'existe pas d'estimateurs convergents pour les paramètres λ_t . La solution consiste donc en l'application de la méthode du maximum de vraisemblance, sous l'hypothèse de normalité des erreurs résiduelles u_t . La fonction de log-vraisemblance pour Y est :

$$\ln L(Y | \lambda_1, \dots, \lambda_T, \beta_1, \dots, \beta_T) = -(NT/2) \ln(2\pi) - (1/2) \ln |\Omega| - (1/2) (y - X\beta)' \Omega^{-1} (y - X\beta) \quad [3.9]$$

Avec : $|\Omega| = |B(\Sigma \otimes I_N) B'| = |\Sigma|^N |B|^2$; $\ln |\Omega| = N \ln |\Sigma| + 2 \ln |B|$. L'annulation des dérivées premières forme un système non linéaire n'admettant pas de solutions analytiques. Ce système doit donc être résolu par des méthodes numériques.

De la même façon, on pourra trouver dans Anselin (1988a), les expressions des fonctions de log-vraisemblance pour le modèle SUR spatial avec variable endogène décalée et pour le modèle avec erreurs composées et autocorrélation spatiale.

322. Deux cas particuliers

En pratique, l'estimation des modèles spatiaux peut s'effectuer à partir de la maximisation de la fonction de log-vraisemblance complète grâce à diverses techniques d'optimisation non-linéaires. Dans le cas de modèles plus simples pourtant, comme le modèle autorégressif et le modèle à erreurs autocorrélées, l'estimation s'effectue à partir de la fonction de log-vraisemblance concentrée. Le principe de la méthode est de résoudre une partie des équations associées aux conditions du premier ordre et d'introduire ensuite les solutions obtenues dans la fonction de log-vraisemblance. On obtient une fonction de log-vraisemblance simplifiée, concentrée qui ne dépend plus que de quelques paramètres, un dans le meilleur des cas. Dès lors, on peut trouver des estimations de ces paramètres par balayage, c'est-à-dire en évaluant la fonction de log-vraisemblance pour un petit intervalle de ces paramètres. On trouve alors un maximum local qui correspond à la solution (Upton et Fingleton, 1985).

1/ Dans le cas du modèle autorégressif [2.2], la fonction de log-vraisemblance est construite en appliquant à ce cas particulier la formule générale :

$$\ln L = |I - \rho W| - \frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) - \frac{(y - \rho W y - X\beta)'(y - \rho W y - X\beta)}{2\sigma^2} \quad [3.10]$$

A partir des conditions de premier ordre usuelles, les estimateurs pour β et σ^2 sont obtenus comme :

$$\hat{\beta}_{ML} = (X'X)^{-1} X'(I - \rho W)y \quad [3.11]$$

$$\hat{\sigma}_{ML}^2 = \frac{(y - \rho W y - X\hat{\beta}_{ML})'(y - \rho W y - X\hat{\beta}_{ML})}{N} \quad [3.12]$$

Conditionnellement à ρ , [3.11] et [3.12] sont les estimateurs des MCO appliqués au modèle filtré [2.3]. La substitution de ces estimateurs dans la fonction de log-vraisemblance aboutit à la fonction de log-vraisemblance concentrée qui ne dépend plus alors que de ρ :

$$\ln L_c = C - \frac{N}{2} \ln \left[\frac{(\hat{\varepsilon}_0 - \rho \hat{\varepsilon}_L)'(\hat{\varepsilon}_0 - \rho \hat{\varepsilon}_L)}{N} + |I - \rho W| \right] \quad [3.13]$$

Où $C = -(N/2) \ln(2\pi) - (N/2)$ et où $\hat{\varepsilon}_0$ et $\hat{\varepsilon}_L$ sont respectivement les résidus des régressions de y sur X et de Wy sur X . Une estimation pour ρ est alors obtenue par une optimisation numérique de la fonction de log-vraisemblance concentrée.

La matrice des variances-covariances asymptotique est donnée par l'inverse de la matrice d'information :

$$\text{AsyV}(\rho, \beta, \sigma^2) = \begin{bmatrix} \text{tr}(W_A)^2 + \text{tr}(W_A W_A') + \frac{[W_A X \beta]' [W_A X \beta]}{\sigma^2} & \frac{(X' W_A X \beta)'}{\sigma^2} & \frac{\text{tr}(W_A)}{\sigma^2} \\ X' W_A X \beta & \frac{X' X}{\sigma^2} & 0 \\ \frac{\text{tr}(W_A)}{\sigma^2} & 0 & \frac{N}{2\sigma^4} \end{bmatrix}^{-1} \quad [3.14]$$

Dans [3.14], $W_A = W(I - \rho W)^{-1}$. On note que la covariance entre β et la variance est nulle comme dans le modèle standard, mais pas entre ρ et la variance, la matrice n'est donc pas bloc-diagonale.

2/ Dans le cas de l'autocorrélation spatiale des erreurs, la fonction de log-vraisemblance prend la forme suivante :

$$\ln L = \ln|I - \lambda W| - \frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) - \frac{(y - X\beta)' \Omega(\lambda)^{-1} (y - X\beta)}{2\sigma^2} \quad [3.15]$$

$\Omega(\lambda)$ est tel que $\sigma^2 \Omega(\lambda) = E(\varepsilon\varepsilon') = \sigma^2 [(I - \lambda W)'(I - \lambda W)]^{-1}$.

Les conditions de premier ordre fournissent donc l'estimateur des MCG pour β et σ^2 , conditionnellement à λ :

$$\hat{\beta}_{ML} = [X' \Omega(\lambda)^{-1} X]^{-1} X' \Omega(\lambda)^{-1} y \quad [3.16]$$

$$\hat{\sigma}_{ML}^2 = \frac{[y - X\hat{\beta}_{ML}]' \Omega(\lambda)^{-1} [y - X\hat{\beta}_{ML}]}{N} \quad [3.17]$$

Si λ est connu, l'estimateur du maximum de vraisemblance est équivalent à celui des MCO appliqué aux variables filtrées de [2.10]⁴.

La substitution de [3.16] et [3.17] dans la fonction de log-vraisemblance [3.15] aboutit à la fonction de log-vraisemblance concentrée suivante :

$$\ln L_c = C - \frac{N}{2} \ln \left(\frac{(y - X\hat{\beta}_{ML})' \Omega(\lambda)^{-1} (y - X\hat{\beta}_{ML})}{N} \right) + |I - \lambda W| \quad [3.18]$$

La matrice des variances-covariances asymptotique pour les estimateurs est similaire à la forme générale de Magnus (1978) et Breusch (1980) et est bloc-diagonale pour β et σ^2 :

$$\text{Asy}V(\lambda, \beta, \sigma^2) = \begin{bmatrix} \text{tr}(W_B)^2 + \text{tr}(W_B' W_B) & 0 & \frac{\text{tr}(W_B)}{\sigma^2} \\ 0 & \frac{X'(I - \lambda W)'(I - \lambda W)X}{\sigma^2} & 0 \\ \frac{\text{tr}(W_B)}{\sigma^2} & 0 & \frac{N}{2\sigma^4} \end{bmatrix} \quad [3.19]$$

où $W_B = W(I - \lambda W)^{-1}$.

323. Problèmes pratiques liés à l'estimation par le maximum de vraisemblance

L'estimation des modèles spatiaux par le maximum de vraisemblance nécessite la manipulation de matrices dont la dimension est égale au nombre d'observations. Par exemple, pour le modèle autorégressif et le modèle à erreurs autorégressives, le calcul de la matrice des variances-covariances nécessite l'évaluation respectivement des matrices $W_A = W(I - \rho W)^{-1}$ et $W_B = W(I - \lambda W)^{-1}$ qui sont des matrices carrées, de dimension N , pleines et qui ne se prêtent

⁴ Si l'on considère la forme moyenne mobile de l'autocorrélation spatiale, l'expression de l'estimateur des MCG [3.16] nécessite l'inversion d'une matrice des variances-covariances de dimension (N, N) : $\Omega(\lambda)^{-1} = [I + \gamma(W + W') + \gamma^2 W W']^{-1}$. Cette complication a sans doute contribué à limiter l'application pratique de ce modèle (Sneek et Rietveld, 1998).

donc pas à l'application des algorithmes de traitement des matrices contenant beaucoup d'éléments nuls. La taille de l'échantillon dépend alors de la plus grande inverse de matrice qui peut être calculée avec une précision numérique suffisante, précision variable selon les logiciels⁵.

Le problème principal de l'estimation par le maximum de vraisemblance est la présence du jacobien dans la fonction de log-vraisemblance. Le jacobien est de la forme $|I - \rho W|$ pour un modèle autorégressif ou $|I - \lambda W|$ pour un modèle à erreurs autocorrélées. Par conséquent, la maximisation de la fonction de log-vraisemblance nécessite une optimisation non-linéaire qui nécessite l'évaluation du terme jacobien pour chaque nouvelle valeur de ρ ou de λ . Même pour des échantillons de taille modérée, cela peut être une opération lourde puisqu'il s'agit de calculer le déterminant d'une matrice carrée de dimension N , N étant le nombre d'observations. Plusieurs solutions ont été proposées dans la littérature.

