Croissance, capital humain et interactions spatiales : une étude économétrique

Cem Ertur et Kalidou Thiaw

Laboratoire d'Economie et de Gestion (UMR 5118 CNRS)

Université de Bourgogne

Pôle d'Economie et de Gestion

B.P. 26611 – 21066 Dijon Cedex

Juin 2005

Abstract

The aim of this paper is to analyze the theoretical and econometric implications of omitting spatial dependence in the Mankiw, Romer, and Weil (1992) model. Indeed, the international distribution of income levels and growth rates suggests the existence of large international disparities, and therefore the important role of location on economic performance. However, taking spatial dependence into account requires resorting to the methods of Spatial Econometrics, not only for a valid statistical inference, but also for revaluating the impact of the variables generally considered as crucial in the growth phenomenon and finding the processes underlying growth rates and income levels.

Keywords: *Economic growth, convergence, spatial econometrics*

JEL: C14, C31, O4

Résumé

L’objectif de cet article est d’analyser les implications économétriques et théoriques de l’omission de la dépendance spatiale dans le cadre de l’estimation du modèle de Mankiw, Romer et Weil (1992). En effet, la distribution internationale des taux de croissance et des niveaux de revenu suggère l’existence de bassins de croissance et de récession, et donc un rôle important de la localisation dans les performances économiques. Cependant, la prise en compte de l’autocorrélation spatiale dans l’estimation nécessite de recourir aux méthodes de l’économétrie spatiale afin, non seulement d’obtenir une inférence statistique valide, mais également d’élucider les processus qui sous-tendent la détermination des rythmes de croissance et des niveaux de revenu.

Mots-clés : Croissance économique, convergence, économétrie spatiale

Introduction

Dans leur article « A contribution to the empirics of economic growth »,Mankiw, Romer et Weil (1992) entreprennent de montrer que, malgré ses insuffisances, le modèle de Solow offre un cadre satisfaisant pour l'analyse de la croissance. Ils montrent en effet qu'en élargissant la notion de capital pour y intégrer le capital humain, puis en conservant l'idée d'un progrès technique exogène et de rendements d'échelle constants dans la production, il est possible d'expliquer une très grande partie des différences internationales de niveaux de revenu et de taux de croissance par tête.

Pour ce faire, Mankiw et *al.* (1992) reprennent les équations de base du modèle de Solow en y introduisant une mesure du capital humain. Cependant, tout comme dans la plupart des études empiriques sur la croissance, ils mettent en œuvre des régressions en coupe transversale estimées par la méthode des moindres carrés ordinaires, omettant entre autres toutes les influences sur le niveau de revenu ou sur le taux de croissance liées à la localisation des différents pays. Or l'observation des données internationales sur la croissance montre qu'il existe une certaine tendance au regroupement géographique des nations les plus riches, et qu'il en est de même pour les pays pauvres. Temple [1999] rappelle d'ailleurs que les variables muettes régionales se sont souvent révélées statistiquement significatives.

Les performances de croissance ne sont donc probablement pas insensibles à la localisation des pays, même si la littérature économique s'est très peu inspirée des questions que pose la prise en compte de l'espace du point de vue empirique. Walter Isard parlait d'ailleurs, déjà dans les années 50, de « biais anglo-saxon » en déplorant l'absence de la dimension spatiale dans l'analyse des phénomènes économiques. Ce n'est donc que très récemment que des études empiriques se sont proposé d'intégrer de manière explicite les effets de l'espace sur la croissance (Cf. Conley et Ligon, [2002] ; Easterly et Levine [1998], Ertur, Le Gallo and Baumont [2005], Fingleton [1999], Moreno et Trehan [1997]).

Pourtant, les problèmes liés à la manipulation des données spatiales sont depuis longtemps mis en évidence dans les modèles économétriques en science régionale. C'est d'ailleurs à cette discipline et plus précisément à J. Paelinck que l'on doit le terme d'« économétrie spatiale » qui, selon Anselin [1988], désigne « l'ensemble des techniques qui traitent des particularités liées à l'espace dans l'analyse statistique des modèles […]».

