



HAL
open science

Econométrie spatiale (2, Hétérogénéité spatiale)

Julie Le Gallo

► **To cite this version:**

Julie Le Gallo. Econométrie spatiale (2, Hétérogénéité spatiale). [Rapport de recherche] Laboratoire d'analyse et de techniques économiques(LATEC). 2000, 36 p., Figure, ref. bib.: 3 p. hal-01526969

HAL Id: hal-01526969

<https://hal.science/hal-01526969>

Submitted on 23 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° 2001-01

**Econométrie spatiale
2. Hétérogénéité spatiale**

Julie Le Gallo*

Novembre 2000

e-mail : JleGallo@aol.com

ECONOMETRIE SPATIALE

2. Hétérogénéité 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. La première partie (document de travail n° 2000-05) est consacrée au problème de l'autocorrélation spatiale alors que cette seconde partie (document de travail n° 2001-01) porte 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. The first part deals with spatial autocorrelation (working paper n°2000-05) and this second part deals with spatial heterogeneity (working paper n°2001-01).

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

Classification *JEL*: C51, C52, R15

* L'auteur remercie C. Baumont, M.-C. Pichery et C. Ertur 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 (1988), Haining (1990), Cressie (1993), Jayet (1993), Bailey et Gatrell (1995), Fotheringham, Brundson et al. (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. La première partie est consacrée au problème de l'autocorrélation spatiale alors que la seconde porte sur le problème de l'hétérogénéité spatiale.

ECONOMETRIE SPATIALE

2. Hétérogénéité spatiale

INTRODUCTION

L'étude des marchés immobiliers dans les espaces urbains fait souvent apparaître une segmentation de ces marchés : les caractéristiques et les prix des logements diffèrent substantiellement selon leurs localisations. Cette segmentation provient entre autres de l'inélasticité de la demande des ménages pour certaines caractéristiques des logements ou encore de diverses barrières institutionnelles. Elle conduit à des variations persistantes et significatives des caractéristiques des logements et de leurs prix dans les différents sous-marchés. Dans ces conditions, estimer une relation "globale" entre le prix du logement et ses caractéristiques, relation s'appliquant de la même façon sur toute l'aire urbaine étudiée, est susceptible de masquer des différences importantes dans l'espace.

L'instabilité dans l'espace des relations économiques illustrée par cet exemple est appelée *hétérogénéité spatiale*. Elle se retrouve à plusieurs échelles : les comportements et les phénomènes économiques ne sont pas les mêmes dans le centre d'une ville et dans sa périphérie, dans une région urbaine et dans une région rurale, dans le "Nord" de l'Europe et dans le "Sud", etc.

En pratique, ces différences peuvent se traduire dans une régression économétrique de deux façons : par des variances différentes ou par des coefficients différents. Dans le premier cas, on est confronté à un problème d'*hétéroscédasticité*, problème par ailleurs fréquent en coupe transversale et qui dû à l'utilisation d'unités spatiales hétérogènes, à des variables omises ou à toute autre forme de mauvaise spécification du modèle. Dans le second cas, on parle d'*instabilité des paramètres* de la régression qui varient systématiquement avec la localisation. Ces deux cas peuvent également être combinés.

Ces aspects ont été souvent pris en compte dans la littérature, comme par exemple dans le cadre de l'analyse des espaces urbains : étude de la variation des densités de population (Brueckner, 1986 ; Kau, Lee et Sirmans, 1986 ; McMillen et McDonald, 1998) et des caractéristiques des logements (Dubin et Sung, 1987 ; Can, 1990, 1992) ou étude de la segmentation des marchés immobiliers (Goodman et Kawai, 1982 ; Bender et Hwang, 1985).

L'objectif de cet article est double.

D'une part, le fait de travailler sur des données géographiques en coupe transversale produit des effets particuliers : hétéroscédasticité et instabilité structurelle dans l'espace. Au contraire de l'autocorrélation spatiale¹, les problèmes causés par cette hétérogénéité spatiale peuvent être traités par les méthodes économétriques standards. Ainsi, les modélisations en coupe transversale relatives aux paramètres variables, aux coefficients aléatoires ou aux changements de régimes peuvent être facilement adaptées aux spécifications particulières nécessitées par l'utilisation de données localisées dans le but d'étudier la variation dans l'espace. Dans une première section sont donc présentées les différentes façons permettant de

¹ L'étude détaillée de cet aspect est fourni dans Le Gallo, 2000.

prendre en compte l'hétéroscédasticité et l'instabilité des paramètres dans l'espace ainsi que les méthodes d'estimation et d'inférence adaptées.

D'autre part, les liens entre autocorrélation spatiale et hétérogénéité spatiale sont explicités dans une deuxième section. Ces deux aspects sont en effet souvent présents ensemble dans une régression et des tests adaptés doivent être développés pour permettre de vérifier la présence de l'un des effets en tenant compte de l'autre ou de vérifier la présence de ces deux effets simultanément.

Dans la section 3, un exemple permet d'illustrer ces différentes techniques.

1. MODELISATION ET ESTIMATION DE L'HETEROGENEITE SPATIALE

11. L'hétéroscédasticité

En général, l'hétéroscédasticité provient de variables manquantes ou de toute autre forme de mauvaise spécification. En outre, lorsqu'on travaille sur des données localisées, les unités spatiales utilisées ne sont généralement ni régulières, ni homogènes : elles peuvent avoir des formes et des aires différentes, des niveaux de développement technologique variables, des populations plus ou moins importantes etc.

Dans ces conditions, deux stratégies sont possibles. White (1980) a fourni un estimateur convergent de la matrice des variances-covariances de l'estimateur des Moindres Carrés Ordinaires (MCO) en présence d'hétéroscédasticité de forme inconnue pour que l'inférence statistique basée sur les MCO soit asymptotiquement fiable (paragraphe 111). Alternativement, des tests plus puissants peuvent être obtenus lorsque la forme de l'hétéroscédasticité est correctement spécifiée (paragraphe 112).

111. La méthode de White

Soit le modèle de régression suivant :

$$y = X\beta + \varepsilon \quad [1.1]$$

où y est le vecteur $(N,1)$ des observations de la variable dépendante, N est le nombre d'observations, X est la matrice (N,K) des observations des variables explicatives, β est le vecteur $(K,1)$ des K coefficients inconnus à estimer et ε est le vecteur $(N,1)$ des termes d'erreurs.

Le modèle hétéroscédastique est un cas particulier du modèle où les erreurs sont non-sphériques. La variance de chaque terme n'est plus constante mais au contraire varie avec chaque observation :

$$V(\varepsilon_i) = \sigma_i^2 \quad [1.2a]$$

$$\text{soit } E(\varepsilon\varepsilon') = \Omega \quad [1.2b]$$

En présence d'hétéroscédasticité, les MCO sont non biaisés mais inefficients. Les tests de significativité doivent dans ce cas être basés sur la vraie matrice des variances-covariances de l'estimateur des MCO $(X'X)^{-1} X'\Omega X(X'X)^{-1}$ au lieu de $\sigma^2(X'X)^{-1}$. Lorsqu'on ne connaît pas la vraie forme de l'hétéroscédasticité, White (1980) a démontré que, même si Ω ne pouvait pas être estimée de manière convergente, on pouvait tout de même estimer de manière convergente (sous des conditions de régularité générales) $X'\Omega X$ par $X'GX$ où G est la

matrice diagonale des erreurs estimées élevées au carré². Par conséquent, un test de significativité pour les coefficients de régression est alors basé sur les MCO et la matrice des variances-covariances estimée $(X'X)^{-1}X'GX(X'X)^{-1}$.

Une autre solution consiste à calculer une estimation de la matrice des variances-covariances basée sur la procédure «jackknife» qui consiste à répéter N fois l'estimation par les MCO, où chaque fois une observation est éliminée. Une estimation pour la covariance des estimateurs MCO est obtenue par :

$$\frac{N-1}{N} \sum_i [\hat{\beta}(i) - \frac{1}{N} \sum_j \hat{\beta}(j)] [\hat{\beta}(i) - \frac{1}{N} \sum_j \hat{\beta}(j)]' \quad [1.3]$$

où $\hat{\beta}(i)$ est l'estimation sur l'échantillon privé de la $i^{\text{ème}}$ observation (Efron, 1982). MacKinnon et White (1985) ont adapté cette méthodologie au cas de l'hétéroscédasticité.

Cette approche conduit au test de White (1980), qui est basé sur le principe du multiplicateur de Lagrange et dans lequel on teste l'homoscédasticité contre une forme tout à fait générale :

$$\begin{aligned} H_0 : \sigma_i^2 &= \sigma^2, \forall i \\ H_1 : \text{non } H_0 \end{aligned} \quad [1.4]$$

Soit $\hat{\varepsilon}_i$ l'erreur estimée pour la $i^{\text{ème}}$ observation. Pour trouver la statistique du test, on calcule NR^2 dans la régression de $\hat{\varepsilon}_i^2$ sur une constante, les régresseurs de l'équation estimée, leurs carrés et leurs produits croisés.

Cette statistique est asymptotiquement distribuée comme un χ^2 à $P - 1$ degrés de libertés, où P est le nombre de régresseurs, sans la constante.

Le test de White est un test très général qui ne nécessite pas de connaissances a priori sur la forme de l'hétéroscédasticité. Par ailleurs, si l'échantillon est assez grand, le test de White est suffisamment puissant pour détecter toutes les formes d'hétéroscédasticité pour lesquelles la matrice des variances-covariances usuelle n'est pas convergente. Cependant, même si le test de White est convergent contre une grande variété d'alternatives hétéroscédastiques, il peut ne pas être très puissant en échantillon fini.

Enfin, il faut souligner que le test de White peut aussi identifier un problème dans la spécification du modèle, une dépendance entre les régresseurs et les erreurs.

112. Spécification de l'hétéroscédasticité

Une deuxième solution consiste à donner à l'hétéroscédasticité une forme particulière. Lorsque l'hétéroscédasticité est correctement spécifiée, les tests d'hétéroscédasticité associés sont alors plus puissants que le test de White. En effet, comme le souligne Greene (1999), la

² Pour les petits échantillons, MacKinnon et White (1985) ont proposé d'ajuster les résidus élevés au carré de la matrice G par le terme $1/(1 - k_{ii})$ où k_{ii} est le $i^{\text{ème}}$ élément de la matrice idempotente $X(X'X)^{-1}X'$.

puissance d'un test d'hétéroscédasticité dépend de la précision avec laquelle on spécifie sa forme. Deux formes possibles adaptées au cas spatial sont présentées dans ce paragraphe³.

1/ La première possibilité consiste à exprimer la variance comme une fonction d'un ensemble de variables :

$$\sigma_i^2 = \sigma^2 f(z' \alpha) \quad [1.5]$$

où σ^2 est un facteur d'échelle, f une forme fonctionnelle, z est un vecteur de T variables, $z' \alpha = \alpha_0 + \alpha_1 z_1 + \dots + \alpha_T z_T$. Enfin, $\alpha_i, i = 1 \dots T$ sont des paramètres inconnus.

Les formes fonctionnelles peuvent être linéaires (cas d'une hétéroscédasticité additive) ou exponentielles (cas d'une hétéroscédasticité multiplicative (Harvey, 1976)). Les variables z incluses dans cette spécification correspondent ou non aux variables explicatives. Si l'on soupçonne une hétéroscédasticité basée sur la différence de taille entre les régions, on peut par exemple inclure l'aire de ces régions ou toute autre variable reliée à la taille (population totale, densité de population, etc.).

Le modèle est homoscedastique si $\alpha_i = 0, i = 1 \dots T$. Cette forme des erreurs est à la base du test de Breusch-Pagan (1979). Il s'agit d'un test du multiplicateur de Lagrange relativement puissant dès lors qu'on a correctement identifié les variables susceptibles d'être à la source de l'hétéroscédasticité. On teste l'hypothèse nulle :

$$H_0 : \alpha_i = 0, i = 1 \dots T \quad [1.6]$$

Soit Z la matrice $(N, T + 1)$ comportant les observations sur $(1, z_i)$ et soit g le vecteur $(N, 1)$ comportant les observations sur $g_i = \hat{\varepsilon}_i^2 / (\hat{\varepsilon}' \hat{\varepsilon} / N)$. La statistique du test s'écrit :

$$BP = \frac{1}{2} (g' Z (Z' Z)^{-1} Z' g - N) \quad [1.7]$$

Sous l'hypothèse nulle d'homoscedasticité, BP est asymptotiquement distribué selon un χ^2 à T degrés de liberté⁴. Une procédure asymptotiquement équivalente consiste à régresser $\hat{\varepsilon}_i^2$ sur Z et à calculer NR^2 . Cette quantité est asymptotiquement distribuée suivant un χ^2 à T degrés de liberté sous l'hypothèse nulle.