1/ *La solution la plus ancienne* pour l'estimation de modèles autorégressifs par le maximum de vraisemblance 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 :

$$|I - \rho W| = \prod_{i=1}^N (1 - \rho \omega_i) \quad [3.20]$$

$$\text{soit : } \ln |I - \rho W| = \sum_{i=1}^N \ln(1 - \rho \omega_i) \quad [3.21]$$

L'avantage de cette forme simplifiée apparaît lors de la maximisation de la fonction de log-vraisemblance complète ou de la procédure de balayage de la fonction de log-vraisemblance concentrée. Si l'on utilise la formulation $|I - \rho W|$, il faut calculer le déterminant de $I - \rho W$ à chaque étape. En revanche, si l'on utilise la formulation simplifiée, il n'est besoin de calculer qu'une fois pour toutes les valeurs propres de W et l'évaluation à chaque étape devient plus facile. Cette propriété permet alors d'écrire la fonction de log-vraisemblance complète en une somme d'éléments correspondant aux observations individuelles. Ainsi, des « pseudo-observations » sont construites pour les éléments du jacobien, chaque terme $1 - \rho \omega_i$ étant relié à une pseudo-variable ω_i . Par exemple, pour le modèle autorégressif, la fonction de log-vraisemblance peut être exprimée comme :

$$\ln L = \sum_i \left[\ln(1 - \rho \omega_i) - \frac{\ln(\sigma^2)}{2} - \frac{(y_i - \rho \{Wy\}_i - x_i \beta)^2}{2\sigma^2} \right] \quad [3.22]$$

Cette formulation additive pour chaque observation est conforme à bon nombre de routines d'optimisation non-linéaires. Certains auteurs fournissent ainsi des codes pour quelques

⁵ Pour les très grands échantillons, même si le calcul du ratio des coefficients à leurs écarts-types est impossible, l'inférence asymptotique reste réalisable dans le cas d'erreurs autocorrélées. Dans ce dernier cas en effet, la matrice des variances-covariances asymptotique est bloc-diagonale et les statistiques asymptotiques pour β peuvent donc être calculées sans connaître la précision du paramètre autorégressif λ (Benirschka et Binkley, 1994 ; Pace et Barry, 1996), l'inférence du paramètre autorégressif étant alors basée sur le test du ratio de vraisemblance. Cette approche n'est pas applicable pour le modèle autorégressif pour lequel la matrice des variances-covariances asymptotique n'est pas bloc-diagonale. Des tests du ratio de vraisemblance doivent donc être considérés pour tout sous-ensemble de coefficients (Pace et Barry, 1997).

logiciels économétriques basés sur cette propriété (Bivand, 1992 ; Griffith, 1988b, 1993 ; Anselin et Hudak, 1992 ; Anselin et al., 1993 ; Li, 1996).

Si le calcul des valeurs propres pour une matrice de poids W asymétrique pose problème, Ord (1975) a donné une propriété intéressante lorsque W est le résultat d'une standardisation d'une matrice de poids symétrique W^* . Les valeurs propres de W peuvent se trouver en calculant les valeurs propres de la matrice symétrique, $D^{1/2}W^*D^{1/2}$ où D est une matrice diagonale, chaque élément de la diagonale principale étant égal à l'inverse de la somme des éléments de la ligne correspondante de W .

Cette méthode est très largement utilisée mais pour de très grands échantillons, le calcul des valeurs propres peut devenir numériquement instable. Une variante de cette procédure a été proposée par Anselin et Smirnov (1999) qui consiste à évaluer directement les coefficients de la fonction caractéristique. Les simulations effectuées par les auteurs indiquent que cette méthode est la seule méthode « directe » capable de calculer les estimations du maximum de vraisemblance pour de très grands échantillons (plus d'un million d'observations).

2/ *D'autres techniques* exploitent la structure particulière des matrices de poids contenant beaucoup d'éléments nuls. Il s'agit des méthodes de factorisation de matrices qui s'avèrent très puissantes pour évaluer rapidement le jacobien : la décomposition de Cholesky pour une matrice symétrique et la décomposition de LU sinon. Pace (1997) et Pace et Barry (1997a et 1997b) ont démontré que ces approches donnent des temps de calcul raisonnables pour des échantillons de dizaines de milliers d'observations. Par exemple, pour une matrice symétrique, la factorisation de Cholesky consiste à résoudre :

$$I - \rho W = LL' \quad [3.23]$$

où L est une matrice triangulaire inférieure, le facteur de Cholesky de la matrice. Le jacobien est alors $|I - \rho W| = |L||L'| = |L|^2$. Puisque le déterminant d'une matrice triangulaire nécessite uniquement les éléments de la diagonale, le jacobien en forme logarithmique est :

$$|I - \rho W| = 2 \sum_{i=1}^N \ln(l_{ii}) \quad [3.24]$$

où l_{ii} sont les éléments de la diagonale de L .

3/ *Enfin*, certains auteurs approximent le jacobien par des fonctions polynomiales ou par les puissances successives de la matrice de poids (Griffith, 1992a ; Martin, 1993 ; Griffith et Sone, 1995). Une approche plus récente suggérée par Barry et Pace (1999) est basée sur des simulations de Monte-Carlo et est capable de traiter des échantillons de plus d'un million d'observations.

33. Autres méthodes d'estimation

Dans le modèle autorégressif, le problème principal 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 (1980, 1988a), Land et Deane (1992), Kelejian et Robinson (1993) ou Kelejian et Prucha (1998). Pour le modèle autorégressif, les estimateurs sont convergents compte tenu d'un choix approprié d'instruments, ce choix déterminant aussi l'efficacité des estimateurs. 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 supérieurs.

En notant Z la matrice (N, P) des instruments (avec $P \geq K + 1$), l'estimateur des variables instrumentales s'écrit :

$$\hat{\beta}_{IV} = [\tilde{X}'Z(Z'Z)^{-1}Z'\tilde{X}]^{-1}\tilde{X}'Z(Z'Z)^{-1}Z'y \quad [3.25]$$

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

Cette approche peut facilement être étendue à des structures d'erreurs plus complexes (Anselin, 1988a) ou à un modèle SUR avec variable endogène décalée. 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 des estimateurs convergents pour le coefficient spatial dans le modèle à erreurs autocorrélées, Kelejian et Prucha (1998, 1999a) ont récemment développé une approche par 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 [2.7] : si $u \sim \text{iid}(0, \sigma^2 I)$, les trois conditions sont les suivantes :

$$\begin{aligned} E[u'u / N] &= \sigma^2 \\ E[u'W'Wu / N] &= \sigma^2(1/N)\text{tr}(W'W) \\ E[u'Wu / N] &= 0 \end{aligned} \quad [3.26]$$

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 λ , λ^2 et σ^2 .

Les deux méthodes d'estimation peuvent être combinées pour estimer les paramètres du modèle général [2.17] contenant à la fois une variable autorégressive et une autocorrélation spatiale des erreurs (Kelejian et Prucha, 1998) pour obtenir des estimateurs convergents.

D'autres méthodes d'estimation ont été proposées dans la littérature. L'une d'entre elles est la *méthode spatiale de filtrage* (Getis, 1990, 1995) qui consiste à « filtrer » les variables spatialement dépendantes pour les transformer en variables indépendantes. Ce filtrage s'effectue à partir des statistiques d'autocorrélation spatiale locales proposées par Getis et Ord (1992, 1995). Il est alors possible d'utiliser les moindres carrés ordinaires ainsi que toutes les autres mesures de régression qui sont biaisées lorsqu'on utilise le maximum de vraisemblance (comme par exemple, le R^2). Cette approche suppose néanmoins que l'autocorrélation spatiale soit systématiquement une nuisance qu'il convient d'éliminer.

Le cas des modèles à variables qualitatives incorporant de l'autocorrélation spatiale a été abordé par Case (1992), McMillen (1992, 1995a, 1995b) et Pinkse et Slade (1998). Ces modèles posent problème dans la mesure où l'interdépendance impliquée par l'autocorrélation spatiale produit une fonction de vraisemblance avec de multiples intégrales, rendant l'estimation directe pratiquement impossible. De plus, l'hétéroscédasticité des erreurs est induite lorsqu'on spécifie l'autocorrélation spatiale par une variable endogène décalée ou une autocorrélation spatiale des erreurs (sauf dans un cas particulier abordé par Case, 1992), les méthodes d'estimation pour les modèles PROBIT avec des observations dépendantes mais des erreurs homoscédastiques proposées par Avery et al. (1983) ou Poirier-Ruud (1988) ne sont donc pas convergentes.

Dans le cadre des modèles de choix discrets, McMillen (1992, 1995b) propose d'utiliser la méthode d'estimation « EM » ou « Espérance - Maximisation » pour les modèles PROBIT avec autocorrélation spatiale des erreurs ou variable endogène décalée. Cette procédure d'estimation remplace la variable discrète dépendante par l'espérance de la variable latente continue. Le modèle est ensuite estimé par maximum de vraisemblance en considérant la variable construite comme une variable standard continue et dépendante. La procédure d'estimation – calcul de l'espérance de la variable dépendante et estimation du modèle par maximum de vraisemblance – est répétée jusqu'à convergence et les estimations des paramètres obtenues sont les estimations du maximum de vraisemblance. Des simulations de Monte-Carlo effectuées par McMillen (1995b) suggèrent cependant qu'il n'y a pas de précision supplémentaire apportée par cette méthode pour les petits échantillons. De plus, cette méthode pose problème dans la mesure où la matrice d'information ne peut pas être déterminée analytiquement (la fonction de vraisemblance comporte N intégrales). LeSage (1999) propose alors l'utilisation de méthodes d'estimation bayésiennes qui fournissent les mêmes estimations que celles données par la méthode du maximum de vraisemblance mais ne souffrent pas de ce problème. Ces méthodes ont par ailleurs été appliquées aux modèles avec variable endogène décalée [2.2] et avec autocorrélation spatiale des erreurs [2.7] par Hepple (1995a, 1995b) et LeSage (1997).

4. LES TESTS EN ECONOMETRIE SPATIALE

La modélisation des données spatialisées peut s'effectuer de différentes façons : on peut inclure des variables décalées (endogènes ou exogènes), une autocorrélation spatiale des erreurs ou estimer les modèles avec différentes matrices de poids. Le choix entre ces différentes alternatives passe par la mise en œuvre de tests de spécification dont les origines remontent jusqu'au test de Moran (1950a, 1950b). Ce test est resté dans l'obscurité jusqu'à sa redécouverte par Cliff et Ord (1972) et Burridge (1980).

Les trois grands principes de tests en économétrie standard sont le test du multiplicateur de Lagrange (ou test du score de Rao (1947)), le test du ratio de vraisemblance et le test de Wald. Ces tests ont également été mis à contribution dans la recherche de la spécification du modèle. Au début, la littérature en économétrie spatiale a été dominée par les deux derniers types de tests (Brandsma et Kelletaper, 1979 ; Anselin, 1980 ; Cliff et Ord, 1981). Cependant, ils nécessitent l'estimation du modèle non-contraint qui doit être estimé par des méthodes non-linéaires. Au contraire, le test du score est basé uniquement sur les résultats du modèle sous l'hypothèse nulle et il s'agit la plupart du temps du modèle linéaire standard estimé par les MCO (paragraphe 41). Durant les 15 dernières années, de tels tests ont été développés en particulier par Anselin (Anselin, 1988a, 1988b, 1998) et ils constituent la première étape dans la recherche de la spécification du modèle (paragraphe 42).