En fait, les données spatiales ou à référence géographique présentent la particularité de véhiculer une information portant non seulement sur la valeur observée d'une variable d'intérêt, mais aussi sur la localisation relative de l'unité d'observation. Dès lors, les régressions en coupe transversale, habituellement mises en œuvre dans les études empiriques sur la croissance, ne sont plus en mesure de fournir des résultats satisfaisants du fait de la présence potentielle de deux effets spatiaux : l'autocorrélation et l'hétérogénéité spatiales.

Tandis que l'hétérogénéité spatiale peut être généralement traitée grâce aux outils de l'économétrie standard, la présence de l'autocorrélation spatiale modifie substantiellement les propriétés des estimateurs des moindres carrés ordinaires ainsi que celles de l'inférence statistique fondée sur ces estimateurs. Ces estimateurs peuvent en effet être biaisés et/ou inefficients. Il devient alors indispensable de recourir aux techniques de l'économétrie spatiale.

Dans cet article, il s'agit donc tout d'abord de tester l'hypothèse de l'omission d'une autocorrélation spatiale dans les résidus d'estimation des spécifications de base du modèle MRW ; en effet, la distribution internationale des niveaux de revenu et des taux de croissance nous laisse penser que ces équations sont mal spécifiées du fait de la non prise en compte des effets de débordement géographique (Moreno et Trehan 1997). Ensuite, nous nous intéressons à la manière dont il convient de modéliser les relations entre pays lorsque l'on tient compte de leur position relative dans l'espace et recourons donc aux différentes spécifications et tests statistiques qu'offre l'économétrie spatiale. En dernier lieu, nous revenons sur les conclusions du modèle de Mankiw et *al.* (1992) afin d'évaluer l'influence de la localisation sur les déterminants de l'état stationnaire et sur les performances de croissance dans les pays considérés.

1. Le modèle de Mankiw, Romer et Weil

Le modèle de Mankiw et *al.* (1992) se fonde sur une fonction de production qui suit les hypothèses traditionnelles du modèle de Solow, et vérifie de ce fait, les conditions inhérentes à la technologie néo-classique : des productivités marginales positives et décroissantes (par rapport à chacun des facteurs de production), des rendements d'échelle constants (par rapport à l'ensemble des facteurs), les conditions d'Inada.

D'un point de vue formel, le modèle MRW se distingue cependant du modèle de Solow du fait de l'introduction d'une variable supplémentaire représentative du « capital humain » dans la fonction de production. On a ainsi :

$$Y(t) = F\big(K(t), H(t), A(t)L(t)\big) = K(t)^{\alpha} H(t)^{\beta} \big(A(t)L(t)\big)^{1-\alpha-\beta} \tag{1}$$

avec $\alpha > 0, \beta > 0, \alpha + \beta < 1$; α et β constants. Y représente le flux de production, K le stock de capital physique, H le stock de capital humain, A le niveau du progrès technique, et L l'offre de travail. Par ailleurs, le niveau de la technologie et l'offre de travail évoluent de manière exogène, plus précisément comme suit :

$$\begin{aligned} A(t) &= A(0)e^{gt} \\ L(t) &= L(0)e^{nt} \end{aligned} \tag{2}$$

Il en résulte que le nombre d'unités de travail efficace de l'économie croît au taux $n+g$. La pondération de la fonction de production globale (grâce à l'hypothèse des rendements constants) par l'inverse du nombre d'unités d'efficience permet d'en obtenir l'écriture sous une forme intensive, soit :

$$y = f(k,h) = k^{\alpha} h^{\beta} \tag{3}$$

Dès lors, le flux de production par unité d'efficience est fonction des stocks de capitaux physique et humain par unité d'efficience. Par ailleurs, à chaque période, des fractions s_k et s_h (s_k et s_h sont des constantes exogènes) du flux de production sont respectivement affectées à l'investissement brut en capital physique et en capital humain.

Dans le même temps, l'obsolescence du capital se traduit par la disparition, à chaque période, d'une fraction δ du stock de capital physique et humain. Le taux effectif de dépréciation du capital (physique et humain) est alors égal à $n+g+\delta$ car plus le nombre d'unités d'efficience augmente, plus le stock de capital diminue.