2/ La deuxième possibilité est de spécifier une hétéroscédasticité en blocs. On considère différents régimes et la variance des erreurs est différente pour chaque régime mais constante à l'intérieur de chaque régime. Par exemple, dans le cas simple de 2 régimes (comme par exemple une partition "Nord/Sud" ou "centre/périphérie"), la matrice des variances-covariances de ε s'écrit de la façon suivante :

$$\Omega = \begin{bmatrix} \sigma_1^2 I_{N_1} & 0 \\ 0 & \sigma_2^2 I_{N_2} \end{bmatrix} \quad [1.8]$$

³ Certaines modélisations visant à capter l'instabilité structurelle produisent aussi une certaine forme d'hétéroscédasticité (voir infra).

⁴ Ce test étant sensible à l'hypothèse de normalité, Koenker (1981) et Koenker et Basset (1982) ont proposé un test plus robuste à la non-normalité des erreurs.

où σ_1^2 est la variance du premier régime, σ_2^2 est la variance du second régime, N_1 et N_2 sont respectivement le nombre d'observations dans le premier et dans le deuxième régime, avec $N_1 + N_2 = N$, I_{N_i} étant la matrice identité d'ordre N_i , $i = 1, 2$.

L'estimation de ce modèle peut s'effectuer par la méthode du maximum de vraisemblance (MV). Considérons G blocs différents. La fonction de log-vraisemblance pour le modèle de régression linéaire avec hétéroscédasticité en blocs s'écrit :

$$\ln L_1(y|\beta, \sigma_1^2, \dots, \sigma_G^2) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \sum_g (N_g \ln \sigma_g^2) - \frac{1}{2} \sum_g \left(\frac{1}{\sigma_g^2} \sum_i \varepsilon_{ig}^2 \right) \quad [1.9]$$

L'estimateur MV de β correspond à l'estimateur des MCG, compte tenu des estimateurs MV de σ_g^2 :

$$\hat{\beta} = \left[\sum_g \frac{1}{\sigma_g^2} X_g' X_g \right]^{-1} \left[\sum_g \frac{1}{\sigma_g^2} X_g' y_g \right] \quad [1.10]$$

L'estimateur MV de σ_g^2 est donné par $\hat{\sigma}_g^2 = (\hat{\varepsilon}_g' \hat{\varepsilon}_g) / N_g$ avec $\varepsilon_g = y_g - X_g \hat{\beta}$ pour $g = 1, \dots, G$.

Dans ces conditions, pour obtenir l'estimateur du maximum de vraisemblance de β , on peut appliquer une méthode itérative. Dans un premier temps, on estime β par les MCO. Ensuite, les variances sont estimées séparément pour chaque groupe : $\hat{\sigma}_g^2 = (\hat{\varepsilon}_g' \hat{\varepsilon}_g) / N_g$. Dans un troisième temps, on recalcule $\hat{\beta}$ avec [1.10]. On recommence les deux dernières étapes tant que le critère de convergence numérique pour $\hat{\beta}$ n'a pas été atteint.

Dans le cas particulier d'une hétéroscédasticité en blocs, sous l'hypothèse de normalité, le test du ratio de vraisemblance est un test uniformément plus puissant.

Sous l'hypothèse nulle d'homoscédasticité (variances identiques), la fonction de log-vraisemblance de [1.1] est :

$$\ln L_0(y|\beta, \sigma^2) = -\frac{N}{2} [\ln(2\pi) + \ln \sigma^2] - \frac{1}{2\sigma^2} \sum_i \varepsilon_i^2 \quad [1.11]$$

où $N = \sum_g N_g$ est le nombre total d'observations. L'estimation par le maximum de vraisemblance de σ^2 est donnée par $\hat{\sigma}^2 = (\hat{\varepsilon}' \hat{\varepsilon}) / N$.

Sous l'hypothèse alternative d'hétéroscédasticité entre les G groupes, cette fonction est donnée par [1.9].

Si on évalue L_0 et L_1 à l'optimum, le test du ratio de vraisemblance est donné par :

$$-2(\ln L_0 - \ln L_1) = N \ln \hat{\sigma}^2 - \sum_g N_g \ln \hat{\sigma}_g^2 \quad [1.12]$$

Sous l'hypothèse nulle, cette statistique est asymptotiquement distribuée selon un χ^2 à $G - 1$ degré de liberté.

12. L'instabilité des paramètres

L'instabilité des paramètres provient de l'absence de stabilité dans l'espace des comportements ou des relations économiques : les paramètres d'une régression (voire la forme fonctionnelle de cette dernière), varient selon leurs localisations. Ce type de phénomène nécessite alors des modélisations permettant la prise en compte des caractéristiques particulières de chaque région.

La variabilité spatiale de la moyenne ou des coefficients d'une régression peut être *discrète*, lorsque des différences systématiques entre les différents régimes sont observés (paragraphe 121), ou *continue* sur l'ensemble de la zone étudiée (paragraphe 122). Enfin, la méthode GWR («Geographically Weighted Regression») est une méthode non-paramétrique visant à capter finement l'instabilité des coefficients d'une régression (paragraphe 123).

121. Variation spatiale discrète

121.1 ANOVA spatiale

Une première question que l'on peut se poser est de savoir dans quelle mesure la moyenne d'une variable varie entre différents sous-groupes d'observations localisées. Par exemple, on peut se demander si la moyenne des prix des logements est la même à l'Est et à l'Ouest d'une ville. Une réponse possible passe par une régression où la variable expliquée est la variable étudiée et les variables explicatives des variables muettes représentant l'appartenance aux différents sous-groupes.

Formellement, considérons une zone partitionnée en G sous-groupes de localisations, avec N_g ($g = 1, \dots, G$) régions dans le $g^{\text{ème}}$ sous-groupe. On a alors $\sum_{g=1}^G N_g = N$ où N est le nombre total de régions. Notons y_i la valeur de la variable aléatoire Y étudiée pour la $i^{\text{ème}}$ unité ($i = 1, \dots, N$). On estime alors le modèle suivant :

$$y_i = \alpha_1 + \sum_{g=2}^G \alpha_g d_{ig} + \varepsilon_i \quad [1.13]$$

où d_{ig} sont les variables muettes définies de la façon suivante :

$$d_{ig} = \begin{cases} 1 & \text{si la région } i \text{ appartient à la zone } g \\ 0 & \text{sinon} \end{cases} \quad (i = 1, \dots, N; g = 2, \dots, G) \quad [1.14]$$

α_1 s'interprète comme la moyenne de la région omise (ici la première région) et les autres coefficients α_g s'interprètent comme la différence entre la moyenne de la région g et celle de la région omise.

Ce modèle peut s'estimer par les MCO si les hypothèses habituelles sont vérifiées. Sous l'hypothèse de normalité des erreurs, un test d'égalité des moyennes s'effectue alors en testant la significativité des coefficients α_g , $g = 2, \dots, G$.

Enfin, la table d'analyse de la variance associée à [1.13] est donnée par Griffith (1978) :

Source	Somme des carrés	Degrés de liberté
Inter-régionale	$\hat{\beta}' X' y - n\bar{y}^2$	$G - 1$
Intra-régionale	$y' y - \hat{\beta}' X' y$	$N - G$
Totale	$y' y - n\bar{y}^2$	$N - 1$

où $\hat{\beta}$ est l'estimateur des MCO, X est la matrice des observations des variables explicatives et y est le vecteur des observations de la variable expliquée.

121.2 Régimes spatiaux

Dans le modèle [1.13], seule la constante varie entre les différents sous-groupes spatiaux. Plus généralement, si l'on considère l'ensemble des coefficients d'une régression, l'hétérogénéité spatiale peut être présente sous la forme de différentes constantes et/ou de différentes pentes. Dans ce cas général, on parle alors d'instabilité structurelle dans l'espace ou de régimes spatiaux (Anselin, 1988a).

Formellement, considérons le cas où deux régimes sont considérés, indiqués par 1 et 2 :

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad [1.15]$$

y_1 et y_2 sont les vecteurs des observations des variables dépendantes, X_1 et X_2 sont les matrices des observations des variables explicatives, β_1 et β_2 sont les vecteurs de coefficients et ε_1 et ε_2 sont les vecteurs des erreurs respectivement des sous-ensembles 1 et 2. Supposons que le nombre total d'observations dans le sous-ensemble 1 est N_1 alors que celui du sous-ensemble 2 est N_2 ($N_1 + N_2 = N$). Les observations dans chaque régime sont supposées être suffisamment importantes pour estimer β_1 et β_2 . Enfin, soit $\varepsilon' = [\varepsilon_1' \ \varepsilon_2']$ et Ω sa matrice des variances-covariances : $\Omega = E[\varepsilon\varepsilon']$. Les notations peuvent être simplifiées :

$$y = X^* \beta + \varepsilon \quad [1.16]$$

où y , X^* , β et ε correspondent à la structure [1.15].

Cette approche en termes de régimes spatiaux nécessite donc la définition de régimes spatiaux. S'il existe des méthodes permettant une détection endogène des régimes en séries temporelles (Bai et Perron, 1998), à l'heure actuelle il n'y en a pas en revanche permettant à la fois une estimation endogène de ces régimes et une prise en compte de l'autocorrélation spatiale. En conséquence, l'analyse exploratoire des données spatiales permettant d'étudier les schémas globaux et locaux d'autocorrélation spatiale (Anselin, 1995) s'avère alors utile pour suggérer les concentrations locales pouvant servir de base à ces régimes (voir par exemple Le Gallo et Ertur, 2000).

122. Variation spatiale continue

Comme précédemment dans le cas discret, on veut étudier la variabilité de la moyenne ou celle des coefficients de régression. Dans le premier cas, on utilise la méthode TSA («Trend

Surface Analysis ») alors que dans le second cas, c'est la méthode VE («Variable Expansion») et ses extensions qui sont les plus adaptées.

122.1 La méthode TSA

Dans cette méthode, on régresse la variable étudiée sur une expansion polynomiale des coordonnées de chaque localisation. Le modèle suivant est estimé par les MCO :

$$y_i = \sum_{r+s \leq p} a_{rs} u_i^r v_i^s + \varepsilon_i \quad [1.17]$$

où p est l'ordre de la TSA, u_i et v_i sont les coordonnées de la $i^{\text{ème}}$ observation (resp. la latitude et la longitude), a_{rs} sont les coefficients inconnus à estimer dont le nombre dépend de l'ordre de la TSA.

Par exemple, les deux premières fonctions, la fonction linéaire ($p = 1$) et la fonction quadratique ($p = 2$), s'écrivent de la façon suivante :

$$y_i = a_{00} + a_{10}u_i + a_{01}v_i + \varepsilon_i \quad [1.18a]$$

$$y_i = a_{00} + a_{10}u_i + a_{01}v_i + a_{11}uv + a_{02}u_i^2 + a_{20}v_i^2 + \varepsilon_i \quad [1.18b]$$

Cette méthode permet de décrire les grandes tendances caractéristiques de la surface de régression – comme des tendances simples "Nord-Sud" ou "Est-Ouest" ou des tendances plus complexes - et sert donc à "lisser" les données (Ripley, 1981; Agterberg, 1984). Enfin, comme la variable expliquée est uniquement une fonction des coordonnées des points, des valeurs prévues pour cette variable peuvent facilement être obtenues pour chaque localisation et représentées sur une carte.

Ce type de modélisation a, entre beaucoup d'autres applications, servi dans l'étude des fonctions de densité urbaines (Schroeder et Sjoquist, 1976) et dans les modèles hédoniques de prix immobiliers (Olmo, 1995; Des Rosiers et Thériault, 1995, 1996) mais elle présente plusieurs défauts (Ripley, 1981). Tout d'abord, les variables des polynômes sont souvent fortement corrélés, ce qui est source de multicollinéarité. De plus, l'autocorrélation spatiale des erreurs n'est pas systématiquement éliminée dans ce type de modèle⁵.

122.2. La méthode VE

La méthode VE a été développée par Casetti (1972, 1997) et elle généralise la méthode précédente : on suppose que l'hétérogénéité spatiale se traduit par des coefficients de la régression (autres que la constante) qui varient pour chaque observation. Cette différenciation est spécifiée par une fonction dépendant d'un certain nombre de variables auxiliaires.

Sans perte de généralité, les propriétés de cette méthode sont illustrées formellement par un exemple simple avec une seule variable explicative. *Le modèle initial* est, pour une observation i :

$$y_i = \beta_0 + \beta_{1i}x_i + \varepsilon_i \quad [1.19]$$

⁵ Dans le cadre des modèles hédoniques, DesRosiers et Thériault (1999) et Thériault et al. (1999) ont développé une stratégie d'estimation et de tests combinant autocorrélation spatiale et variables polynomiales tout en évitant les problèmes de multicollinéarité.

β_0 et β_{1i} sont les coefficients de régression et x_i correspond à la $i^{\text{ème}}$ observation de la variable explicative. On suppose que le second coefficient est variable pour chaque observation⁶ : β_1 est une fonction exacte de variables d'"expansion" ou variables augmentées, notées $z_i, i = 1, \dots, P$:

$$\beta_{1i} = f(z_i, \gamma) \quad [1.20]$$

où z_i est un vecteur de P variables augmentées, γ est le vecteur des P paramètres inconnus correspondants, f est la relation fonctionnelle qui exprime la forme de la variation du coefficient β_1 .