41. Le test de l'autocorrélation spatiale

Les tests de l'autocorrélation spatiale peuvent se diviser en plusieurs catégories. Tout d'abord, le test de Moran est le test le plus ancien et encore le plus utilisé. Il vise à tester l'autocorrélation spatiale des résidus lorsque les erreurs suivent un processus autorégressif, moyenne mobile ou la spécification de Kelejian et Robinson (paragraphe 411). Plus récemment, les tests du multiplicateur de Lagrange ont été développés et ils peuvent être soit unidirectionnels – lorsqu'une hypothèse simple est testée en supposant une spécification correcte pour le reste du modèle (paragraphe 412) –, soit multidirectionnels – lorsque plusieurs types de dépendance spatiale sont testés (paragraphe 413) –. La puissance et la

robustesse de ces tests ont été étudiées dans quelques simulations de Monte-Carlo (paragraphe 414).

411. Le test de Moran

Le test I a été développé par Moran (1950a, 1950b) comme une extension à deux dimensions du test de corrélation temporelle dans les séries temporelles univariées. Il a été adapté aux résidus d'une régression et se présente formellement de la façon suivante en notation matricielle :

$$I = \frac{N}{S_0} \left(\frac{\tilde{\varepsilon}' W \tilde{\varepsilon}}{\tilde{\varepsilon}' \tilde{\varepsilon}} \right) \quad [4.1]$$

où $\tilde{\varepsilon} = y - X\tilde{\beta}$ est le vecteur des résidus estimés des MCO ($\tilde{\beta} = (X'X)^{-1}X'y$) 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 : $S_0 = N$.

Sous l'hypothèse nulle d'indépendance spatiale, le test I de Moran est un test localement meilleur invariant (King, 1981) et est asymptotiquement un test du ratio de vraisemblance de $\lambda = 0$ dans [2.7] ou de $\delta = 0$ dans [2.12] (BurrIDGE, 1980). Sous l'hypothèse nulle, Cliff et Ord (1972) ont dérivé les deux premiers moments de I :

$$E(I) = \frac{tr(MW)}{N - K} \quad [4.2]$$

$$\text{et } V(I) = \frac{tr(MWMW') + tr(MW)^2 + \{tr(MW)\}^2}{(N - K)(N - K + 2)} - [E(I)]^2 \quad [4.3]$$

où $M = I - (X'X)^{-1}X'$.

Le test se base alors sur la statistique de Moran centrée et réduite : $Z(I) = [I - E(I)] / V(I)$. Pour des résidus normalement distribués et une matrice de poids « bien élevée », $Z(I)$ suit asymptotiquement une loi normale centrée et réduite. Pinkse (1998) et Kelejian et Prucha (1999b) donnent ainsi des conditions formelles et des preuves pour la normalité asymptotique du test de Moran dans plusieurs types de modèles. L'inférence statistique peut également être basée sur la distribution exacte de I en échantillon fini. Tiefelsdorf et Boots (1995) et Hepple (1998) ont dérivé un test exact en utilisant les résultats sur les ratios de formes quadratiques de variables normales (une présentation complète de ces méthodes se trouve dans Tiefelsdorf, 1998, 2000). Le test de Moran a par ailleurs été étendu aux modèles comportant des variables explicatives endogènes par Anselin et Kelejian (1997) et aux modèles PROBIT par Pinkse (1998, 1999).

412. Tests unidirectionnels

1/ On considère tout d'abord le cas où *les erreurs suivent un processus spatial autorégressif* [2.7] : $\varepsilon = \lambda W\varepsilon + u$ et pour lequel on teste $H_0 : \lambda = 0$. Sous l'hypothèse nulle, on retrouve le modèle linéaire classique [2.1]. D'une façon générale, la statistique du test du multiplicateur de Lagrange se calcule de la façon suivante :

$$ML = d'(\tilde{\theta})I(\tilde{\theta})^{-1}d(\tilde{\theta}) \quad [4.4]$$

Dans [4.4], $d(\tilde{\theta})$ et $I(\tilde{\theta})$ sont respectivement le vecteur score et la matrice d'information évalués sous l'hypothèse nulle. Pour le modèle à erreurs autorégressives, $\theta = (\beta', \sigma^2, \lambda)'$ et la fonction de log-vraisemblance est donnée en [3.15]. Quelques développements matriciels permettent de trouver l'expression suivante (Burridge, 1980) :

$$LM_{ERR} = \frac{[\tilde{\varepsilon}' W \tilde{\varepsilon} / \tilde{\sigma}^2]^2}{T} \quad [4.5]$$

avec $T = tr[(W' + W)W]$. $\tilde{\varepsilon}$ et $\tilde{\sigma}^2$ sont les estimations obtenues sous l'hypothèse nulle.

Puisqu'il n'y a qu'une seule contrainte : $LM_{ERR} \xrightarrow{D} \chi_1^2$.

La statistique de test est la même si on spécifie comme hypothèse alternative le processus moyenne mobile [2.12] et comme test $H_0 : \delta = 0$. LM_{ERR} est donc localement optimal pour les deux alternatives (autorégressive et moyenne mobile) et lorsque l'hypothèse nulle est rejetée, le test ne donne pas d'indications quant à la nature du processus des erreurs.

On trouvera dans Anselin (1988a, 1988c) une extension de ce test appliqué au modèle SUR avec autocorrélation des erreurs [2.25] et au modèle à erreurs composées et erreurs spatialement autocorrélées. Par ailleurs, pour des processus spatiaux d'ordre supérieur, on généralise facilement la statistique LM_{ERR} . Par exemple, si l'on considère un processus autorégressif d'ordre q :

$$\varepsilon = \lambda_1 W_1 \varepsilon + \lambda_2 W_2 \varepsilon + \dots + \lambda_q W_q \varepsilon + u \quad [4.6]$$

Pour le test $H_0 : \lambda_1 = \lambda_2 = \dots = \lambda_q = 0$, la statistique devient :

$$LM_{ERR_{\lambda_1, \lambda_2, \dots, \lambda_q}} = \sum_{l=1}^q \frac{[\tilde{\varepsilon}' W_l \tilde{\varepsilon} / \tilde{\sigma}^2]}{T_l} \quad [4.7]$$

où : $T_l = tr[W_l' W_l + W_l^2]$, $l = 1, 2, \dots, q$. Sous l'hypothèse nulle $LM_{ERR_{\lambda_1, \lambda_2, \dots, \lambda_q}} \xrightarrow{D} \chi_q^2$. Il s'agit simplement de la somme des tests individuels (Anselin et Bera, 1998). La même statistique de test reste valable si on prend le processus moyenne mobile [2.12] au lieu de [2.7].

2/ Pour la spécification de l'erreur proposée par Kelejian et Robinson (1995), un test du multiplicateur de Lagrange peut être dérivé suivant le même principe.

En reprenant les notations du modèle [2.14] et pour le test de l'hypothèse nulle $H_0 : \varphi = 0$, Anselin (1998) montre que la statistique de test s'écrit :

$$KR = \left[\frac{\tilde{\varepsilon}' W W' \tilde{\varepsilon}}{\tilde{\sigma}^2} - T_1 \right]^2 / 2 \left(T_2 - \frac{T_1}{N} \right) \quad [4.8]$$

où $T_1 = tr(WW')$ et $T_2 = tr(WW'WW')$. Sous $H_0 : KR \xrightarrow{D} \chi_1^2$.

3/ Finalement, le test d'une variable endogène décalée a été proposé par Anselin (1988b). Soit l'hypothèse nulle $H_0 : \rho = 0$ dans [2.2] en utilisant la fonction de vraisemblance de [3.10].

On trouve la statistique de test suivante :

$$LM_{LAG} = \frac{[\tilde{\varepsilon}' W y / \tilde{\sigma}^2]^2}{\tilde{T}_1} \quad [4.9]$$

avec $\tilde{T}_1 = \left[(WX\tilde{\beta})' (I - X(X'X)^{-1}X') (WX\tilde{\beta}) + T\tilde{\sigma}^2 \right] / \tilde{\sigma}^2$. Sous $H_0 : LM_{LAG} \xrightarrow{D} \chi_1^2$.

413. 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. Par exemple, Anselin et Bera (1998) remarquent que LM_{ERR} est la statistique de test correspondant à $H_0 : \lambda = 0$ en supposant que $\rho = 0$. En revanche, si $\rho \neq 0$, ce test n'est plus valide, même asymptotiquement et il n'est plus distribué selon un χ^2 central à 1 degré de liberté. Pour une inférence statistique valide, il est donc nécessaire de prendre en compte la possible variable endogène décalée lorsqu'on teste l'autocorrélation spatiale des erreurs et vice-versa. Face à ce problème, plusieurs stratégies sont possibles. On peut déjà effectuer un test joint de présence d'une variable décalée et d'une autocorrélation des erreurs mais si l'hypothèse nulle est rejetée, on ne connaît pas la nature exacte de la dépendance spatiale. Une autre solution consiste alors à estimer un modèle avec une variable endogène décalée et tester ensuite s'il y a encore une autocorrélation des erreurs et vice-versa (Anselin, 1988b). Dans ce cas, il faut estimer les modèles par le maximum de vraisemblance. Anselin et al. (1996) ont finalement proposé des tests basés sur les résidus des MCO dans le modèle simple (2.1) mais qui sont capables de prendre en compte une autocorrélation des erreurs lorsqu'on teste la présence d'une variable endogène décalée et vice-versa.