En général, cette fonction f est linéaire ; pour deux variables augmentées, on aura :

$$\beta_{1i} = \gamma_0 + \gamma_1 z_{1i} + \gamma_2 z_{2i} \quad [1.21]$$

L'introduction des variables augmentées dans la formulation originale donne le modèle terminal suivant, estimé par les MCO :

$$y_i = \beta_0 + \gamma_0 x_i + \gamma_1 (z_{1i} x_i) + \gamma_2 (z_{2i} x_i) + \varepsilon_i \quad [1.22]$$

Comme dans la méthode TSA, les variables z_i correspondent le plus souvent aux coordonnées du point i (latitude et longitude). Dans le cas linéaire le plus simple, seules des tendances simples dans la variation des paramètres sont captées (tendance du type "Nord/Sud" ou "Est/Ouest") et des expansions quadratiques ou d'ordre supérieur doivent être spécifiées pour capter des tendances plus complexes. A nouveau, plus l'ordre de l'expansion est grand, plus la multicollinéarité risque de poser problème⁷.

Mises à part les coordonnées des observations, les variables augmentées peuvent plus généralement correspondre à toutes les variables susceptibles d'être à l'origine de la différenciation spatiale des coefficients. Par exemple, LeSage (1999) a utilisé la distance au centre de la ville pour modéliser la variation spatiale des prix marginaux des caractéristiques dans un modèle de prix immobiliers. Toujours dans le même cadre, Can (1988, 1992) a préféré utiliser une variable composite de la "qualité" du voisinage de chaque logement : revenu moyen, composition raciale de la population, taux de chômage, etc.⁸

Revenons au cas simple linéaire. Sous forme matricielle, le modèle [1.22] s'écrit de la façon suivante :

$$y = S\beta_0 + \gamma_0 x + \gamma_1 (z_1 x) + \gamma_2 (z_2 x) + \varepsilon \quad [1.23]$$

où $z_1 = \text{diag}(z_{1i})$ et $z_2 = \text{diag}(z_{2i})$.

Si le modèle terminal est la spécification correcte, les estimateurs des paramètres du modèle initial sont biaisés. En effet, il s'agit d'un cas particulier du problème traditionnel de variables omises. Soient $b' = [\beta_0 \ \gamma_0]$, $X = [S \ x]$, $Z = [z_1 x \ z_2 x]$ et $\gamma' = [\gamma_1 \ \gamma_2]$. On vérifie immédiatement que l'espérance de l'estimateur de b du modèle initial si le modèle final est

⁶ On peut également supposer la constante variable pour chaque observation.

⁷ Casetti et Jones (1983) proposent alors de remplacer les variables augmentées initiales par leur composantes principales (méthode VE orthogonale).

⁸ Voir Casetti et Jones, 1992 pour les différentes applications de cette méthode dans le cadre spatial ou non.

vrai est $E(\hat{b}) = b + (X'X)^{-1} X'Z\gamma$. L'estimateur de b dans le modèle initial est biaisé (Anselin, 1988a ; Anselin et Griffith, 1988).

122.3 La méthode non-paramétrique GWR

D'après Fotheringham et al. (2000), la méthode VE permet de modéliser des paramètres variables dans l'espace mais elle souffre de deux limitations principales. La première est que cette technique ne peut capter que des *tendances* dans les relations dans l'espace, la complexité de ces tendances étant dépendante de la complexité des équations d'expansion spécifiées. Les estimations des paramètres spatialement variables obtenues à partir de cette méthode ne captent donc pas toujours d'éventuelles variations locales importantes. La seconde est que la forme des équations d'expansion doit être supposée a priori.

Pour pallier ces problèmes, trois auteurs, Brundson, Fotheringham et Charlton, ont développé la méthode d'estimation non-paramétrique GWR ("Geographically Weighted Regression") présentée dans un ouvrage (Fotheringham et al., 2000) et dans une série d'articles (Brundson et al., 1996, 1998a, 1998b, 1999 ; Fotheringham, 1997 ; Fotheringham et Brundson, 1999 ; Fotheringham et al., 1996a, 1996b, 1997, 1998). Cette méthode est destinée à capter au niveau de chaque observation les variations des coefficients d'une régression dans l'espace. Pour cela, un paramètre différent est estimé pour chaque observation en utilisant les valeurs des caractéristiques prises par les observations voisines.

Formellement, considérons la régression globale suivante pour une observation i :

$$y_i = \sum_k \beta_k x_{ki} + \varepsilon_i \quad [1.24]$$

Dans ce modèle, un seul paramètre est estimé pour mesurer la relation entre chaque variable indépendante et la variable expliquée. De plus, cette relation est supposée constante dans l'espace. Ce modèle s'estime par les MCO si les conditions sont vérifiées.

Au contraire, dans la formulation GWR, les paramètres à estimer sont locaux. En d'autres termes, les coefficients β_k ne sont pas uniques mais ils sont variables pour chaque observation i . Le modèle s'écrit :

$$y_i = \sum_k \beta_{ki} x_{ki} + \varepsilon_i \quad [1.25]$$

Cette spécification permet l'estimation d'une surface continue des valeurs des paramètres dans l'espace. Le modèle global [1.24] est un cas spécial du modèle GWR représenté par le modèle [1.25] dans lequel la surface des paramètres est supposée constante dans l'espace.

Pour estimer les paramètres β_{ki} du modèle [1.25], on suppose que les données observées près du point i ont plus d'influence dans l'estimation de β_{ki} que les données situées loin de ce point i . On utilise alors une pondération variable avec le point i . Notons $\hat{\beta}_i$ le vecteur des K coefficients estimés pour l'observation i . Cet estimateur des moindres carrés pondéré s'écrit de la façon suivante :

$$\hat{\beta}_i = (X'V_iX)^{-1} X'V_iY \quad [1.26]$$

où X est la matrice des observations des variables explicatives et Y représente le vecteur des observations de la variable dépendante. Enfin, V_i est une matrice (N, N) , variable pour chaque observation i , dont les éléments hors de la diagonale principale sont nuls et dont les éléments de la diagonale principale dénotent le poids géographique accordé aux données autour du point i :

$$V_i = \text{diag}(v_{i1}, \dots, v_{in}) \quad [1.27]$$

où v_{ij} dénote le poids de la $j^{\text{ème}}$ observation sur l'estimation du modèle autour du $i^{\text{ème}}$ point.

Cette méthodologie diffère de la régression non-paramétrique du noyau habituelle (Cleveland, 1979 ; Cleveland et Devlin, 1988, Cleveland et al., 1988) dans laquelle les pondérations se réfèrent à l'espace des *attributs* de X : les données plus similaires à x_i sont pondérées davantage que les données moins similaires et le résultat est un ensemble d'estimation de paramètres dans l'espace de X . Au contraire, la méthode GWR utilise des pondérations basées sur l'espace *géographique*. Selon les auteurs, différents schémas de pondération V_i , ou "noyaux", ont été utilisés (on trouvera dans Fotheringham et al. (2000) une description des principaux poids utilisés). Une possibilité consiste par exemple à retenir des poids de la forme suivante :

$$v_{ij} = \exp(-d_{ij}^2 / h^2) \quad [1.28]$$

Dans la terminologie non-paramétrique, h est appelé la fenêtre : c'est le paramètre qui détermine le nombre d'observations utilisées pour estimer les coefficients associés à chaque observation i .

Pour étudier la répartition spatiale des densités de population et des valeurs foncières à Chicago, McMillen (1996) et McMillen et McDonald (1997) ont utilisé une fonction "tri-cube". On attribue aux h observations les plus proches de l'observation i des poids qui diminuent avec la distance à cette observation. Cette fonction s'écrit de la façon suivante :

$$v_{ij} = [1 - (d_{ij} / d_i)^3]^3 \quad I(d_{ij} < d_i) \quad [1.29]$$

où d_{ij} est la distance euclidienne entre l'observation i et l'observation j et d_i est la distance du $h^{\text{ème}}$ voisin le plus proche de l'observation i . $I()$ une fonction indicatrice égale à 1 si la condition est vérifiée et 0 sinon.

Quel que soit le noyau utilisé, comme dans toute estimation non-paramétrique, il existe un arbitrage entre le biais et la variance. Dans [1.28] par exemple, lorsque h tend vers l'infini, les poids tendent vers 1 pour toutes les paires de points, les estimations GWR deviennent équivalentes à celles des MCO et le biais augmente. Quand la fenêtre devient plus petite, les estimations des paramètres vont dépendre de plus en plus des observations proches de i et auront donc une variance accrue. La fenêtre détermine donc le degré de "lissage" des données ainsi que le biais et la variance de l'estimateur non-paramétrique associé. Le problème est donc de sélectionner un noyau et une fenêtre appropriés. Le choix d'un noyau est moins important que celui de la fenêtre qui a une plus grande incidence sur les propriétés des estimateurs.

Pour choisir la valeur de h dans [1.28], Cleveland (1979) et Bowman (1984) ont proposé de minimiser la quantité suivante (procédure de "cross-validation") :

$$z = \sum_{i=1,n} [y_i - y_{\neq i}(h)]^2 \quad [1.30]$$

où $y_{\neq i}(h)$ est la valeur estimée de y_i lorsque les observations associées au point i sont omises du processus de calibrage.

Le résultat d'une GWR est un ensemble d'estimations localisées des paramètres (ainsi que des versions localisées de mesures de qualité de la régression comme le R^2). Ces estimations locales étant toutes associées à des localisations spécifiques, chaque ensemble de paramètres peut être cartographié pour illustrer les variations spatiales de la relation mesurée. Pour savoir si ces estimations locales des paramètres sont significativement différentes, des tests paramétriques ont été proposés par Brundson et al. (1999) et Leung et al. (2000).

13. Instabilité des paramètres et hétéroscédasticité

L'hétérogénéité spatiale, qui est due à l'instabilité dans l'espace des relations économiques, peut être modélisée par exemple par la spécification de régimes spatiaux sous la forme [1.15]. Pourtant, lorsque l'instabilité spatiale n'est pas complètement prise en compte de cette façon, il peut subsister une hétéroscédasticité qui représente l'hétérogénéité non observée. Ainsi, il est possible qu'une hétéroscédasticité résiduelle se manifeste dans le cadre d'un modèle avec régimes spatiaux (paragraphe 131). En revanche, la situation est différente dans le cadre du modèle VE [1.22] dans lequel l'hétéroscédasticité des erreurs est induite par la structure même du modèle (paragraphe 132).

131. Régimes spatiaux et hétéroscédasticité

Reprenons les notations du modèle [1.15] où deux régimes spatiaux sont considérés :

$$\begin{aligned} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \\ \Rightarrow y &= X^* \beta + \varepsilon \end{aligned}$$

Dans Quandt (1958), l'estimation est basée sur la fonction de vraisemblance du modèle global, dans lequel on suppose en plus que les termes d'erreurs ont une variance différente pour chaque sous-ensemble (modèle [1.8]). La fonction de log-vraisemblance correspondante est de la forme (en ignorant le terme constant) :

$$L = -\left(\frac{N_1}{2}\right) \ln(\sigma_1)^2 - \left(\frac{N_2}{2}\right) \ln(\sigma_2)^2 - \frac{(y_1 - X_1 \beta_1)'(y_1 - X_1 \beta_1)}{2(\sigma_1)^2} - \frac{(y_2 - X_2 \beta_2)'(y_2 - X_2 \beta_2)}{2(\sigma_2)^2} \quad [1.31]$$

Cette fonction de vraisemblance peut être modifiée pour tenir compte de structures d'erreurs plus complexes (comme par exemple, l'incorporation d'une autocorrélation spatiale des erreurs, voir infra, section 2).

132. Variation spatiale continue et hétéroscédasticité

132.1 Expansion aléatoire

Considérons le modèle VE :

$$y_i = \beta_0 + \beta_{1i}x_i + \varepsilon_i$$

avec $\beta_{1i} = \gamma_0 + \gamma_1z_{1i} + \gamma_2z_{2i}$

Jusqu'à présent, on a supposé que la forme exacte de l'expansion est connue a priori. Pourtant, en pratique, l'hypothèse d'une relation exacte (déterministe) entre les coefficients et leurs variables augmentées est difficile à tenir (Anselin, 1988a, 1992). Il est donc plus pertinent de rajouter un terme aléatoire dans l'expansion linéaire [1.22] :

$$\beta_{1i} = \gamma_0 + \gamma_1z_{1i} + \gamma_2z_{2i} + v_i \quad [1.32]$$

v_i est un terme aléatoire normalement et indépendamment distribué d'espérance nulle et de variance σ_v^2 . La substitution de cette expression dans le modèle initial conduit à un modèle terminal différent :

$$y = S\beta_0 + \gamma_0x + \gamma_1(z_1x) + \gamma_2(z_2x) + vx + \varepsilon \quad [1.33]$$

Soit w le nouveau terme aléatoire : $w = vx + \varepsilon$. Cette erreur est normalement et indépendamment distribuée d'espérance nulle et de matrice des variances-covariances $V(w) = \sigma_v^2 \text{diag}(x_i^2) + \sigma_\varepsilon^2 I_N$ (si les erreurs v et ε sont indépendantes).