1/ *La première approche* consiste à tester l'hypothèse nulle jointe $H_0 : \lambda = \rho = 0$ dans le modèle [2.17] grâce au principe du multiplicateur de Lagrange. Ainsi, le test peut être effectué à partir des résidus des MCO dans le modèle simple [2.1]. La statistique qui en résulte est la suivante :

$$SARMA = \tilde{E}^{-1} \left[(\tilde{d}_\lambda)^2 \frac{\tilde{D}}{\tilde{\sigma}^2} + (\tilde{d}_\rho)^2 T_{22} - 2\tilde{d}_\lambda \tilde{d}_\rho T_{12} \right] \quad [4.10]$$

où \tilde{d}_λ et \tilde{d}_ρ sont respectivement les scores par rapport à λ et ρ évalués sous l'hypothèse nulle, $T_{ij} = \text{tr}[W_i W_j + W_i' W_j']$, $D = (W_1 X \beta)' M (W_1 X \beta) + T_{11} \sigma^2$, $E = (D / \sigma^2) T_{22} - (T_{12})^2$. Si $W_1 = W_2 = W$ alors $T_{11} = T_{21} = T_{22} = T = \text{tr}[(W' + W)]$ et [4.10] se simplifie à :

$$SARMA = \frac{\tilde{d}_\lambda^2}{T} + \frac{(\tilde{d}_\lambda - \tilde{d}_\rho)^2}{\tilde{\sigma}^{-2}(\tilde{D} - T\tilde{\sigma}^2)} \quad [4.11]$$

Sous $H_0 : \lambda = \rho = 0$, $SARMA$ converge vers un χ^2 à deux degrés de libertés.

2/ *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, cela consiste à tester l'hypothèse nulle $H_0 : \lambda = 0$ en présence de ρ . Sous l'hypothèse nulle, on retrouve le modèle autorégressif [2.2] alors que sous l'hypothèse alternative, on retrouve le

modèle général [2.17]. Le test est alors basé sur les résidus de l'estimation par le maximum de vraisemblance dans le modèle autorégressif [2.2]. La statistique est :

$$LM_{ERR}^* = \frac{\hat{d}_\rho^2}{T_{22} - (T_{21A})^2 \hat{V}(\hat{\rho})} \quad [4.12]$$

où $T_{21A} = \text{tr}[W_2 W_1 A^{-1} + W_2' W_1' A^{-1}]$, $A = I - \hat{\rho} W_1$ et où le chapeau désigne les estimateurs évalués par le maximum de vraisemblance dans le modèle [2.2] obtenus par optimisation non linéaire, $\hat{V}(\hat{\rho})$ est la variance de $\hat{\rho}$ dans le modèle [2.2]. Sous $H_0 : \lambda = 0$, $LM_{ERR}^* \xrightarrow{D} \chi_1^2$.

On peut également tester l'hypothèse nulle $H_0 : \rho = 0$ en présence de λ , le test est alors basé sur les résidus de l'estimation par le maximum de vraisemblance dans le modèle avec autocorrélation des erreurs [2.7]. La statistique est :

$$LM_{LAG}^* = \frac{[\hat{\varepsilon}' B' B W_1 y]^2}{H_\rho - H_{\theta\rho} \hat{V}(\hat{\theta}) H_{\theta\rho}'} \quad [4.13]$$

où $\hat{\varepsilon}$ est le vecteur des résidus estimés par le maximum de vraisemblance dans le modèle avec erreurs autorégressives [2.7], $\theta = (\beta', \lambda, \sigma^2)$, $B = I - \hat{\lambda} W_2$. Les autres termes sont :

$$H_\rho = \text{tr} W_1^2 + \text{tr}(B W_1 B^{-1})' (B W_1 B^{-1}) + \frac{(B W_1 X \beta)' (B W_1 X \beta)}{\sigma^2} \quad [4.14]$$

$$\text{et : } H_{\theta\rho}' = \begin{bmatrix} \frac{(B X)' B W_1 X \beta}{\sigma^2} \\ \text{tr}(W_2 B^{-1})' B W_1 B^{-1} + \text{tr} W_2' W_1 B^{-1} \\ 0 \end{bmatrix} \quad [4.15]$$

et $\hat{V}(\hat{\theta})$ est la matrice des variances-covariances estimée de $\hat{\theta}$ dans le modèle [2.7]. Sous H_0 : LM_{LAG}^* converge vers un χ^2 à un degré de libertés.

Ces tests nécessitent donc une estimation par le maximum de vraisemblance. Une dernière approche de ces tests nécessite uniquement une estimation du modèle simple [2.1] par les MCO.

3/ *La dernière approche* est celle de Bera et Yoon (1993) qui a été reprise par Anselin et al. (1996). Elle consiste à utiliser des tests robustes à une mauvaise spécification locale. Par exemple, on ajuste LM_{ERR} pour que sa distribution asymptotique reste un χ^2 central, même en présence *locale* de ρ . Ce test s'effectue à partir des résidus des MCO du modèle simple [2.1]. La statistique modifiée pour le test de $H_0 : \lambda = 0$ est :

$$RLM_{ERR} = \frac{[\tilde{d}_\lambda - T_{12} \tilde{\sigma}^2 \tilde{D}^{-1} \tilde{d}_\rho]^2}{T_{22} - (T_{12})^2 \tilde{\sigma}^2 \tilde{D}} \Rightarrow RLM_{ERR} = \frac{[\tilde{d}_\lambda - T \tilde{\sigma}^2 \tilde{D}^{-1} \tilde{d}_\rho]^2}{T(1 - T \tilde{\sigma}^2 \tilde{D})} \text{ si } W_1 = W_2 = W \quad [4.16]$$

De la même façon, les résidus du modèle [2.1] servent au test de $H_0 : \rho = 0$ dans la présence locale de λ . La statistique de ce test est :

$$RLMLAG = \frac{[\tilde{d}_\rho - T_{12}T_{22}^{-1}\tilde{d}_\lambda]^2}{\tilde{\sigma}^{-2}\tilde{D} - (T_{12})^2T_{22}^{-1}} \Rightarrow RLMLAG = \frac{[\tilde{d}_\rho - \tilde{d}_\lambda]^2}{\tilde{\sigma}^{-2}\tilde{D} - T} \quad \text{si } W_1 = W_2 = W \quad [4.17]$$

414. 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 (1995) fournissent quelques indications quant aux performances de ces tests asymptotiques en échantillon fini.

Anselin et Rey (1991) et Anselin et Florax (1995) ont comparé les performances des tests de Moran, $LMERR$, $LMLAG$, $RLMERR$ et $RLMLAG$ pour différentes matrices de poids (sur zonage régulier ou non), différentes distributions des erreurs et différentes tailles d'échantillon. Plusieurs résultats ressortent.

- D'une manière générale, les puissances des tests diminuent dans les petits échantillons et augmentent avec des valeurs plus fortes des paramètres spatiaux, $LMLAG$ étant le test le plus puissant et robuste à la non-normalité des erreurs.
- Le test I de Moran apparaît 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.
- Les tests du score $LMERR$ et $LMLAG$ ont les plus grandes puissances pour leur alternative respective.
- Les tests ajustés $RLMERR$ et $RLMLAG$ ont également de bonnes performances en termes de puissance et de taille empiriques, le « prix à payer » pour la correction est faible ($RLMLAG$ reste plus performant que $RLMERR$).

Florax et Rey (1995) ont étudié les conséquences d'une mauvaise spécification de la matrice de poids sur la puissance des tests précédents. Ils distinguent entre une « sur-spécification » et une « sous-spécification ». Dans le premier cas, 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 deuxième cas, des liens spatiaux sont omis à tort. Les simulations effectuées par ces auteurs suggèrent qu'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. Le coefficient de Moran est moins affecté que les autres tests par une mauvaise spécification de la matrice de poids.

42. A la recherche de la spécification du modèle

Dans le paragraphe 41 sont développés des tests de spécification permettant de détecter une omission de l'autocorrélation spatiale et la forme prise par cette dernière dans le modèle. D'autres tests de spécification ont pour but de tester la présence de variables exogènes décalées (paragraphe 421) et de déterminer la structure de la dépendance spatiale, reflétée par le choix de la matrice de poids (paragraphe 422). Tous ces tests peuvent alors servir à établir des règles de décision permettant de rechercher la meilleure spécification du modèle (paragraphe 423) et une étude de la convergence entre les régions européennes fournit une illustration de ces techniques (paragraphe 424).

421. Le test du facteur commun

Le test du facteur commun permet de choisir entre un modèle avec autocorrélation des erreurs et un modèle avec l'ensemble des variables explicatives décalées. Rappelons qu'en séries temporelles, l'approche en termes de facteur commun est basée sur l'équivalence de deux spécifications de modèles, l'une exprimée en termes des erreurs et l'autre en termes de variables endogènes retardées. Si certaines contraintes sont respectées sur les coefficients du dernier modèle, la spécification se réduit à une forme plus simple avec des erreurs autocorrélées. En effet, soit le modèle comportant une variable endogène décalée :

$$y_t = \alpha y_{t-1} + \beta_1 x_t + \beta_2 x_{t-1} + u_t \quad [4.18]$$

où les u_t sont indépendants de X et des valeurs passées de u , alors, si $\beta_2 = -\beta_1 \alpha$, le modèle devient un modèle avec autocorrélation des erreurs :

$$\begin{aligned} y_t &= \beta_1 x_t + v_t \\ v_t &= \alpha v_{t-1} + u_t \end{aligned} \quad [4.19]$$

En économétrie spatiale, l'approche est similaire et prend pour point de départ la formulation de Durbin (Burrige, 1980 ; Bivand, 1984). Nous avons vu dans le paragraphe 221 que les deux modèles suivants étaient équivalents :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \end{aligned} \quad [4.20]$$

$$\Leftrightarrow y = \lambda Wy + X\beta - \lambda WX\beta + u \quad [4.21]$$

Le modèle (4.21) est estimé par :

$$y = \lambda Wy + X\beta + WX\delta + u \quad [4.22]$$

Par conséquent, pour savoir si le modèle [4.22] peut se réduire au modèle [4.20], il faut tester l'hypothèse suivante : $H_0 : \lambda\beta + \delta = 0$.

Au contraire des séries temporelles, un modèle avec erreurs autocorrélées 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 du maximum de vraisemblance. Le seul avantage de la forme [4.20] sur la forme [4.22] est qu'il y a moins de paramètres à estimer : $K + 1$ au lieu de $2K + 1$.

Le test sur les contraintes des paramètres se fait avec l'un des trois tests traditionnels : test de Wald, test du rapport de vraisemblance ou test du multiplicateur de Lagrange. Tous trois sont distribués selon une loi du χ^2 à $K-1$ degrés de liberté (en ignorant le terme constant). Par exemple, le test de Wald est de la forme :

$$W = g'[G'VG]^{-1}g$$

où : $g = \lambda\beta + \delta$; $G = \partial g / \partial \theta$; $\theta = [\lambda \beta \delta]$; V est la matrice estimée des variances-covariances du modèle [4.22].

Le test du facteur commun part donc d'un modèle général sur lequel des tests sont faits pour voir si un modèle plus simple est plus pertinent ou non. Cette manière de faire peut

constituer un désavantage lorsqu'il y a peu d'observations : les degrés de liberté risquent de devenir trop faibles. Ce test a été étendu à plusieurs décalages spatiaux et retards temporels par Blommestein (1983) et Blommestein et Nijkamp (1986).

22. La détermination de la structure spatiale

Lorsque deux modèles concurrents sont formulés comme des hypothèses alternatives et lorsqu'un des modèles ne peut pas être exprimé comme un cas particulier de l'autre modèle (comme c'était le cas jusqu'à présent), les modèles non-emboîtés nécessitent des procédures spéciales de tests. Par exemple, deux modèles autorégressifs peuvent être en compétition (Anselin, 1984) :

$$H_0 : y = \rho_0 W_0 y + X_0 \beta_0 + \varepsilon_0 \quad [4.24a]$$

$$H_1 : y = \rho_1 W_1 y + X_1 \beta_1 + \varepsilon_1 \quad [4.24b]$$

Dans ces expressions [4.24a) et [4.24b), les matrices de poids, les variables explicatives sont différentes et les erreurs n'ont pas forcément la même distribution. Au contraire des tests précédents, l'hypothèse alternative ne peut pas être considérée comme un modèle contraint par rapport à l'hypothèse nulle. Il faut donc trouver des procédures de tests permettant de tester ces hypothèses non-emboîtées. Différentes procédures ont été listées par Anselin (1984) dont le test J .

Le test J a été proposé par Davidson et MacKinnon (1981) dans un cadre aspatial. Deux modèles sont en compétition :

$$H_0 : y = X\beta + \varepsilon \quad [4.25a]$$

$$H_1 : y = Z\gamma + u \quad [4.25b]$$

y est le vecteur de la variable expliquée de dimension $(N,1)$. X et Z sont des matrices de variables explicatives de dimensions respectives (N,k_1) et (N,k_2) . β et γ sont des vecteurs de paramètres de dimensions respectives $(k_1,1)$ et $(k_2,1)$. ε et u sont des vecteurs d'erreurs de dimension $(N,1)$ avec :

$$E(\varepsilon\varepsilon') = \sigma_\varepsilon^2 \quad [4.26a]$$

$$E(uu') = \sigma_u^2 \quad [4.26b]$$

L'idée est de construire un modèle enveloppe qui contient les deux modèles précédents pour pouvoir tester l'un ou l'autre des modèles de départ contre le modèle reconstruit suivant :

$$y = (1 - \alpha)X\beta + \alpha Z\gamma + v \quad [4.27]$$

α est un paramètre inconnu. On voit que si $\alpha = 0$, on obtient le modèle [4.25a) et on rejette l'hypothèse alternative. En revanche, si $\alpha = 1$, on obtient le modèle [4.25b) et on rejette l'hypothèse nulle. L'idée est donc de développer le test : $H_0 : \alpha = 0$. Pourtant, les paramètres α , β et γ du modèle [4.27) ne sont pas identifiables. Par conséquent, la procédure de test va s'effectuer en deux étapes. Premièrement, γ est estimé par les MCO dans le modèle [4.25b). On en déduit l'estimation $\hat{\gamma}$. Deuxièmement, on régresse par les MCO y sur X et $Z\hat{\gamma}$:

$$y = X\beta + \alpha Z\hat{\gamma} + u \quad [4.28]$$

On teste alors $H_0 : \alpha = 0$ et si l'hypothèse nulle n'est pas rejetée, c'est le modèle [4.25a] qui est acceptable. Le ratio du coefficient $\hat{\alpha}$ à son écart-type associée suit asymptotiquement une loi normale d'espérance nulle et de variance unitaire.

Anselin (1986) a montré que les résultats précédents pouvaient inclure des variables dépendantes décalées sous des conditions assez générales : une variance bornée (ce qui est vérifié lorsque le paramètre spatial est inclus dans son espace des paramètres) et une structure de dépendance spatiale qui décroît lorsque la distance entre les observations augmente (ce qui est vérifié avec le choix d'une matrice de poids appropriée).

423. Les règles de décision

Les différents tests de spécification peuvent être combinés afin de choisir la meilleure spécification du modèle (tableau 1).

1/ La première étape consiste à estimer le modèle simple [2.1] par les MCO et effectuer le test de Moran [4.1] et le test SARMA [4.11] (test joint d'une présence de présence d'autocorrélation des erreurs et d'une variable autorégressive). Le rejet de l'hypothèse nulle indique une mauvaise spécification du modèle et une omission de l'autocorrélation spatiale. Les tests LM_{ERR} [4.5], LM_{LAG} [4.9] et leurs versions robustes [4.16] et [4.17] permettent de spécifier la forme de l'autocorrélation spatiale (voir 3/)

2/ Si les résultats des tests indiquent une 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 (si cette dernière provient d'une mauvaise spécification) ou des variables exogènes décalées spatialement (Florax et Folmer, 1992).

3/ 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 en comparant les niveaux de significativité des tests du multiplicateur de Lagrange s'effectue selon les valeurs relatives des tests du multiplicateur de Lagrange LM_{ERR} [4.5], LM_{LAG} [4.9] et leurs versions robustes [4.16] et [4.17].

* Si l'on ne considère que LM_{ERR} et LM_{LAG} , Anselin et Rey (1991) et Florax et Folmer (1992) proposent de choisir l'une ou l'autre forme fonctionnelle [2.2] ou [2.7] en appliquant la règle de décision simple suivante :

- Si le test du modèle autorégressif n'aboutit pas au rejet de l'hypothèse nulle alors que le test de l'autocorrélation des erreurs rejette l'hypothèse nulle, ou si les deux tests aboutissent au rejet de l'hypothèse nulle et que le deuxième test est plus significatif que le premier, on choisit le modèle avec autocorrélation des erreurs.

- Si le test du modèle autorégressif aboutit au rejet de l'hypothèse nulle, ou si les deux tests aboutissent au rejet de l'hypothèse nulle et que le premier test est plus significatif que le test de l'autocorrélation des erreurs, on choisit le modèle autorégressif.

Les simulations de Monte-Carlo effectuées par Florax et Folmer (1992) indiquent que la probabilité de trouver le vrai modèle à l'aide de cette stratégie est plus élevée si le vrai modèle est un modèle autorégressif et non un modèle où les erreurs sont autocorrélées. Il est aussi à noter que les probabilités sont plus élevées lorsque le paramètre spatial est positif.

* *Anselin et Florax (1995) affinent cette règle de décision en la complétant par l'utilisation des tests robustes : si LM_{LAG} est plus significatif que LM_{ERR} et RLM_{LAG} est significatif mais pas RLM_{ERR} , on inclut une variable endogène décalée. D'une façon similaire, la présence d'une autocorrélation des erreurs peut être identifiée à travers RLM_{ERR} . Dans ce dernier cas, le test du facteur commun doit encore être effectué pour vérifier que le modèle avec autocorrélation des erreurs est le meilleur.*

4/ Une fois que le modèle spatial adéquat a été estimé ([2.2] ou [2.7]), trois tests supplémentaires peuvent être mobilisés.

* Pour un modèle autorégressif [2.2], le test LM_{ERR}^* permet de savoir si une autocorrélation spatiale des erreurs est encore nécessaire.

* Pour un modèle avec autocorrélation des erreurs [2.7], le test LM_{LAG}^* permet de savoir si une variable endogène décalée est encore nécessaire. Le test du facteur commun (voir section 421) indique si la restriction $\delta + \lambda\beta = 0$ peut être rejetée ou non. Si elle ne l'est pas, le modèle [4.22] se réduit au modèle avec autocorrélation des erreurs [4.20].

5/ Si plusieurs modèles restent encore en compétition, le test J sert à comparer des modèles spatiaux comportant des matrices de poids différentes. Le choix entre modèles peut aussi se faire avec les critères traditionnels tels que les critères d'information :

$$INF = -2\ln L + q(K) \quad [4.29]$$

$\ln L$ 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 $q = \log NK$ pour le critère de Schwartz. Lorsqu'on compare deux modèles selon leur critère d'information, on choisit celui qui minimise ce coefficient.

424. Illustration : Convergence des régions européennes

Aujourd'hui, les logiciels économétriques les plus courants n'intègrent pas spécifiquement des options permettant d'estimer les modèles spatiaux et d'effectuer les différents tests. Ceci contraste avec le grand nombre de logiciels disponibles pour l'analyse de données spatiales dans les sciences physiques, avec une attention particulière accordée aux données géostatistiques. Par exemple, il y a une librairie pour GSLIB (Deutsch et Journel, 1992) et un module S+SpatialStat pour S+plus (Mathsoft, 1996). Ce module contient quelques-unes des procédures exposées ici. Il existe en revanche un logiciel économétrique destiné spécifiquement à l'analyse des données spatiales : SpaceStat (Anselin, 1999), qui contient les méthodes d'estimation par le maximum de vraisemblance, par les variables instrumentales, ou la méthode GMM pour les modèles spatiaux et hétéroscédastiques. Il contient également un grand nombre de tests de spécification et de nombreuses fonctions destinées à l'analyse exploratoire des données spatiales.