Le modèle final induit donc une certaine forme d'hétéroscédasticité dont il faut tenir compte pour l'inférence statistique, en particulier si l'on cherche à tester la significativité de γ_1 et γ_2 . Cette hétéroscédasticité peut être testée avec le test de Breusch-Pagan. Cette approche revient alors à tester la version aléatoire contre la version déterministe de la méthode VE : sous l'hypothèse nulle d'homoscédasticité, on retrouve le modèle [1.23] alors que sous l'hypothèse alternative d'hétéroscédasticité, on retrouve le modèle [1.33].

Dans le cas où l'hétéroscédasticité impliquée par la méthode des variables augmentées est liée à la distance à un point du plan (par exemple la distance à un point central), Casetti (1982) et Casetti et Can (1998) ont proposé la modélisation DARP («Drift Analysis of Regression Parameters»). Les auteurs supposent que la variance de chaque observation i dans [1.26] est une fonction de la distance entre cette observation et le point de référence. Il s'agit alors d'un modèle hétéroscédastique et l'estimation des paramètres de ce modèle requiert soit les MCQG, soit la méthode du maximum de vraisemblance.

132.2 Les coefficients aléatoires

Parfois, il n'y a pas de variables supplémentaires disponibles pouvant servir à la détermination d'une forme spécifique de la variation spatiale des coefficients de régression. Dans ce cas, une approche alternative a été proposée par Hildreth-Houck (1968) sous la forme de coefficients aléatoires. L'hétérogénéité est modélisée comme une variation aléatoire autour de la moyenne. Au contraire des modèles précédents, ce modèle n'est donc pas explicitement "spatial".

Formellement, pour chaque observation i , on pose :

$$y_i = x_i' \beta_i \quad \text{avec} \quad \beta_i = \beta + u_i \quad [1.34]$$

Le vecteur aléatoire u_i est supposé suivre une loi normale multivariée d'espérance nulle et de matrice des variances-covariances $\Sigma_{(K,K)}$. On suppose cette matrice diagonale, c'est-à-dire des covariances entre les erreurs des paramètres individuels nulles. La substitution de cette expression dans le modèle standard de régression donne :

$$y_i = x_i' \beta + v_i \quad \text{avec} \quad v_i = x_i' u_i \quad [1.35]$$

Les moments de v_i se déduisent directement de cette écriture : $E(v_i) = 0$ et $V(v_i) = x_i' \Sigma x_i$.

Le modèle [1.35] est estimé par la méthode du maximum de vraisemblance ou par la méthode des moindres carrés généralisés. La présence de coefficients aléatoires pourra être testée, avant l'estimation du modèle plus complexe, grâce à un test d'hétéroscédasticité, comme celui de Breusch-Pagan (1979). Sous l'hypothèse nulle d'homoscédasticité, on retrouve alors le modèle de régression simple.

2. L'INTERACTION ENTRE AUTOCORRELATION ET HETEROGENEITE SPATIALES

2.1. La position du problème

La complexité du lien entre autocorrélation spatiale et hétérogénéité spatiale se retrouve à plusieurs niveaux.

Premièrement, il n'est pas toujours facile de distinguer les conséquences de l'autocorrélation de celles de l'hétérogénéité spatiales dans une coupe transversale : il peut y avoir équivalence observationnelle entre ces deux effets (Anselin et Bera, 1998 ; Anselin, 2000). Par exemple, l'observation d'un regroupement spatial de valeurs très élevées des erreurs estimées peut provenir d'une hétérogénéité (qui se manifeste par exemple sous la forme d'une hétéroscédasticité en blocs) ou une autocorrélation spatiale (i.e. un processus stochastique impliquant le regroupement de points extrêmes).

Deuxièmement, les tests pour l'hétéroscédasticité ne sont pas fiables en présence d'autocorrélation spatiale. Des simulations de Monte-Carlo effectuées par Anselin et Griffith (1988) montrent par exemple que la présence d'autocorrélation spatiale des erreurs a des répercussions sur le biais et sur la puissance des tests de White et de Breusch-Pagan. En particulier, sous l'hypothèse nulle d'absence d'hétéroscédasticité, les fréquences empiriques de rejet du test de White est inférieure au niveau de signification choisi en présence d'autocorrélation spatiale alors que pour le test de Breusch-Pagan, les fréquences de rejet sont trop importantes. Les puissances des tests sont également affectées : elles diminuent pour des valeurs positives du paramètre spatial. Inversement, les tests d'autocorrélation spatiale voient leurs propriétés affectées lorsqu'il y a hétéroscédasticité de forme inconnue. La présence jointe d'hétéroscédasticité et de dépendance spatiale nécessite donc des tests spécialisés et des méthodes d'estimation adaptées (Anselin, 1988a, 1988b).

Troisièmement, l'autocorrélation spatiale est parfois le résultat d'une instabilité des paramètres qui n'est pas prise en compte (Anselin et Griffith, 1988 ; Brundson et al., 1999). En d'autres termes, si des relations variables dans l'espace sont modélisées à l'aide d'une

régression "globale", les termes d'erreurs peuvent être spatialement autocorrélés. Ainsi, supposons une région pour laquelle y est régressé sur x et que le modèle global qui en résulte est :

$$y_i = 1.5 + 2x_i \quad [2.1]$$

Cependant, supposons que la vraie relation n'est pas la même dans les deux zones de cette région :

$$\text{zone 1 : } y_i = 1.5 + 2.5x_i \quad [2.2a]$$

$$\text{zone 2 : } y_i = 1.5 + 1.5x_i \quad [2.2b]$$

Appliquer l'équation [2.1] à la partie de la région pour laquelle l'équation [2.2a] est valable conduit à la sous-estimation des valeurs de y_i dans cette région et des erreurs estimées positives (pour des valeurs positives de y_i). Inversement, appliquer l'équation [2.1] à la partie de la région pour laquelle l'équation [2.2b] est valable conduit à la surestimation des valeurs de y_i dans cette région et des erreurs estimées négatives. Il y aura donc une forte autocorrélation spatiale des erreurs résultant de l'incapacité du modèle global à traiter l'instabilité structurelle de la relation.

Dans le même ordre d'idée, la modélisation de l'hétérogénéité par l'utilisation de variables augmentées est susceptible d'éliminer l'autocorrélation spatiale des erreurs estimées du modèle initial (Jones, 1983 ; Casetti et Jones, 1988). Anselin (1988) et Anselin et Griffith (1988) montrent en effet que, si des variables augmentées sont omises à tort, le test de Moran basé sur les erreurs estimées du modèle initial mal spécifié peut indiquer de l'autocorrélation spatiale, même si les erreurs ne sont pas spatialement autocorrélées. Le test de Moran semble donc puissant contre des alternatives indiquant de l'hétérogénéité spatiale.

Enfin, il faut souligner que les mauvaises spécifications du modèle et les erreurs de mesure qui peuvent causer de l'hétéroscédasticité sont aussi susceptibles de conduire à de l'autocorrélation spatiale des erreurs. Tous ces éléments suggèrent que les liens entre hétérogénéité spatiale et dépendance spatiale sont forts et complexes. Les deux aspects ne peuvent pas être considérés indépendamment l'un de l'autre et ils doivent être structurés très précisément pour pouvoir identifier les paramètres du problème et tester chaque type de mauvaise spécification. D'une part, la dépendance spatiale ne se teste plus de la même façon en présence d'hétéroscédasticité, connue ou non (paragraphe 22) et d'autre part, il faut adapter les tests d'hétéroscédasticité et d'instabilité structurelle à la présence d'autocorrélation spatiale (paragraphe 23 et 24).

22. Dépendance spatiale et hétéroscédasticité

Les tests de l'autocorrélation spatiale varient suivant que la forme précise de l'hétéroscédasticité est connue ou non.

221. Le test de l'autocorrélation spatiale en présence d'hétéroscédasticité de forme connue

Dans ce premier cas, il s'agit des modèles de régression linéaire avec une forme préspecifiée d'hétéroscédasticité : soit une hypothèse particulière est faite sur le modèle, soit ce cas de figure peut intervenir dans diverses situations, comme par exemple lorsque des

coefficients aléatoires sont inclus. Considérons le modèle avec autocorrélation des erreurs (sans variable autorégressive) et hétéroscédasticité.

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \quad u \sim N(0, \Omega) \end{aligned} \quad [2.3]$$

Si la forme de Ω est connue, Anselin a dérivé un test du multiplicateur de Lagrange sur le paramètre λ qui s'obtient à partir des résultats de l'estimation du modèle sous l'hypothèse nulle $H_0 : \lambda = 0$ (Anselin, 1988a, 1988b). Dans ce cas, la statistique du test du multiplicateur de Lagrange est :

$$[\hat{\varepsilon}'\Omega^{-1}W\hat{\varepsilon}]^2 / T^* \stackrel{as}{\sim} \chi^2(1) \quad [2.4]$$

$\hat{\varepsilon}$ est le vecteur contenant les erreurs estimées du modèle sous l'hypothèse nulle estimé par les MCO et $T^* = \text{tr}\{(W + W')W\}$.

Dans le modèle avec variables augmentées et hétéroscédasticité [1.33], la forme de l'hétéroscédasticité est aussi connue. Dans ce cas, Anselin (1992) a également proposé un test de l'autocorrélation spatiale (test du multiplicateur de Lagrange).

222. Le test de l'autocorrélation spatiale en présence d'hétéroscédasticité de forme inconnue

Souvent, l'hétéroscédasticité est susceptible d'être présente sans que sa forme précise soit connue. Par conséquent, le test précédent n'est pas applicable. De plus, un mauvais choix de la spécification de l'hétéroscédasticité diminue la puissance du test du multiplicateur de Lagrange. Il faut donc avoir un test robuste de l'autocorrélation spatiale en présence d'une hétéroscédasticité dont la forme précise est inconnue.

1/ Dans le cas général non spatial, Davidson et MacKinnon (1985) ont développé, comme extension de l'approche de White, un test de significativité des coefficients de régression robuste à l'hétéroscédasticité. Par exemple, ils testent l'hypothèse nulle $\gamma = 0$ dans le modèle suivant :

$$\begin{aligned} H_0 : y &= X\beta + u \\ H_1 : y &= X\beta + Z\gamma + u \end{aligned} \quad [2.5]$$

X et Z sont des matrices de variables explicatives exogènes de dimensions respectives (N, K) et (N, R) , β et γ sont des vecteurs de paramètres inconnus de dimensions respectives $(K, 1)$ et $(R, 1)$ et u est un terme aléatoire indépendant mais hétéroscédastique : $V(u_i) = \sigma_i^2$, pour $i = 1, \dots, N$. La statistique du test DM_1 pour ce cas est :

$$DM_1 = y' MZ (Z' M \Omega(u) MZ)^{-1} Z' M y \stackrel{as}{\sim} \chi^2(R) \quad [2.6]$$

M_x est la matrice de projection habituelle et $\Omega(u)$ est la matrice diagonale des erreurs estimées élevées au carré sous H_0 . Il apparaît que cette statistique est en fait égale à N moins la somme des carrés des erreurs estimées de la régression auxiliaire :

$$S = UMZ\gamma + \text{erreur} \quad [2.7]$$

S est le vecteur somme et U est la matrice diagonale des erreurs estimées des MCO.

2/ Dans le cadre précédent, un test de l'autocorrélation spatiale des erreurs peut être effectué en utilisant l'équation de Durbin (Anselin, 1990b) :

$$\begin{aligned} H_0 : y &= X\beta + u \\ H_1 : y &= X\beta + \lambda W_y - \lambda WX\beta + u \end{aligned} \quad [2.8]$$

Le modèle de l'hypothèse alternative doit être estimé par $y = X\beta + \lambda W_y + WX\gamma + u$ avec la contrainte $\lambda\beta = -\gamma$. L'estimation peut se faire par la méthode des variables instrumentales. Soit Q la matrice des instruments et Z la matrice de dimension $(N, K+1)$: $[W_y \ WX]$. Il faut prendre en compte les contraintes sur les paramètres explicitement dans une formulation non linéaire, estimée par la méthode des variables instrumentales non linéaire $y = f(\beta, \lambda)$, avec les dérivées :

$$\partial f / \partial \beta = X - \lambda WX \quad [2.9a]$$

et $\partial f / \partial \lambda = W_y - WX\beta \quad [2.9b]$

La statistique du test est alors :

$$DM_2 = (y - f)' MPF(\lambda) [F(\lambda)' PM\Omega(u) MPF(\lambda)]^{-1} F(\lambda)' PM(y - f) \stackrel{as}{\sim} \chi^2(1) \quad [2.10]$$

où $y - f$ sont les erreurs estimées des MCO, $\Omega(u)$ la matrice diagonale des erreurs estimées des MCO élevées au carré, $M = I - PF(\beta)[F(\beta)'PF(\beta)]^{-1}F(\beta)'P$, $P = Q(Q'Q)^{-1}Q'$, $F(\lambda)$ et $F(\beta)$ sont les matrices des dérivées partielles, évaluées sous l'hypothèse nulle $\lambda = 0$.