Dans Baumont et al. (2000b), le logiciel SpaceStat a été utilisé afin d'appliquer les techniques de l'économétrie spatiale à l'étude de la convergence des régions européennes. L'hypothèse de convergence est basée sur les modèles de croissance néoclassiques et implique qu'une région « pauvre » tend à croître plus rapidement qu'une région « riche », de telle sorte que la région « pauvre » rattrape à long terme le niveau de revenu ou de production par tête de la région « riche ». Cette propriété correspond au concept de β -convergence (Barro and Sala-I-Martin, 1995). La β -convergence est absolue lorsqu'elle est indépendante des

conditions initiales et cette hypothèse est habituellement testée sur le modèle en coupes transversales suivantes :

$$\frac{1}{T} \ln \left(\frac{y_{i,T}}{y_{i,0}} \right) = \alpha + \beta \ln(y_{i,0}) + \varepsilon_i \quad \varepsilon_i \sim iid(0, \sigma_\varepsilon^2) \quad [4.30]$$

où $y_{i,t}$ est le produit par tête de la région i ($i = 1, \dots, N$) à la date t , T est la longueur de la période, α et β sont les paramètres inconnus et ε_i le terme d'erreur. Il y a β -convergence lorsque β est négatif et statistiquement significatif puisque dans ce cas le taux de croissance moyen entre les dates 0 et T est négativement corrélé avec le niveau initial du produit par tête.

Pour détecter une éventuelle autocorrélation spatiale des observations, divers modèles économétriques spatiaux ont été estimés sur un échantillon comprenant les PIB par tête de 122 régions européennes pour la période 1980-1995⁶.

Les résultats de l'estimation par les MCO de [4.30] et des différents tests sont présentés dans le tableau 2. Le coefficient associé au niveau du PIB initial est significatif et négatif, ce qui confirme l'hypothèse de convergence pour les régions européennes. Les résultats des tests appellent plusieurs remarques. Tout d'abord, le test de Jarque-Bera ne rejette pas l'hypothèse de normalité. Par conséquent, la validité de l'estimation par le maximum de vraisemblance et le calcul des tests du multiplicateur de Lagrange sont assurés. Ensuite, le test de Moran et le test SARMA conduisent au rejet de l'hypothèse nulle d'absence d'autocorrélation spatiale. Le modèle est donc mal spécifié. Enfin, pour déterminer la forme qui doit être prise par l'autocorrélation spatiale, on étudie la significativité des tests du multiplicateur de Lagrange : LM_{ERR} , LM_{LAG} , RLM_{ERR} et RLM_{LAG} . LM_{ERR} et LM_{LAG} sont tous les deux significatifs mais RLM_{ERR} est significatif alors que RLM_{LAG} ne l'est pas. Ces tests indiquent donc la présence d'une autocorrélation spatiale des erreurs plutôt qu'une variable endogène décalée. Pour confirmer ce résultat, les différents modèles économétriques spatiaux sont estimés.

Pour vérifier que l'autocorrélation spatiale des erreurs n'est pas le résultat d'une variable spatiale décalée omise, le modèle régressif croisé a été estimé par les MCO. Le coefficient associé à la variable mesurant l'influence des PIB initiaux des régions voisines n'est pas significatif. De plus, les tests indiquent la présence d'une variable endogène décalée omise et ce modèle est inférieur au modèle précédent en termes des critères d'information. Par conséquent, il est nécessaire d'estimer un modèle incorporant explicitement l'autocorrélation spatiale.

Dans le modèle avec variable endogène décalée, tous les coefficients sont significatifs. L'estimation du paramètre autorégressif $\hat{\rho}$ est hautement significatif et positif ($\hat{\rho} = 0,635$) et indique un effet de débordement géographique : la croissance du PIB par tête est influencée par celle des régions voisines. Le test LM_{ERR}^* indique qu'une autocorrélation des erreurs supplémentaire n'est pas nécessaire. Ce modèle est meilleur que les modèles précédents en termes des critères d'information. On remarque également que la variance estimée a diminué.

Les résultats de l'estimation du modèle avec erreurs autocorrélées font apparaître à nouveau des coefficients tous significatifs : le coefficient associé au niveau du PIB initial est plus grand que dans le modèle [4.30] et l'autocorrélation spatiale des erreurs est positive ($\hat{\lambda} = 0,702$). Ce modèle est correctement spécifié. En effet, le test LM_{LAG}^* indique qu'une variable endogène supplémentaire dans ce modèle n'est pas nécessaire. De plus, le test du

⁶ Les données sont tirées de la base de données EUROSTAT-REGIO. L'échantillon est composé de régions au niveau NUTS1 (Danemark, Luxembourg, Royaume-Unis) et au niveau NUTS2 (Belgique, Espagne, France, Allemagne, Italie, Pays-Bas et Portugal).

facteur commun indique que la restriction $\delta + \lambda\beta = 0$ ne peut pas être rejetée dans le modèle [4.22] : le modèle [4.22] se réduit donc bien au modèle avec autocorrélation des erreurs [4.20]. Ce modèle est le meilleur en termes des critères d'information et la variance estimée est la plus basse. Il apparaît donc bien que le modèle avec autocorrélation spatiale des erreurs est la spécification la plus adéquate.

CONCLUSION

L'objectif de cet article était d'examiner la façon dont l'autocorrélation spatiale pouvait être introduite dans les modèles économétriques. L'autocorrélation spatiale se réfère à la dépendance des observations provenant de la disposition géographique des données. Elle se modélise grâce aux matrices de poids et différents coefficients sont destinés à détecter la présence de l'autocorrélation spatiale dans une série. 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 celle des moments généralisés. Les tests de spécification permettent enfin de déterminer la forme prise par l'autocorrélation spatiale.

Longtemps ignorées dans les articles contenant des applications empiriques, les techniques de l'économétrie spatiale sont de plus en plus appliquées dans la littérature, spécialisée ou non. Les développements les plus nouveaux, l'incorporation des effets spatiaux dans les modèles à variables qualitatives et les modèles spatio-temporels, sont en revanche encore relativement peu utilisés car beaucoup de progrès relatifs à la théorie économétrique dans ces deux domaines restent encore à faire, au contraire des modèles linéaires à variables quantitatives autour desquels un consensus s'est établi. Les principales avancées à faire encore en économétrie spatiale sont donc la systématisation de l'introduction des effets spatiaux dans les modèles à variables qualitatives et les modèles spatio-temporels.

BIBLIOGRAPHIE

- Amemiya T., 1985, *Advanced Econometrics*, Harvard University Press, Cambridge.
- Ancot J.P., Kuiper J.H., Paelinck J.H.P., 1983, Econométrie spatiale : une synthèse décennale, in J.H.P. Paelinck et A. Salles (Eds.), *Espace et Localisation*, Economica, Paris.
- Anselin L., 1980, *Estimation Methods for Spatial Autoregressive Structures*, Cornell University, Regional Science Dissertation and Monograph Series #8, Ithaca, NY.
- Anselin L., 1982, A note on small sample properties of estimators in a first-order spatial autoregressive model, *Environment and Planning A*, 14, 1023-1030.
- Anselin L., 1984, Specification tests on the structure of interaction in spatial econometric models, *Papers of the Regional Science Association*, 54, 165-182.
- Anselin L., 1986, Non-tested tests on the weight structure in spatial autoregressive models : some Monte-Carlo results, *Journal of Regional Science*, 26, 267-284.
- Anselin L., 1988a, *Spatial Econometrics : Methods and Models*, Kluwer Academic Publishers, Dordrecht.
- Anselin L., 1988b, Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity, *Geographical Analysis*, 20, 1-17.
- Anselin L., 1988c, A test for spatial autocorrelation in seemingly unrelated regressions, *Economics Letters*, 28, 335-341.
- Anselin L., 1998, Rao's score test in spatial econometrics, Working Paper, Bruton Center, School of Social Sciences, University of Texas at Dallas.
- Anselin L., 1999, *SpaceStat, a Software Program for the Analysis of Spatial Data*, Version 1.90. Bio Medware, Ann Arbor.
- Anselin L., Bera A., 1998, Spatial dependence in linear regression models with an application to spatial econometrics, in A. Ullah et D.E.A. Giles (Eds.), *Handbook of Applied Economics Statistics*, Springer-Verlag, Berlin, 21-74.
- Anselin L., Bera A.K., Florax R., Yoon M.J., 1996, Simple diagnostic tests for spatial dependence, *Regional Science and Urban Economics*, 26, 77-104.
- Anselin L., Can A., 1986, Model comparison and model validation issues in empirical work on urban density models, *Geographical Analysis*, 18, 179-197.
- Anselin L., Florax R., 1995, Small sample properties of tests for spatial dependence in regression models, in L. Anselin et R. Florax (Eds.), *New Directions in Spatial Econometrics*, Springer, Berlin, 21-74.
- Anselin L., Hudak S., 1992, Spatial econometrics in practice, a review of software options, *Regional Science and Urban Economics*, 22, 509-536.
- Anselin L., Kelejian H., 1997, Testing for spatial error autocorrelation in the presence of endogenous regressors, *International Regional Science Review*, 20, 153-182.
- Anselin L., Rey S., 1991, Properties of tests for spatial dependence in linear regression models, *Geographical Analysis*, 23, 112-131.
- Anselin L., Smirnov O., 1996, Efficient algorithms for constructing proper higher order spatial lag operators, *Journal of Regional Science*, 36, 67-89.
- Anselin L., Smirnov O., 1999, Fast maximum likelihood estimation of very large spatial autoregressive models : a characteristic polynomial approach, Working Paper, Bruton Center, School of Social Sciences, University of Texas at Dallas.
- Anselin L., Varga A., Acs Z., 1997, Local geographic spillovers between university research and high technology innovations, *Journal of Urban Economics*, 42, 422-448.
- Anselin L., Varga A., Acs Z., 1998, Geographic and sectoral characteristics of academic knowledge externalities, Working Paper, Bruton Center, School of Social Sciences, University of Texas at Dallas.
- Arbia G., 1989, *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems*, Kluwer Academic Publishers, Boston.
- Arora S., Brown M., 1977, Alternatives approaches to spatial autocorrelation : an improvement over current practice, *International Regional Science Review*, 2, 67-78.
- Aten B., 1996, Evidence of spatial autocorrelation in international prices, *Review of Income and Wealth*, 42, 149-163.
- Aten B., 1997, Does space matter? International comparisons of the prices of tradables and nontradables, *International Regional Science Review*, 20, 35-52.
- Avery R.B., Hansen L.P., Hotz V.J., 1983, Multiperiod probit models and orthogonality condition estimation, *International Economic Review*, 24, 21-35.
- Bailey T.C., Gatrell A.C., 1995, *Interactive Spatial Data Analysis*, Longman.