La statistique DM_2 est égale à N moins la somme des carrés des erreurs estimées de l'équation auxiliaire :

$$S = UMPF(\lambda)\gamma + erreur \quad [2.11]$$

où U est la matrice diagonale des erreurs estimées des MCO, $PF(\lambda)$ sont les erreurs estimées de la régression de $PF(\lambda)$ sur $PF(\hat{\beta})$ ou de $P(W_y - WX\hat{\beta})$ sur PX .

3/ Il est enfin possible de faire un test robuste à l'hétéroscédasticité pour la présence d'une variable endogène décalée (Anselin, 1990b). Dans ce cas, le test porte sur la nullité de ρ dans les modèles suivants :

$$\begin{aligned} H_0 : y &= X\beta + u \\ H_1 : y &= X\beta + \rho W_y + u \end{aligned} \quad [2.12]$$

La présence d'une variable endogène décalée nécessite une estimation par le maximum de vraisemblance ou les variables instrumentales. L'extension de la statistique dans ce dernier cas donne :

$$DM_3 = y' MPW_y [y' W' PM\Omega(u) MW_y]^{-1} y' W' PM_y \stackrel{as}{\sim} \chi^2(1) \quad [2.13]$$

où $P = Q(Q'Q)^{-1}Q'$, Q est une matrice d'instruments, $\Omega(u)$ est la matrice diagonale des erreurs estimées élevées au carré sous l'hypothèse nulle. Il s'avère que cette statistique est équivalente à N moins la somme des carrés des erreurs estimées de la régression auxiliaire :

$$S = UMPW\gamma\gamma + \text{erreur} \quad [2.14]$$

U est la matrice diagonale des erreurs estimées des MCO.

Aucun travail de simulation permettant d'établir les propriétés de biais et de puissance de ces tests en échantillon fini n'a encore été réalisé.

23. Le test de l'hétéroscédasticité en présence de dépendance spatiale

En présence d'autocorrélation spatiale, les propriétés des tests d'hétéroscédasticité ne sont plus valables (Anselin et Griffith, 1988). Face à ce problème, deux stratégies sont possibles (Anselin, 1988a, 1988b).

1/ *La première stratégie* consiste tout d'abord à faire un test joint d'hétéroscédasticité et d'autocorrélation spatiale (dans un modèle non-autorégressif) et ensuite, si l'hypothèse nulle est rejetée, à tester séparément ces deux effets.

Ainsi, on teste l'hypothèse jointe $H_0 : \lambda = 0$ et $\alpha_i = 0, i = 1, \dots, T$ dans le modèle suivant :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \quad u \sim N(0, \Omega) \end{aligned} \quad [2.15]$$

où les éléments de la diagonale principale de Ω est de forme [1.5] : $\sigma_i^2 = \sigma^2 f(\alpha_0 + \alpha_1 z_1 + \dots + \alpha_T z_T)$.

Le modèle [2.15] peut s'estimer avec la méthode du maximum de vraisemblance. La fonction de log-vraisemblance associée à ce modèle s'écrit:

$$\ln L(y|\beta, \lambda, \Omega) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln|\Omega| + \ln|I - \lambda W| - \frac{1}{2} v'v \quad [2.16a]$$

$$v'v = (y - X\beta)'(I - \lambda W)'\Omega^{-1}(I - \lambda W)(y - X\beta) \quad [2.16b]$$

La statistique du multiplicateur de Lagrange associée à l'hypothèse nulle jointe $H_0 : \lambda = 0$ et $\alpha_i = 0, i = 1, \dots, T$ dans [2.15] est la suivante (Anselin, 1988a) :

$$JLM = (1/2)\hat{f}'Z(Z'Z)^{-1}Z'\hat{f} + (1/T^*)[\hat{\varepsilon}'W\hat{\varepsilon}/\hat{\sigma}^2]^2 \stackrel{as}{\sim} \chi^2(T+1) \quad [2.17]$$

où Z est la matrice $(N, T+1)$ comportant les observations sur $(1, z_i)$, f est le vecteur $(N, 1)$ comportant les observations sur $f_i = (\hat{\sigma}^{-1}\hat{\varepsilon}_i) - 1$, $T^* = \text{tr}\{(W'+W)W\}$, $\hat{\varepsilon}$ et $\hat{\sigma}^2$ sont respectivement les erreurs estimées et la variance estimée du modèle de régression standard estimé par les MCO.

Cette statistique JLM est la somme de la statistique de Breusch-Pagan et de la statistique LM_{ERR} .

Kelejian et Robinson (1998) proposent une autre formulation de ce test qui n'est pas basée sur le principe du maximum de vraisemblance et qui n'exige pas a priori un modèle

linéaire, des erreurs normales et une forme d'hétéroscédasticité connue (seules les variables susceptibles d'être à l'origine de l'hétéroscédasticité doivent être connues à l'avance).

Des simulations de Monte-Carlo ont été effectuées par ces deux auteurs sur les performances de ces différents tests. Ces simulations suggèrent qu'en l'absence d'hétéroscédasticité, le test de Moran est le meilleur pour détecter l'autocorrélation spatiale même si les erreurs ne sont pas normales. Si le paramètre d'autocorrélation spatiale est élevé, le test de Moran, les tests LM_{ERR} et JLM et le test de Kelejian et Robinson (noté $KR-SPHET$) sont puissants qu'il y ait ou non de l'hétéroscédasticité. En revanche, si de l'hétéroscédasticité est présente, JLM et le test $KR-SPHET$ sont beaucoup plus puissants que les tests de Moran et LM_{ERR} pour détecter au moins un des deux problèmes : autocorrélation spatiale ou hétéroscédasticité.

En conclusion, si l'on pense que la présence d'hétéroscédasticité est peu probable, il est préférable d'utiliser le test de Moran et/ou le test LM_{ERR} . En revanche, s'il y a possibilité d'hétéroscédasticité, alors il vaut mieux utiliser les deux autres tests JLM et $KR-SPHET$. Si l'hypothèse nulle est rejetée, un test pour chaque cas doit être effectué afin de déterminer l'origine du rejet.

2/ La deuxième stratégie consiste à effectuer un test d'hétéroscédasticité en présence d'autocorrélation spatiale des erreurs. On teste donc $H_0 : \alpha_i = 0, i = 1, \dots, T$ dans [2.15]. Par conséquent, sous H_0 , ce modèle devient un modèle homoscedastique avec autocorrélation spatiale des erreurs :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \quad u \sim N(0, \sigma^2 I) \end{aligned} \quad [2.18]$$

Sous l'hypothèse nulle, la statistique du test (noté BPS comme "Breusch-Pagan Spatial") s'écrit de la façon suivante :

$$BPS = (1/2) \hat{f}' Z^{-1} [Z' \hat{D} Z]^{-1} Z' \hat{f} \stackrel{as}{\sim} \chi^2(T) \quad [2.19]$$

avec : $D_{(N,N)} = I - (1/2\sigma^4) dVd'$, $d_{(N,2)} = [S \ 2\sigma^2 w]$, $w_{(N,1)}$ est le vecteur composé des éléments diagonaux de $W(I - \lambda W)^{-1}$, $V_{(N,N)}$ est la matrice estimée des variances-covariances pour σ^2 et λ .

Ce test s'effectue après avoir estimé le modèle avec autocorrélation des erreurs [2.18] alors que le test précédent [2.17] pouvait s'effectuer à partir des erreurs estimées des MCO du modèle de régression standard. Un test de Breusch-Pagan ajusté peut aussi être dérivé pour le modèle avec variable autorégressive et hétéroscédasticité. Cependant, comme pour les tests précédents, les propriétés de ce test en échantillon fini ne sont pas connues.

24. Le test de l'instabilité des paramètres en présence de dépendance spatiale

Lorsqu'on teste l'hypothèse d'une différence significative entre diverses régions (ANOVA spatiale [1.13]) et s'il y a autocorrélation spatiale des erreurs, les MCO fournissent des estimateurs sans biais mais inefficients et l'inférence statistique n'est pas fiable. En particulier, dans le cas d'une autocorrélation spatiale positive, la variance intra-groupe est sous-estimée et l'hypothèse nulle d'égalité entre les moyennes des régions est rejetée trop souvent (Cliff et Ord, 1981 ; Legendre et al., 1990 ; Sokal et al., 1993). Il faut donc estimer le

modèle [1.13] par une méthode adaptée⁹ dans le but d'ajuster les sommes des carrés associés aux variances intra- et inter-groupe : par exemple, dans la table d'analyse des variances associée au modèle [1.13], les variables X et Y doivent être transformées en les prémultipliant par $I - \rho W$.

Plus généralement, lorsque les hypothèses d'homoscédasticité et de non-corrélation ne sont pas vérifiées, le modèle [1.13] doit être estimé comme un modèle hétéroscédastique (avec variances différentes), un modèle avec erreurs autocorrélées ou un modèle autorégressif (Griffith, 1992). De la même façon, il est nécessaire de tester la présence d'hétéroscédasticité et d'autocorrélation spatiale dans les modèles TSA [1.17] et VE [1.22] pour une inférence statistique fiable.

Enfin, dans le cas des régimes spatiaux et sous l'hypothèse d'homoscédasticité, le test de Chow, basé sur une statistique F permet de tester l'instabilité structurelle. En revanche, la présence d'autocorrélation spatiale nécessite son adaptation (Anselin, 1990a).

241. Les différents cas possibles

Formellement, reprenons les notations du modèle avec deux régimes [1.15] :

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

$$\Rightarrow y = X^* \beta + \varepsilon$$

Soit $\varepsilon' = [\varepsilon_1 \ \varepsilon_2]$ et Ψ sa matrice des variances-covariances : $\Psi = E(\varepsilon\varepsilon')$.

On veut tester l'hypothèse d'homogénéité régionale, soit : $H_0 : \beta_1 = \beta_2$. La forme de la statistique du test dépend des hypothèses faites sur les erreurs.

1^{er} cas : variance constante sur l'ensemble des observations

C'est le cas le plus simple :

$$\Psi = \sigma^2 I_N \tag{2.20}$$

où σ^2 est la variance constante et I_N est la matrice identité d'ordre N .

2^{ème} cas : hétéroscédasticité en blocs

On peut, comme dans la première section, envisager une structure d'hétéroscédasticité en blocs (modèle [1.8]) :

$$\Psi = \begin{bmatrix} \sigma_1^2 I_{N_1} & 0 \\ 0 & \sigma_2^2 I_{N_2} \end{bmatrix} \tag{2.21}$$

Plus généralement, lorsque l'autocorrélation spatiale est présente dans les erreurs, quatre situations différentes peuvent être distinguées selon la structure de la variance Ψ dans le

⁹ Voir Anselin et Bera (1998) et Le Gallo (2000) pour une présentation complète des méthodes d'estimation des modèles comportant de l'autocorrélation spatiale.

modèle [1.15]. Dans les deux premières, la dépendance des erreurs est reliée au système spatial complet alors que dans les deux dernières, la structure de la dépendance spatiale est différente pour chaque régime.

3^{ème} cas : *Homoscédasticité et autocorrélation spatiale*

Supposons une autocorrélation spatiale autorégressive pour le terme d'erreur complet ε :

$$\varepsilon = \lambda W \varepsilon + u \quad [2.22]$$

où λ est le paramètre autorégressif et u un vecteur de termes d'erreurs indépendants. Notons $B = (I - \lambda W)$ pour simplifier les écritures. On a alors $\varepsilon = B^{-1}u$.

Si les erreurs u sont homoscédastiques, soit $E(uu') = \sigma^2 I$, la matrice des variances-covariances complète prend alors la forme :

$$\Psi = \sigma^2 (B' B)^{-1} \quad [2.23]$$

4^{ème} cas : *Hétéroscédasticité en blocs et autocorrélation spatiale*

Reprenons la structure [2.22] pour ε , mais supposons le modèle hétéroscédastique où chaque régime a une variance des erreurs différentes. Dans ce cas, la matrice des variances-covariances pour u^* devient :

$$E[u u'] = \Omega = \begin{bmatrix} \sigma_1^2 I_1 & 0 \\ 0 & \sigma_2^2 I_2 \end{bmatrix} \quad [2.24]$$

$$\Rightarrow \Psi = B^{-1} \Omega (B')^{-1} \quad [2.25]$$

5^{ème} cas : *Processus spatiaux différents entre sous-groupes indépendants*

Lorsque la structure de dépendance spatiale est différente pour chaque sous-régime, cela se traduit pas le fait que les erreurs ε_1 et ε_2 dans [1.15] suivent chacune un processus spatial différent, réfléchés par les matrices de poids W_1 et W_2 respectivement :

$$\begin{aligned} \varepsilon_1 &= \lambda_1 W_1 \varepsilon_1 + u_1 \\ \varepsilon_2 &= \lambda_2 W_2 \varepsilon_2 + u_2 \end{aligned} \quad [2.26]$$

où λ_1 et λ_2 sont les coefficients spatiaux pour chaque régime.