- Barro R.J., Sala-i-Martin X., 1995, *Economic Growth Theory*, MIT Press.
- Barry R. P., Pace R. K., 1999, Monte-Carlo estimates of the log-determinant of large sparse matrices, *Linear Algebra and its Applications*, 289, 42-54.
- Baumont C., Ertur C., Le Gallo J., 2000a, Convergence des régions européennes : une approche par l'économétrie spatiale, Working Paper n° 2000-03, LATEC, Université de Bourgogne.
- Baumont C., Ertur C., Le Gallo J., 2000b, Geographic spillover and growth. A spatial econometric analysis for european regions, 6th RSAI World Congress 2000, Lugano, Suisse, 16-20 Mai 2000.
- Bavaud F., 1998, Models for spatial weights: a systematic look, *Geographical Analysis*, 30, 153-171.
- Benirschka A., Binkley J., 1996, Land price volatility in a geographically dispersed market, *American Journal of Agricultural Economics*, 76, 185-195.
- Bera A.K., Yoon M., 1993, Specification testing with locally misspecified alternatives, *Econometric Theory*, 9, 649-658.
- Bernat G.A., 1996, Does manufacturing matter? A spatial econometric view of Kaldor's laws, *Journal of Regional Science*, 463-477.
- Bivand R.S., 1984, Regression modeling with spatial dependence : an application of some class selection and estimation methods, *Geographical Analysis*, 16, 25-37.
- Bivand R.S., 1992, Systat compatible software for modeling spatial dependence among observations, *Computers and Geoscience*, 18, 951-963.
- Blommestein H.J., 1983, Specification and estimation of spatial econometric models, *Regional Science and Urban Economics*, 13, 251-270.
- Blommestein H.J., 1985, Elimination of circular routes in spatial dynamic regression equations, *Regional Science and Urban Economics*, 15, 121-130.
- Blommestein H.J., Koper N.A.M., 1992, Recursive algorithms for the elimination of redundant paths in spatial lag operators, *Journal of Regional Science*, 32, 91-111.
- Blommestein H.J., Nijkamp P., 1986, Testing the spatial scale and the dynamic structure in regional models (a contribution to spatial econometric specification analysis), *Journal of Regional Science*, 26, 1-17.
- Boarnet M.G., 1998, Spillovers and the locational effects of public infrastructure, *Journal of Regional Science*, 38, 381-400.
- Brandsma A.S., Kelletaper R.H., 1979, A biparametric approach to spatial autocorrelation, *Environment and Planning A*, 11, 51-58.
- Breusch T., 1980, Useful invariance results for generalized regression models, *Journal of Econometrics*, 13, 327-340.
- Brueckner J.K., 1998, Testing for strategic interaction among local governments: The case of growth controls, *Journal of Urban Economics*, 44, 438-467.
- Burrige P., 1980, On the Cliff-Ord test for spatial autocorrelation among regression residuals, *Geographical Analysis*, 4, 267-284.
- Burrige P., 1981, Testing for a common factor in a spatial autoregressive model, *Environment and Planning A*, 13, 795-800.
- Can A. et Megboluge I., 1997, Spatial dependence and house price index construction, *Journal of Real Estate Finance and Economics*, 14, 203-222.
- Can A., 1990, The measurement of neighborhood dynamics in urban house prices, *Economic Geography*, 66, 254-272.
- Can A., 1992, Specification and estimation of hedonic housing price models, *Regional Science and Urban Economics*, 22, 453-474.
- Case A., 1991, Spatial patterns in household demand, *Econometrica*, 59, 953-965.
- Case A., 1992, Neighborhood influence and technological change, *Regional Science and Urban Economics*, 22, 491-508.
- Case A.C., Rosen H.S., Hines J.R., 1993, Budget spillovers and fiscal policy interdependence: Evidence from the States, *Journal of Public Economics*, 52, 285-307.
- Chou Y.H., 1991, Map resolution and spatial autocorrelation, *Geographical Analysis*, 23, 228-246.
- Cliff A.D., Ord J.K., 1972, Testing for spatial autocorrelation among regression residuals, *Geographical Analysis*, 4, 267-284.
- Cliff A.D., Ord J.K., 1973, *Spatial Autocorrelation*, Pion, Londres.
- Cliff A.D., Ord J.K., 1981, *Spatial Processes : Models and Applications*, Pion, Londres.
- Cressie N., 1993, *Statistics for Spatial Data*, John Wiley, New York.
- Davidson R., Mackinnon J.G., 1981, Several tests for model specification in the presence of alternative hypotheses, *Econometrica*, 49, 781-793.

- Deutsch C., Journel A., 1992, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, Oxford.
- Dubin R., Pace R.K., Thibodeau T.G., Spatial autoregression techniques for real estate data, *Journal of Real Estate Techniques*, 7, 79-95.
- Dubin R.A., 1995, Estimating logit models with spatial dependence in L Anselin et R. Florax (Eds.), *New Directions in Spatial Econometrics*, Springer, Berlin, 229-242.
- Dubin R.A., 1997, A note on the estimation of spatial logit models, *Geographical Systems*, 4, 181-193.
- Dubin R.A., Spatial autocorrelation: a primer, *Journal of Housing Economics*, 7, 304-327.
- Fingleton B., 1999, Estimates of time to convergence: An analysis of regions of european union, *International Regional Science Review*, 22, 5-34.
- Florax R., 1992, *The University: A Regional Booster? Economic Impacts of Academic Knowledge Infrastructure*, Avebury, Aldershot.
- Florax R., Folmer H., 1992, Specification and estimation of spatial linear regression models, *Regional Science and Urban Economics*, 22, 405-432.
- Florax R., Rey S., 1995, The impacts of misspecified spatial interaction in linear regression models, in L Anselin et R. Florax (Eds.), *New Directions in Spatial Econometrics*, Springer, Berlin, 111-135.
- Fotheringham A.S., Brundson C., Charlton M., *Quantitative Geography. Perspectives on spatial data analysis*, Sage Publications, Londres.
- Getis A., 1990, Screening for spatial dependence in regression analysis, *Papers of the Regional Science Association*, 69, 69-81.
- Getis A., 1995, Spatial filtering in a regression framework : examples using data on urban crime, regional inequality, and government expenditures, in L. Anselin, R. Florax (Eds.), *New Directions in Spatial Econometrics*, Springer, Berlin.
- Getis A., Ord J.K., 1992, The analysis of spatial association by use of distance statistics, *Geographical Analysis*, 24, 189-206.
- Getis A., Ord J.K., 1995, Local spatial autocorrelation statistics : distributional issues and an application, *Geographical Analysis*, 27, 286-306.
- Griffith D.A., 1981, Modelling urban population density in a multi-centered city, *Journal of Urban Economics*, 9, 298-310.
- Griffith D.A., 1988a, *Advanced Spatial Statistics, Special Topics in the Exploration of Quantitative Spatial Data Series*, Kluwer Academic Publishers, Dordrecht.
- Griffith D.A., 1988b, Estimating spatial autoregressive model parameters with commercial statistical packages, *Geographical Analysis*, 20, 176-186.
- Griffith D.A., 1992a, Simplifying the normalizing factor in spatial autoregressions for irregular lattices, *Papers in Regional Science*, 71, 71-86.
- Griffith D.A., 1992b, What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics, *L'Espace Géographique*, 21, 265-280.
- Griffith D.A., 1993, *Spatial Regression on the PC : Spatial Statistics using SAS*, Association of the American Geographers, Washington, DC.
- Griffith D.A., 1999, Statistical and mathematical sources of regional science theory: map pattern analysis as an example, *Papers in Regional Science*, 78, 21-45.
- Griffith D.A., Can A., 1996, Spatial Statistical/Econometric Versions of Simple Urban Population Density Models, in D.A. Griffith, S.L. Arlinghaus, *Practical Handbook of Spatial Statistics*, Boca Raton, CRC Press.
- Griffith D.A., Layne L.J., 1999, *A Casebook for Spatial Statistical Data Analysis*, Oxford University Press, Oxford.
- Griffith D.A., Sone A., 1995, Trade-offs associated with normalizing constant computational simplifications for estimating spatial statistical models, *Journal of Statistical Computation and Simulation*, 51, 165-183.
- Haining R., 1978, The moving average model for spatial interaction, *Transactions and Papers of the Institute of British Geographers*, 3, 202-225.
- Haining R., 1990, *Spatial Data Analysis in the Social and Environmental Sciences*, Cambridge University Press, Cambridge.
- Heijmans R.D.H, Magnus J.R, 1986a, Asymptotic normality of maximum likelihood estimators obtained from normally distributed but dependent observations, *Econometric Theory*, 12, 374-412.

- Heijmans R.D.H., Magnus J.R., 1986b, Consistent maximum-likelihood estimation with dependent observations : the general (non-normal) case and the normal case, *Journal of Econometrics*, 32, 253-285.
- Hepple L., 1995a, Bayesian techniques in spatial and network econometrics: 1. Model comparison and posterior odds, *Environment and Planning A*, 27, 447-469.
- Hepple L., 1995b, Bayesian techniques in spatial and network econometrics: 2. Computational methods and algorithms, *Environment and Planning A*, 27, 615-644.
- Hepple L., 1998, Exact testing for spatial autocorrelation among regression residuals, *Environment and Planning A*, 30, 85-108.
- Hordijk L., 1979, Problems in estimating econometric relations in space, *Papers of the Regional Science Association*, 42, 99-115.
- Huang J.S., 1984, The autoregressive moving average model for spatial analysis, *Australian Journal of Statistics*, 26, 169-178.
- Jayet H., 1993, *Analyse spatiale quantitative*, Economica, Paris.
- Kelejian H.H., 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*, 20, 103-111.
- Kelejian H.H., 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*, 17, 99-121.
- Kelejian H.H., Prucha I.R., 1999a, A generalized moments estimator for the autoregressive parameter in a spatial model, *International Economic Review*, 40, 509-534.
- Kelejian H.H., Prucha I.R., 1999b, On the asymptotic distribution of the Moran *I* test statistic with applications, *Working Paper, Department of Economics, University of Maryland, College Park, MD*.
- Kelejian H.H., Robinson D.P., 1993, A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure country police expenditure, *Papers in Regional Science*, 72, 297-312.
- Kelejian H.H., Robinson D.P., 1995, Spatial correlation: a suggested alternative to the autoregressive model, in L. Anselin et R. Florax (Eds.), *New Directions in Spatial Econometrics*, Springer, Berlin, 75-95.
- Kelejian H.H., Robinson D.P., 1997, Infrastructure productivity estimation and its underlying econometric specifications: a sensitivity analysis, *Papers in Regional Science*, 76, 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*, 43, 263-264.
- Land K., Deane G., 1992, On the large sample estimation of regression models with spatial -or network - effects terms : a two stage least squares approach, in P. Marsden (Ed.), *Sociological Methodology*, Josey Bass, San Francisco, 221-248.
- LeSage J., 1997, Bayesian estimation of spatial autoregressive models, *International Regional Science Review*, 20, 113-129.
- LeSage J., 1999, Bayesian estimation of limited dependent variable spatial autoregressive models, *Geographical Analysis*, 32, 19-35.
- Li B., 1996, Implementing spatial statistics on parallel computers, in S. Arlinghaus (Ed.), *Practical Handbook of Spatial Statistics*, CRC Press, Boca Raton, 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*, 52, 63-81.
- Magnus, J., 1978, Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix, *Journal of Econometrics*, 7, 281-312 (Corrigenda, *Journal of econometrics*, 10, 261).
- Martin R.J., 1993, Approximations to the determinant term in gaussian maximum likelihood estimation of some spatial models, *Communications in Statistics: Theory and Methods*, 22, 189-205.
- 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*, 32, 335-348.
- McMillen D.P., 1995a, Spatial effects in probit models, in L. Anselin, R. Florax (Eds.), *New Directions in Spatial Econometrics*, Springer, Berlin, 189-228.
- McMillen D.P., 1995b, Selection bias in spatial econometric models, *Journal of Regional Science*, 35, 417-436.
- Moore M., 1988, Spatial linear models, *Communications in Statistics, Stochastic models*, 4, 45-75.

- Moran P., 1948, The interpretation of statistical maps, *Journal of the Royal Statistical Society*, 10B, 243-251.
- Moran P., 1950a, Notes on continuous stochastic phenomena, *Biometrika*, 37, 17-23.
- Moran P., 1950b, A test for serial independence of residuals, *Biometrika*, 37, 178-181.
- Murdoch J.C., Rahmatian M., Thayer M.A., 1993, A spatially autoregressive median voter model of recreation expenditures, *Public Finance Quarterly*, 21, 334-350.
- Ord K., 1975, Estimation methods for models of spatial interaction, *Journal of the American Statistical Association*, 70, 120-126.
- Pace R. K., 1997, Performing large spatial regressions and autoregression, *Economics Letters*, 54, 283-291.
- Pace R. K., Barry R., 1997a, Sparse spatial autoregressions, *Statistics and Probability Letters*, 33, 291-297.
- Pace R. K., Barry R., 1997b, Quick computation of spatial autoregressive estimators, *Geographical Analysis*, 29, 232-246.
- Pace R. K., Gilley O.W., 1997, Using the spatial configuration of the data to improve estimation, *Journal of Real Estate Finance and Economics*, 14, 333-340.
- Paelinck J.H.P., Klaassen L.H., 1979, *Spatial econometrics*, Saxon House, Farnborough.
- Pinkse J., 1998, Asymptotic properties of the Moran and related tests and a test for spatial correlation in probit models, Working Paper, Department of Economics, University of British Columbia, Vancouver, BC.
- Pinkse J., 1999, Moran-flavoured tests with nuisance parameters, in L. Anselin, Florax R. (Eds.), *Advances in Spatial Econometrics*, Springer-Verlag, Heidelberg.
- Pinkse J., Slade, 1998, Contracting in space : an application of spatial statistics to discrete-choice models, *Journal of Econometrics*, 85, 125-154.
- Poirier D.J., Ruud P.A., 1988, Probit with dependent observations, *Review of Economic Studies*, 55, 593-614.
- Rao C.R., 1947, Large sample tests to statistical hypotheses concerning several parameters with applications to problems of estimation, *Proceedings of the Cambridge Philosophical Society*, 44, 50-57.
- Rey S.J., Montouri B.D., 1999, U.S. regional income convergence: a spatial econometric perspective, *Regional Studies*, 33, 145-156.
- Rietveld P., Wintershoven P., 1998, Border effects and spatial autocorrelation in the supply of network infrastructure, *Papers in Regional Science*, 77, 265-276.
- Sneek J.M., Rietveld P., 1998, On the estimation of the spatial moving average model, Working Paper, Tinbergen Institute, Rotterdam.
- Stephan F., 1934, Sampling errors and interpretations of social data ordered in time and space, in: F.A. Ross, *Proceedings of the American Statistical Journal, New Series No 185A*, *Journal of the American Statistical Association*, 29, Suppl., 165-166.
- Student, 1914, The elimination of spurious correlation due to position in time or space, *Biometrika*, 5, 351-360.
- Tiefelsdorf M., 1998, Some practical approximations of Moran I 's exact distribution, *Papers in Regional Science*, 77, 101-129.
- Tiefelsdorf M., 2000, *Modelling Spatial Processes*, Lecture Notes in Earth Sciences, n°87, Springer, Berlin.
- Tiefelsdorf M., Boots B., 1995, The exact distribution of Moran's I , *Environment and Planning A*, 27, 985-999.
- Tobler W., 1979, Cellular geography, in S. Gale et G. Olsson (Eds), *Philosophy in Geography*, 379-386, Reidel, Dordrecht.
- Upton G.J.G., Fingleton B., 1985, *Spatial Data Analysis by Example*. John Wiley, New York.
- Varga A., 1998, *University Research and Regional Innovation: a Spatial Econometric Analysis of Academic Technology Transfers*, Kluwer Academic Publishers, Boston.

TABLEAU 1 : RESUME DES DIFFERENTS TESTS

Modèle 1 : $y = X\beta + \varepsilon$	$\varepsilon \sim \text{Nid}(0, \sigma^2 I)$	
Modèle 2 : $y = \rho W y + X\beta + \varepsilon$	$\varepsilon \sim \text{Nid}(0, \sigma^2 I)$	
Modèle 3 : $y = X\beta + \varepsilon$	$\varepsilon = \lambda W \varepsilon + u$	$u \sim \text{Nid}(0, \sigma^2 I)$
Modèle 4 : $y = \lambda W y + X\beta + W X \delta + u$	$u \sim \text{Nid}(0, \sigma^2 I)$	
Modèle 5 : $y = \rho W_1 y + X\beta + \varepsilon$	$\varepsilon = \lambda W_2 \varepsilon + u$	$u \sim \text{Nid}(0, \sigma^2 I)$

Test	Hypothèse nulle	Modèle non-contraint	Modèle contraint	Estimation
LM_{LAG}	$\rho = 0$	modèle 2	modèle 1	M_1 par MCO
LM_{ERR}	$\lambda = 0$	modèle 3	modèle 1	M_1 par MCO
SARMA	$\rho = 0$ et $\lambda = 0$	modèle 5	modèle 1	M_1 par MCO
LM_{LAG}^*	$\rho = 0$	modèle 5	modèle 3	M_3 par MV
LM_{ERR}^*	$\lambda = 0$	modèle 5	modèle 2	M_2 par MV
FC	$\lambda\beta + \delta = 0$	modèle 4	modèle 3	M_3 par MV

Note : Les deux tests unidirectionnels LM_{LAG} et LM_{ERR} existent aussi en version robuste : RLM_{LAG} et RLM_{ERR} .

**TABLEAU 1 : RESULTATS DES ESTIMATIONS ET DES TESTS
MATRICE DE CONTIGUÏTE D'ORDRE 1**

MODELE	β -convergence (4.30)	Variable exogène décalée	Variable endogène décalée	Autocorrélation des erreurs
Estimation	<i>MCO</i>	<i>MCO</i>	<i>MV</i>	<i>MV</i>
$\hat{\alpha}$	0.170 (0.000)	0,164 (0,000)	0.089 (0.000)	0.196 (0.000)
$\hat{\beta}$	-0.012389 (0.000)	-0,01453 (0,000)	-0.00761 (0.000)	-0.01523 (0.000)
$\hat{\lambda}$	-	-	-	0.694 (0.000)
$\hat{\rho}$	-	-	0.631 (0.000)	-
$\hat{\delta}$	-	0,0026 (0,557)	-	-
LIK	406,721	406,898	431,893	436,702
AIC	-809,442	-807,796	-857,787	-869,403
SC	-803,834	-799,384	-849,375	-863,795
$\hat{\sigma}^2$	$7,569.10^{-3}$	$7,610.10^{-3}$	$4,389.10^{-3}$	$3,931.10^{-3}$
Tests				
JB	2,429 (0,297)	3,348 (0,187)	-	-
Moran	8,812 (0,000)	8,695 (0,000)	-	-
SARMA	69,610 (0,000)	78,108 (0,000)	-	-
LM_{ERR}	69,256 (0,000)	68,199 (0,000)	-	-
RLM_{ERR}	11,563 (0,000)	8,777 (0,003)	-	-
LM_{LAG}	(58,047) (0,000)	69,331 (0,000)	-	-
RLM_{LAG}	0,354 (0,552)	9,909 (0,002)	-	-
LM_{ERR}^*	-	-	1,686 (0,194)	-
LM_{LAG}^*	-	-	-	0,113 (0,737)
FC (Wald)	-	-	-	0,00062 (0,980)

Notes : les “ p-values ” sont entre parenthèses. LIK est la valeur de la fonction de log-vraisemblance à l’optimum. AIC est le critère d’information d’Akaike, SC est le critère d’information de Schwartz, JB est le test de Jarque-Bera de normalité des résidus, FC est la statistique du test du facteur commun.