Supposons tout d'abord que les erreurs des deux sous-régimes sont indépendantes : $E[u_{h1} u_{k2}] = 0$ pour tout (h, k) appartenant aux deux sous-régimes. Alors la matrice des variances-covariances de chaque sous-régime sera de la forme :

$$\Psi_i = \sigma_i^2 (B_i' B_i)^{-1} \text{ pour } i = 1, 2 \quad [2.27]$$

La matrice des variances-covariances totale sera :

$$\Psi = \begin{bmatrix} \sigma_1^2 (B_1' B_1)^{-1} & 0 \\ 0 & \sigma_2^2 (B_2' B_2)^{-1} \end{bmatrix} \quad [2.28]$$

6^{ème} cas : *Processus spatiaux différents entre sous-groupes dépendants*

Reprenons la structure [2.26] pour les erreurs. Lorsqu'il y a autocorrélation spatiale à la fois à l'intérieur et entre les différents régimes, alors les erreurs entre les deux sous-régimes ne sont plus indépendantes : $E[u_{h1}u_{h2}] = \sigma_{12}$. Dans ce cas, la variance totale est de la forme :

$$\Psi = \begin{bmatrix} B_1^{-1} & 0 \\ 0 & B_2^{-1} \end{bmatrix} \cdot \begin{bmatrix} \sigma_1^2 I_1 & \sigma_{12} SS' \\ \sigma_{12} SS' & \sigma_2^2 I_2 \end{bmatrix} \cdot \begin{bmatrix} (B_1')^{-1} & 0 \\ 0 & (B_2')^{-1} \end{bmatrix} \quad [2.29]$$

242. Les tests

Dans le premier cas [2.20], la variance est constante sur l'ensemble des observations et le test de Chow traditionnel s'applique. Ce test exact est construit à partir des erreurs estimées sous l'hypothèse nulle et sous l'hypothèse alternative :

$$F = \frac{(\hat{e}_c' \hat{e}_c - \hat{e}_L' \hat{e}_L) / K}{\hat{e}_L' \hat{e}_L / (N - 2K)} \sim F(K, N - 2K) \quad [2.30]$$

où \hat{e}_c est le vecteur des erreurs estimées du modèle contraint et \hat{e}_L celui du modèle libre. De la même manière, on peut effectuer ce test sur chacun des coefficients du modèle.

En revanche, ce test n'est plus valide dans tous les autres cas lorsque les erreurs ne sont pas homoscedastiques. Le test d'instabilité structurelle doit alors être établi sur des bases asymptotiques.

Pour $\Psi = \sigma^2 \Omega$, l'expression correspondante est :

$$C_G = \frac{\hat{e}_c' \hat{\Omega}^{-1} \hat{e}_c - \hat{e}_L' \hat{\Omega}^{-1} \hat{e}_L}{\hat{\sigma}^2} \stackrel{as}{\sim} \chi^2(K) \quad [2.31]$$

Ainsi, les tests pour toutes les formes de modèles présentés dans le paragraphe 242 peuvent être construits de cette façon, avec des estimations convergentes des paramètres du modèle. Par exemple, dans Anselin (1988a, 1990), le test de Chow spatial pour le modèle avec autocorrélation spatiale [2.22] et erreurs homoscedastiques [2.23] est de la forme :

$$C_s = \frac{\hat{e}_c' (I - \hat{\lambda} W)' (I - \hat{\lambda} W) \hat{e}_c - \hat{e}_L' (I - \hat{\lambda} W)' (I - \hat{\lambda} W) \hat{e}_L}{\hat{\sigma}^2} \stackrel{as}{\sim} \chi^2(K) \quad [2.32]$$

Les simulations de Monte-Carlo effectuées par Anselin (1990a) indiquent qu'ignorer la présence de l'autocorrélation spatiale affecte le biais et la puissance du test de Chow traditionnel. En particulier, les fréquences empiriques de rejet sont significativement différentes du seuil de signification choisi ($p = 0.05$) : elles sont supérieures ou inférieures en fonction de la partition de l'échantillon choisie (partition nord/sud ou centre/périphérie). La

puissance du test est également affectée mais différemment selon que le paramètre spatial est positif ou négatif.

En conclusion, il apparaît nécessaire de tester la présence d'une autocorrélation spatiale des erreurs avant d'effectuer un test d'instabilité structurelle et de la prendre en compte si elle est présente. En effet, la présence d'autocorrélation spatiale dans les analyses de régression invalide les résultats des tests de stabilité structurelle ; on pourrait par exemple conclure que deux régions appartenant à des régimes différents répondent différemment à une politique économique alors que ce n'est pas le cas.

3. UNE ILLUSTRATION

Le modèle étudié ici est destiné à illustrer certaines des techniques présentées dans ce papier. Ce modèle relie le nombre de délits (vols et cambriolages) au revenu et aux valeurs immobilières. Une partie des résultats est disponible dans Anselin (1988, 1990a) et LeSage (1999) mais nous avons produit d'autres résultats à partir de la base de données disponible sur le site Internet de SpaceStat¹⁰ et à partir de deux logiciels : d'une part le logiciel d'analyse des données spatiales SpaceStat 1.90 et d'autre part le logiciel de Système d'Information Géographique ArcView 3.2. Comme le souligne Anselin (1988), il ne s'agit pas de fournir une explication aux schémas spatiaux des délits mais d'illustrer l'interaction entre autocorrélation et hétérogénéité spatiales.

L'échantillon est composé de 49 observations pour 1980 sur des "census tracts" à Columbus, ville de l'Ohio. La variable DELIT est le nombre de cambriolages et de vols de voiture par milliers de ménage dans la zone. La variable de revenu (REV) et celle de valeurs immobilières (IMMO) sont exprimés en milliers de dollars :

$$\text{DELIT} = a + b\text{REV} + c\text{IMMO} + \varepsilon \quad [3.1]$$

Dans cet exemple, la dépendance spatiale est prise en compte à travers une matrice de contiguïté d'ordre 1 standardisée.

3.1. Estimation du modèle simple

La première étape consiste à estimer le modèle simple reliant le crime au revenu et aux valeurs immobilières. Les résultats de l'estimation de [3.1] sont donnés dans la première colonne du tableau 1. Tous les coefficients sont significatifs, à l'exception de c qui n'est pas significatif lorsqu'on applique la correction de White. Le test de Jarque-Bera ne permet pas de rejeter 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. Les tests suivants concernent des diagnostics pour l'autocorrélation spatiale des erreurs et l'hétéroscédasticité. Le test joint *JLM* indique la présence jointe de ces deux effets et donc des sources multiples de mauvaise spécification du modèle sont donc présentes.

- Concernant l'autocorrélation spatiale, si l'on applique la règle de décision de Anselin et Florax (1995), l'étude de la significativité des différents tests nous conduit à choisir un modèle autorégressif avec variable endogène décalée W_DELIT :

$$\text{DELIT} = a + b\text{REV} + c\text{IMMO} + \rho W_DELIT + \varepsilon \quad [3.2]$$

¹⁰ <http://www.spacestat.com>

- Concernant l'hétéroscédasticité, le test de White indique une hétéroscédasticité en général. Pour préciser la cause de cette hétéroscédasticité, nous avons effectué le test de Breusch-Pagan associé à plusieurs variables : les variables initiales REV et IMMO, la variable associée à l'aire de chaque voisinage, des variables dichotomiques Est/Ouest, Nord-Sud, centre-périphérie. De toutes ces variables, seul le test de Breusch-Pagan associé à la variable Est-Ouest (EO) permet de rejeter l'hypothèse nulle d'homoscédasticité.

Pour confirmer ce diagnostic, nous avons estimé [3.2] par la méthode du maximum de vraisemblance, en prenant en compte l'autocorrélation spatiale sous la forme d'une variable endogène décalée. Les résultats se trouvent dans la deuxième colonne du premier tableau. Tous les coefficients sont hautement significatifs, y compris la variable endogène décalée ce qui indique une forte concentration spatiale du nombre de crimes. Les coefficients structurels sont similaires à ceux estimés par les MCO mais plus petits en valeur absolue. Le modèle est meilleur que le précédent en termes des critères d'information (critère AIC d'Akai ke et SC de Schwartz) et le test LM_{ERR}^* ne permet pas de rejeter l'absence d'une autocorrélation spatiale des erreurs supplémentaire. En revanche, le test de Breusch-Pagan (BP) et sa version ajustée de l'autocorrélation spatiale (BPS) indiquent qu'il subsiste une certaine forme d'hétéroscédasticité liée à la localisation des zones dans l'Est ou dans l'Ouest de Columbus.

32. Méthode des variables augmentées

Une première façon d'incorporer l'hétérogénéité spatiale est de spécifier des variables augmentées. Ici, quatre nouvelles variables sont incluses, U_REV, V_REV, U_IMMO et V_IMMO, qui sont construites en multipliant les variables initiales REV et IMMO par les coordonnées u et v du centroïde de chaque zone. Dans cet exemple, nous avons estimé le modèle suivant :

$$\text{DELIT} = a + b_1\text{REV} + c_1\text{IMMO} + b_2\text{U_REV} + c_2\text{U_IMMO} + b_3\text{V_REV} + c_3\text{V_IMMO} + \varepsilon \quad [3.3]$$

Les résultats de l'estimation de ce modèle par les MCO (avec correction de White pour les écart-types estimés) sont donnés dans la première colonne du tableau 2. La prise en compte de cette forme particulière d'hétérogénéité spatiale n'a pas éliminé l'autocorrélation spatiale : les différents tests indiquent l'omission à tort d'une variable endogène décalée et l'inférence statistique effectuée sur le modèle [3.3] n'est donc pas fiable. Le modèle est alors réestimé avec une variable autorégressive :

$$\text{DELIT} = \rho\text{W_DELIT} + a + b_1\text{REV} + c_1\text{IMMO} + b_2\text{U_REV} + c_2\text{U_IMMO} + b_3\text{V_REV} + c_3\text{V_IMMO} + \varepsilon \quad [3.4]$$

Le coefficient autorégressif est significatif. Par rapport au modèle [3.3], on remarque que la prise en compte de l'autocorrélation spatiale a fait basculer le test global et les tests individuels sur les variables augmentées. Dans le modèle [3.3], le test global et les tests individuels (test F) sur les variables augmentées ne permettaient pas de rejeter l'hypothèse d'une stabilité des coefficients sur l'ensemble de la ville. En revanche, dans [3.4], les coefficients b et c ne peuvent plus être considérés comme stables sur toute la ville. Ainsi, les variables REV et IMMO sont affectées par une tendance Nord/Sud (U_REV et U_IMMO sont significatifs) et une tendance Est/Ouest (V_INC et V_HOUSE sont significatifs).

Pour chaque zone i , nous avons calculé les coefficients initiaux \hat{b}_i et \hat{c}_i comme fonction des coordonnées de chaque localisation i :

$$\hat{b}_i = \hat{b}_1 + \hat{b}_2 u_i + \hat{b}_3 v_i \quad [3.5a]$$

$$\hat{c}_i = \hat{c}_1 + \hat{c}_2 u_i + \hat{c}_3 v_i \quad [3.5b]$$

La représentation sur une carte de la variation spatiale des coefficients \hat{b}_i et \hat{c}_i illustre bien les tendances dont sont affectées les coefficients initiaux associés à REV et IMMO : plus les logements sont localisés dans le Sud-Est, plus l'effet négatif du revenu sur le nombre de délits est important. Au contraire, plus les logements sont situés au Nord-Ouest, plus la valeur immobilière a un effet négatif important.

Enfin, le modèle est bien spécifié. D'une part, le test de Breusch-Pagan et sa version ajustée ne permettent pas de rejeter l'hypothèse d'homoscédasticité. Une expansion aléatoire de la forme [1.33] n'est pas nécessaire et la forme déterministe [1.23] suffit. D'autre part, le test LM_{ERR}^* ne permet pas de rejeter l'absence d'une autocorrélation spatiale des erreurs supplémentaire. Ce modèle est également le meilleur que les précédents en terme des critères d'information.

33. Régimes spatiaux et hétéroscédasticité en groupe

Si l'on s'intéresse à la modélisation d'une hétérogénéité discrète, on doit passer à une analyse en termes de régimes spatiaux. Compte tenu des résultats obtenus par la méthode VE, on pourrait penser à spécifier 4 régimes : Nord-Est, Nord-Ouest, Sud-Est et Sud-Ouest. Pourtant, dans ce cas, le nombre d'observations dans chaque régime n'est pas suffisant (une douzaine). Nous nous sommes donc limités à spécifier des modèles à deux régimes. Pour une partition Nord-Sud, les tests d'instabilité structurelle et d'hétéroscédasticité en blocs n'ont jamais permis de rejeter l'hypothèse de coefficients et de variances variables selon les régimes, que l'autocorrélation spatiale soit prise en compte ou non. Nous avons donc uniquement considéré une partition Est-Ouest de la ville de Columbus.

Soit les deux variables muettes suivantes :

$$D_{1i} = \begin{cases} 1 & \text{si la zone } i \text{ est à l'Est} \\ 0 & \text{si la zone } i \text{ est à l'Ouest} \end{cases} \quad [3.6a]$$

et
$$D_{2i} = \begin{cases} 1 & \text{si la zone } i \text{ est à l'Est} \\ 0 & \text{si la zone } i \text{ est à l'Ouest} \end{cases} \quad [3.6b]$$

On estime le modèle suivant incorporant à la fois des régimes spatiaux et une hétéroscédasticité en groupe :

$$\text{DELIT} = a_1 D_1 + a_2 D_2 + b_1 D_1 \text{REV} + b_2 D_2 \text{REV} + c_1 D_1 \text{IMMO} + c_2 D_2 \text{IMMO} + \varepsilon \quad [3.7a]$$

avec
$$\varepsilon \sim \begin{bmatrix} \sigma_1^2 I_1 & 0 \\ 0 & \sigma_2^2 I_2 \end{bmatrix} \quad [3.7b]$$

Les résultats sont fournis dans les deux premières colonnes du tableau 3. Tous les coefficients sont significatifs à l'exception de celui associé à IMMO dans l'Ouest. Le test du ratio de vraisemblance portant sur l'égalité des variances indique que l'hypothèse de variances différentes par régimes ne peut pas être rejetée. En revanche, le test de Chow asymptotique (test de Wald) portant sur l'instabilité structurelle ne permet pas de rejeter l'hypothèse de coefficients identiques pour les deux régimes. Concernant les tests individuels, seul le coefficient associé à HOUSE peut être considéré comme différent selon le régime. Pourtant, l'inférence statistique faite sur ce modèle n'est pas fiable puisque les derniers tests indiqués dans le tableau indiquent l'omission à tort de l'autocorrélation spatiale sous la forme d'une variable endogène décalée.

Le dernier modèle estimé est alors le modèle [3.8] incluant une variable autorégressive :

$$\text{DELIT} = \rho \text{W_DELIT} + a_1 D_1 + a_2 D_2 + b_1 D_1 \text{REV} + b_2 D_2 \text{REV} + c_1 D_1 \text{IMMO} + c_2 D_2 \text{IMMO} + \varepsilon \quad [3.8a]$$

$$\text{avec } \varepsilon \sim \begin{bmatrix} \sigma_1^2 I_1 & 0 \\ 0 & \sigma_2^2 I_2 \end{bmatrix} \quad [3.8b]$$

La prise en compte de l'autocorrélation spatiale modifie substantiellement certains des résultats précédents. En particulier, le coefficient associé à REV dans l'Ouest n'est plus significatif. Certains tests d'instabilité structurelle basculent également : le test global devient significatif ainsi que le test individuel sur REV (à 10%). On obtient donc un schéma spatial distinct dans les deux régimes : à l'Est, c'est le revenu qui influence le plus le nombre de crimes commis alors qu'à l'Ouest, ce sont les valeurs immobilières qui jouent le rôle le plus important. Les variances sont toujours différentes d'un régime à l'autre et le coefficient spatial autorégressif est hautement significatif. Enfin, ce modèle est le meilleur du point de vue des critères d'information.

CONCLUSION

L'objectif de cet article est d'examiner les méthodes permettant de prendre en compte l'hétérogénéité spatiale sous la forme d'une hétéroscédasticité ou d'une instabilité des paramètres.

Dans le premier cas, la variance des erreurs n'est plus constante mais varie avec chaque observation localisée. Deux stratégies différentes peuvent alors être adoptées. La première est l'approche de White (1980) dans laquelle on utilise pour l'inférence statistique une estimation convergente de la matrice des variances-covariances basée sur les MCO en présence d'une hétéroscédasticité de forme inconnue. La seconde est de donner à l'hétéroscédasticité une forme spécifique, telle qu'une hétéroscédasticité en blocs. Lorsque cette spécification est adéquate, les tests associés sont plus puissants.

Dans le second cas, les coefficients de la régression ne sont plus constants dans l'espace. La variation géographique des paramètres d'une régression peut être discrète, auquel cas il faut spécifier des régimes spatiaux, ou continue, les paramètres sont alors supposés être une fonction des coordonnées de chaque localisation.

La présence simultanée d'autocorrélation spatiale et d'hétérogénéité spatiale dans un modèle nécessite des méthodes d'estimation et d'inférence adaptées. En ce qui concerne l'instabilité structurelle, le test de Chow traditionnel doit être modifié pour tenir compte d'une

autocorrélation spatiale dans les erreurs. Concernant l'hétéroscédasticité, on peut à nouveau adopter deux méthodologies différentes. D'une part, s'il est difficile de spécifier la forme de l'hétéroscédasticité, il faut alors utiliser les tests d'autocorrélation spatiale robuste à une forme inconnue d'hétéroscédasticité développés par Anselin (1990). D'autre part, si l'on utilise une forme particulière d'hétéroscédasticité, il est possible de tester conjointement ou séparément les deux effets spatiaux. Le test de Breusch-Pagan doit également être amendé pour tenir compte d'une éventuelle autocorrélation spatiale des erreurs.

Cependant, au contraire du cas où seule l'autocorrélation spatiale est envisagée, il n'existe pas à l'heure actuelle de règles de décision permettant de choisir entre les différents modèles combinant les deux effets spatiaux. En outre, les propriétés de plusieurs de ces tests en échantillon fini (biais et puissance) ne sont pas encore bien connues.

TABLEAU 1 : ESTIMATION DU MODELE SIMPLE

MODELE	Modèle simple	Modèle autorégressif
Estimation	<i>MCO</i> (écarts-types robustes)	<i>MV</i>
Constante	68,619 (0,000)	45,265 (0.000)
REV	-1,597 (0,001)	-1,036 (0.000)
IMMO	-0,274 (0,123)	-0,259 (0,003)
W_DELIT	-	0.423 (0.000)
Log-vrais	-187,377	-182,518
AIC	380,754	373,035
SC	386,430	380,603
$\hat{\sigma}^2$	130,759	95,723
Test de normalité		
<i>JB</i> (χ^2_2)	1,836 (0,399)	-
Tests joints hétéroscédasticité et autocorrélation spatiale		
<i>JLM</i> (χ^2_3)	13,715 (0,003)	-
Tests d'autocorrélation spatiale		
Moran	2,938 (0,003)	-
<i>LM_{ERR}</i> (χ^2_1)	5,815 (0,016)	-
<i>RLM_{ERR}</i> (χ^2_1)	0,127 (0,721)	-
<i>LM_{LAG}</i> (χ^2_1)	8,760 (0,003)	-
<i>RLM_{LAG}</i> (χ^2_1)	3,072 (0,002)	-
<i>SARMA</i> (χ^2_2)	8,887 (0,012)	-
<i>LM*_{ERR}</i> (χ^2_1)	-	0,505 (0,477)
Tests d'hétéroscédasticité		
White (χ^2_5)	19,947 (0,001)	-
<i>BP</i> (EO) (χ^2_1)	7,055 (0,008)	5,796 (0,016)
<i>BPS</i> (EO) (χ^2_1)	-	5,797 (0,016)

Notes : les "p-values" sont entre parenthèses, log-vrais est la valeur de la fonction de log-vraisemblance à l'optimum, AIC est le critère d'information d'Akai ke, SC est le critère d'information de Schwartz, JB est le test de Jarque-Bera de normalité des résidus, *JLM* est le test joint pour l'autocorrélation spatiale des erreurs et l'hétéroscédasticité, Moran est le test *I* de Moran, *LM_{ERR}* et *LM_{LAG}* sont les tests du multiplicateur de Lagrange respectivement pour l'autocorrélation spatiale des erreurs et la variable endogène décalée, *RLM_{ERR}* et *RLM_{LAG}* sont leurs équivalents robustes, *SARMA* est le test du multiplicateur de Lagrange pour la présence des deux effets spatiaux, *LM*_{ERR}* est le test du multiplicateur de Lagrange d'une autocorrélation des erreurs dans un modèle autorégressif, *BP* est le test de Breusch-Pagan, *BPS* est sa version ajustée.

TABLEAU 2 :

MODELE	Variables augmentées	Variables augmentées et autocorrélation spatiale
Estimation	<i>MCO</i> (écarts-types robustes)	<i>MV</i>
Constante	69,505 (0,000)	44.134 (0,000)
REV	-4,091 (0,029)	-1,939 (0,242)
IMMO	0,405 (0,532)	0,005 (0,994)
U_REV	-0,046 (0,244)	-0,064 (0,018)
U_IMMO	0,027 (0,096)	0,033 (0,001)
V_REV	0,121 (0,074)	0,094 (0,030)
V_IMMO	-0,049 (0,090)	-0,043 (0,009)
W_DELIT	-	0,464 (0,000)
log-vrais	-182,488	-176,010
AIC	378,977	368,019
SC	392,219	383,154
$\hat{\sigma}^2$	117,305	72,543
Test de normalité		
<i>JB</i> (χ^2_2)	1,888 (0,389)	-
Tests d'autocorrélation spatiale		
Moran	3,264 (0,001)	-
<i>LM_{ERR}</i> (χ^2_1)	4,139 (0,042)	-
<i>RLM_{ERR}</i> (χ^2_1)	0,601 (0,438)	-

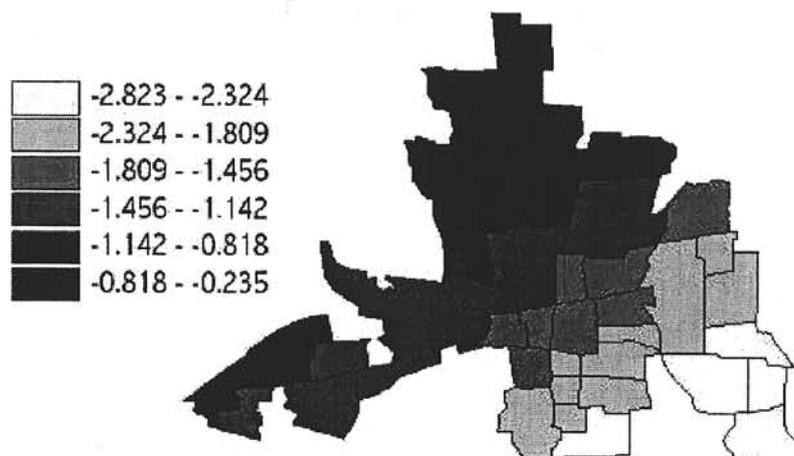
: METHODE VE

MODELE	Variables augmentées	Variables augmentées et autocorrélation spatiale
Tests d'autocorrélation spatiale (suite)		
$LM_{LAG} (\chi_1^2)$	12,153 (0,000)	-
$RLM_{LAG} (\chi_1^2)$	8,615 (0,003)	-
$SARMA (\chi_2^2)$	12,754 (0,002)	-
$LM_{ERR}^* (\chi_1^2)$	-	0,047 (0,828)
Test global sur les variables augmentées		
$(F_{4,4})$ et (χ_4^2)	5,521 (0,238)	15,032 (0,005)
Tests individuels sur les variables augmentées		
REV $(F_{4,2})$ et (χ_2^2)	3,193 (0,202)	8,739 (0,013)
IMMO $(F_{4,2})$ et (χ_2^2)	1,698 (0,428)	13,284 (0,001)
Tests d'hétéroscédasticité		
BP (REV et IMMO) (χ_2^2)	0,650 (0,722)	1,877 (0,391)
BPS (REV et IMMO) (χ_2^2)	-	1,877 (0,391)

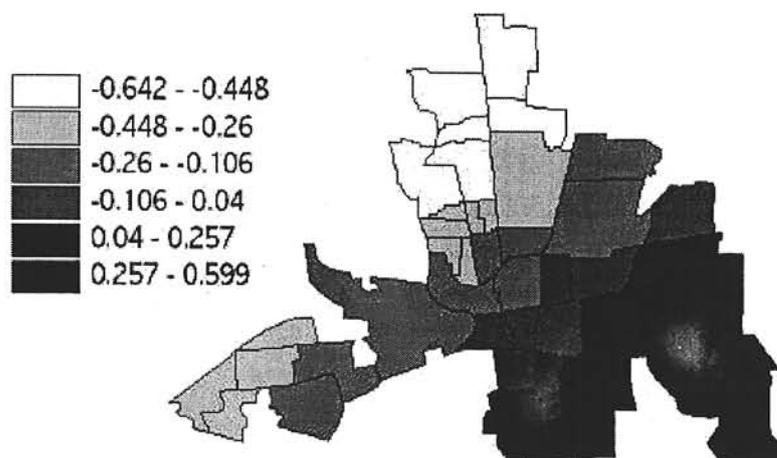
Notes : les "p-values" sont entre parenthèses, log-vrais est la valeur de la fonction de log-vraisemblance à l'optimum, AIC est le critère d'information d'Akai ke, SC celui de Schwartz, JB est le test de Jarque-Bera de normalité des résidus, Moran est le test I de Moran, LM_{ERR} et LM_{LAG} sont les tests du multiplicateur de Lagrange respectivement pour l'autocorrélation spatiale des erreurs et la variable endogène décalée, RLM_{ERR} et RLM_{LAG} leurs équivalents robustes, $SARMA$ est le test du multiplicateur de Lagrange pour la présence des deux effets spatiaux, LM_{ERR}^* est le test du multiplicateur de Lagrange d'une autocorrélation des erreurs dans un modèle autorégressif, les tests des variables augmentées sont effectués à partir d'un test de Chow (test exact ou asymptotique), BP est le test de Breusch-Pagan, BPS est sa version ajustée.

CARTES

Variation spatiale du coefficient associé à REV



Variation spatiale du coefficient associé à IMMO



TABEAU 3 : REGIMES SPATIAUX ET HETEROSCEDASTICITE EN GROUPE

MODELE	Régimes spatiaux et Hétéroscédasticité en blocs		Régimes spatiaux, hétéroscédasticité en blocs et autocorrélation spatiale	
	MV		MV	
Estimation	Ouest	Est	Ouest	Est
constante	76,650 (0,000)	67,294 (0,000)	50,058 (0,000)	45,337 (0,000)
REV	-1,455 (0,016)	-2,014 (0,000)	-0,652 (0,203)	-1,650 (0,000)
IMMO	-0,545 (0,004)	-0,064 (0,499)	-0,546 (0,000)	-0,011 (0,893)
W_CRIME	-		0,431 (0,000)	
$\hat{\sigma}^2$	182,805 (0,001)	57,932 (0,000)	129,07 (0,001)	42,609 (0,000)
R ²	0,6029		0,6611	
log-vrais	-180,472		-173,826	
AIC	372,943		351,652	
SC	384,294		374,895	
Test d'égalité des variances				
LR (χ_1^2)	7,944 (0,005)		7,137 (0,007)	
Test d'instabilité structurelle				
Chow (Wald) (χ_3^2)	5,35 (0,148)		10,053 (0,018)	
Tests de stabilité des coefficients individuels				
Constante (χ_1^2)	0,776 (0,378)		0,278 (0,598)	
REV (χ_1^2)	0,672 (0,412)		3,003 (0,083)	
IMMO (χ_1^2)	5,162 (0,023)		8,942 (0,003)	
Tests d'autocorrélation spatiale				
LM_{ERR} (χ_1^2)	5,972 (0,014)		-	
LM_{LAG} (χ_1^2)	11,267 (0,000)		-	

Notes : les "p-values" sont entre parenthèses, log-vrais est la valeur de la fonction de log-vraisemblance à l'optimum, AIC est le critère d'information d'Akai ke, SC est le critère d'information de Schwartz, LR est le test du ratio de vraisemblance pour tester l'égalité des variances dans les sous-régimes, les tests de stabilité des coefficients sont effectués à partir du test de Chow asymptotique. Enfin, LM_{ERR} et LM_{LAG} sont les tests du multiplicateur de Lagrange respectivement pour l'autocorrélation spatiale des erreurs et la variable endogène décalée.

BIBLIOGRAPHIE

- Agterberg F., 1984, Trend surface analysis, dans : Gaile G.L., Wilmot C.J. (Eds.), *Spatial Statistics and Models*, Reidel, Boston.
- 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., 1990a, Spatial dependence and spatial structural instability in applied regression analysis, *Journal of Regional Science*, 30, 185-207.
- Anselin L., 1990b, Some robust approach to testing and estimating in spatial econometrics, *Regional Science and Urban Economics*, 20, 141-163.
- Anselin L., 1992, Spatial dependence and spatial heterogeneity : model specification issues in the spatial expansion paradigm, dans : Casetti E., Jones J.P. III (Eds.), *Applications of the Expansion Method*, Routledge, Londres.
- Anselin L., 1995, Local indicators of spatial association-LISA, *Geographical Analysis*, 27, 93-115.
- Anselin L., 2000, Spatial econometrics, dans : Baltagi B. (Ed.), *Companion to Econometrics*, Basil Blackwell, Oxford.
- Anselin L., Bera A., 1998, Spatial dependence in linear regression models with an application to spatial econometrics, dans : Ullah A. et Giles D.E.A. (Eds.), *Handbook of Applied Economics Statistics*, Springer-Verlag, Berlin.
- Anselin L., Griffith D.A., 1988, Do spatial effects really matter in regression analysis?, *Papers of the Regional Science Association*, 65, 11-34.
- Bai J., Perron P., 1998, Estimating and testing linear models with multiple structural changes, *Econometrica*, 66, 47-78.
- Bailey T.C., Gatrell A.C., 1995, *Interactive Spatial Data Analysis*, Longman, Harlow.
- Bender B., Hwang H., 1985, Hedonic house price indices and secondary employment centers, *Journal of Urban Economics*, 17, 90-107.
- Bowman A.W., 1984, An alternative method for cross-validation for the smoothing of density estimates, *Biometrika*, 71, 353-360.
- Breusch T., Pagan A., 1979, A simple test for heteroskedasticity and random coefficient variation, *Econometrica*, 47, 1287-1294.
- Brueckner J.K., 1986, A switching regression analysis of urban population densities, *Journal of Urban Economics*, 19, 174-189.
- Brundson C., Fotheringham A.S., Charlton M., 1996, Geographically weighted regression : a method for exploring spatial nonstationarity, *Geographical Analysis*, 28, 281-298.
- Brundson C., Fotheringham A.S., Charlton M., 1998a, Spatial nonstationarity and autoregressive models, *Environment and Planning A*, 30, 957-973.
- Brundson C., Fotheringham A.S., Charlton M., 1998b, Geographically weighted regression-modelling spatial non-stationarity, *The Statistician*, 47, 431-443.
- Brundson C., Fotheringham A.S., Charlton M., 1999, Some notes on parametric significance tests for geographically weighted regression, *Journal of Regional Science*, 39, 497-524.
- 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.
- Casetti E., 1972, Generating models by the expansion method : applications to geographical research, *Geographical Analysis*, 4, 81-91.
- Casetti E., 1982, Drift analysis of regression parameters : an application to the investigation of fertility development relations, *Modeling and Simulation*, 3, 961-966.
- Casetti E., 1997, The expansion method, mathematical modeling, and spatial econometrics, *International Regional Science Review*, 20, 9-33.
- Casetti E., Can A., 1999, The econometric estimation and testing of DARP models, *Journal of Geographical Systems*, 1, 91-106.
- Casetti E., Jones J.P. III (Eds.), 1992, *Applications of the Expansion Method*, Routledge, Londres.

- Casetti E., Jones J.P. III, 1988, Spatial parameter variation by orthogonal trend surface expansions : an application to the analysis of welfare program participation rates, *Social Science Research*, 16, 285-300.
- Cleveland W.S., 1979, Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, 74, 829-836.
- Cleveland W.S., Devlin S.J., 1988, Locally weighted regression : an approach to regression analysis by local fitting, *Journal of the American Statistical Association*, 83, 596-610.
- Cleveland W.S., Devlin S.J., Grosse E., 1988, Regression by local fitting, methods, properties, and computational algorithms, *Journal of Econometrics*, 37, 87-114.
- 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., 1985, Heteroskedasticity-robust tests in regression directions, *Annales de l'INSEE*, 59-60, 183-217.
- Des Rosiers F., Thériault M., 1995, Household preferences and the changing structure of residential markets : combining hedonic modeling, GIS and spatial statistics, Document de travail 1995-32, Faculté des Sciences de l'Administration, Université Laval, Québec.
- Des Rosiers F., Thériault M., 1996, House prices, neighborhood factors and spatial dynamics: improving hedonic modeling using trend surface analysis, Document de travail 1996-04, Faculté des Sciences de l'Administration, Université Laval, Québec.
- Des Rosiers F., Thériault M., 1999, House Prices and Spatial Dependence : Towards an Integrated Procedure to Model Neighborhood Dynamics, Document de travail 1999-002, Faculté des Sciences de l'Administration, Université Laval, Québec.
- Dubin R.A., Sung R.A., 1987, Spatial variation in the price of housing : rent gradients in non-monocentric cities, *Urban Studies*, 24, 193-204.
- Efron B., 1982, *The Jackknife, the Bootstrap and Other resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphie.
- Fotheringham A.S., Brundson C., 1999, Local forms of spatial analysis, *Geographical Analysis*, 31, 340-358.
- Fotheringham A.S., Brundson C., Charlton M., 2000, *Quantitative Geography, Perspectives on Spatial Data Analysis*, Sage Publications, Londres.
- Fotheringham A.S., Charlton M., Brundson C., 1996a, The geography of parameter space : an investigation of spatial non-stationarity, *International Journal of Geographical Information Systems*, 10, 605-627.
- Fotheringham A.S., Charlton M., Brundson C., 1996b, Two techniques for exploring non-stationarity in geographical data, *Geographical Systems*, 28, 281-298.
- Fotheringham A.S., Charlton M., Brundson C., 1997, Measuring spatial variations in relationships with geographically weighted regression, dans : Fisher M.M., Getis A. (Eds.), *Recent Developments in Spatial Analysis*, Springer, Berlin.
- Fotheringham A.S., Charlton M., Brundson C., 1998, Geographically weighted regression : a natural evolution of the expansion method for spatial data analysis, *Environment and Planning A*, 30, 1905-1927.
- Goodman A.C., Kawai M., 1984, Functional form and rental housing market analysis, *Urban Studies*, 21, 367-376.
- Griffith D.A., 1978, A spatially adjusted ANOVA model, *Geographical Analysis*, 10, 296-301.
- Griffith D.A., 1988a, *Advanced Spatial Statistics, Special Topics in the Exploration of Quantitative Spatial Data Series*, Kluwer Academic Publishers, Dordrecht.
- Griffith D.A., 1992, A spatially adjusted N-way ANOVA model, *Regional Science and Urban Economics*, 22, 347-369.
- 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.
- Harvey A., 1976, Estimating regression models with multiplicative heteroskedasticity, *Econometrica*, 47, 387-406.

- Hildreth C., Houck J., 1968, Some estimators for a linear model with random coefficients, *Journal of the American Statistical Association*, 63, 584-595.
- Jones J.P. III, 1983, Parameter variation via the expansion method with tests for autocorrelation, *Modeling and Simulation*, 14, 853-857.
- Kau J., Lee C, Sirmans C., 1986, *Urban Econometrics : Model Developments and Empirical Results*, Greenwich, JAI Press.
- Kelejian H.H., Robinson D.P., 1998, A suggested test for spatial autocorrelation and/or heteroskedasticity and corresponding Monte Carlo results, *Regional Science and Urban Economics*, 28, 389-417.
- Koenker R, Bassett G., 1982, Robust tests for heteroskedasticity based on regression quantiles, *Econometrica*, 50, 43-61.
- Koenker, 1981, A note on studentizing a test for heteroskedasticity, *Journal of Econometrics*, 17, 107-112.
- Le Gallo J., 2000, Econométrie spatiale 1 : Autocorrélation spatiale, document de travail du LATEC n°2000-05, Université de Bourgogne, Dijon.
- Le Gallo J., Ertur C., 2000, Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995, document de travail du LATEC, n°2000-09, Université de Bourgogne, Dijon.
- Legendre P., Oden N.L., Sokal R.R., Vaudor A., Kim J., 1990, Approximate analysis of variance of spatially autocorrelated regional data, *Journal of Classification*, 7, 53-75.
- LeSage J.P., 1999, *Spatial Econometrics*, WebBook of Regional Science.
- Leung Y., Mei C., Zhang W., 2000, Statistical tests for spatial non-stationarity based on the geographically weighted regression model, *Environment and Planning A*, 32, 9-32.
- Mackinnon J., White H., 1985, Some heteroskedasticity-consistent covariance matrix estimators with improved finite samples properties, *Journal of Econometrics*, 29, 305-325.
- McMillen D.P., 1996, One hundred fifty years of land values in Chicago : A nonparametric approach, *Journal of Urban Economics*, 40, 100-124.
- McMillen D.P., McDonald J.F., 1997, A nonparametric analysis of employment density in a polycentric city, *Journal of Regional Science*, 37, 591-612.
- McMillen D.P., McDonald J.F., 1998, Population density in Chicago : a bid rent approach, *Urban Studies*, 7, 1119-1130.
- Olmo J.C., 1995, Spatial estimation of housing prices and locational rents, *Urban Studies*, 32, 1331-1344.
- Ord K., 1975, Estimation methods for models of spatial interaction, *Journal of the American Statistical Association*, 70, 120-126.
- Quandt R., 1958, The estimation of the parameters of a linear regression system obeying two separate regimes, *Journal of the American Statistical Association*, 53, 873-880.
- Ripley B., 1981, *Spatial Statistics*, New York, Wiley.
- Schroeder L.D., Sjoquist D.L., 1976, Investigation of population density gradients using trend surface analysis, *Land Economics*, 52, 382-392.
- Sokal R.R., Oden N.L., Thomson B.A., Kim J., 1993, Testing for regional differences in means: distinguishing inherent from spurious spatial autocorrelation by restricted randomization, *Geographical Analysis*, 25, 199-209.
- Thériault M., Des Rosiers F., Vandersmissen M.-H., 1999, GIS-based simulation of accessibility to enhance hedonic modeling and property value appraisal : an application to the Quebec city metropolitan area, Document de travail 1999-011, Faculté des Sciences de l'Administration, Université Laval, Québec.
- Upton G.J.G., Fingleton B., 1985, *Spatial Data Analysis by Example*. John Wiley, New York.
- White H., 1980, A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, 48, 817-838.