De façon formelle, on en déduit le système dynamique fondamental traduisant l'évolution des stocks de capitaux physique et humain par unité d'efficience. On obtient :

$$\begin{cases} \dot{k} = s_k k^{\alpha} h^{\beta} - (n+g+\delta)k \\ \dot{h} = s_h k^{\alpha} h^{\beta} - (n+g+\delta)h \end{cases} \tag{4}$$

Par ailleurs, on caractérise l'état stationnaire de l'économie en annulant chacune des équations différentielles du premier ordre ci-dessus (à l'état stationnaire, les stocks de capitaux physique et humain ne varient plus). La solution du système qui en résulte définit alors l'équilibre de long terme de l'économie $(\overline{E})$, et correspond aux valeurs suivantes pour k et h :

$$\overline{k} = \left[\frac{s_k^{1-\alpha} s_h^{\alpha}}{n+g+\delta}\right]^{\frac{1}{1-\alpha-\beta}} \quad \text{et} \quad \overline{h} = \left[\frac{s_k^{1-\beta} s_h^{\beta}}{n+g+\delta}\right]^{\frac{1}{1-\alpha-\beta}} \tag{5}$$

Dès lors, de l'expression de la fonction de production intensive sous forme logarithmique, et pour les valeurs d'état stationnaire ($\bar{k}$ et $\bar{h}$), on tire l'« équation de revenu » qui représente la première équation estimable du modèle de Mankiw, Romer, et Weil, soit[1] :

$$\ln \hat{y} = a + \frac{\alpha}{1-\alpha-\beta} \ln s_k + \frac{\beta}{1-\alpha-\beta} \ln s_h - \frac{\alpha+\beta}{1-\alpha-\beta} \ln(n+g+\delta) + \varepsilon \qquad (6)$$

Cette équation explique le niveau de revenu par tête de long terme par les taux d'accumulation des stocks de capitaux physique et humain, et par le taux effectif de dépréciation du capital. Dès lors, Mankiw et *al.* se fondent sur cette spécification pour montrer que lorsqu'on élargit la définition du capital pour lui rajouter le capital humain, l'amplitude de la réaction du niveau de revenu de long terme au taux d'épargne et au taux de croissance démographique en est plus importante. Les élasticités de la production par tête par rapport à ces dernières variables sont en effet plus fortes dans le modèle de Solow que dans le modèle originel, ce qui conduit Mankiw et *al.* à considérer que leur modèle est en mesure de rendre compte des grandes disparités de revenus observées à l'échelle internationale[2].

Leur analyse repose alors sur l'« équation de convergence » suivante :

$$\ln\left[\frac{y(T)}{y(0)}\right] = a + \theta\frac{\alpha}{1-\alpha-\beta} \ln s_k + \theta\frac{\beta}{1-\alpha-\beta} \ln s_h - \theta\frac{\alpha+\beta}{1-\alpha-\beta} \ln(n+g+\delta) - \theta \ln y(0) + \varepsilon \quad (7)$$

Dans la mesure où l'équation du revenu met en évidence le fait que des économies disposant de taux d'épargne et de taux de croissance démographique substantiellement différents connaissent des niveaux de revenu de long terme tout aussi différents, Mankiw et *al.* (1992) soutiennent que l'hypothèse d'une convergence « absolue », longtemps retenue à propos du modèle de Solow, ne tient plus et qu'il convient plutôt de retenir l'idée d'une convergence « conditionnelle » lorsqu'il s'agit de rendre compte des expériences de croissance au niveau international.

Mankiw et *al.* (1992) montrent par ailleurs que l'estimation de l'équation de convergence fournit des résultats en conformité avec les prédictions du modèle. En effet, en retenant la valeur standard de $1/3$ pour les paramètres α et β, et un taux de croissance démographique de 1%, on peut s'attendre à un taux de convergence de 2% (ce résultat est

---

[1] L'« équation de convergence » s'obtient en procédant à une approximation log-linéaire du taux de croissance autour de l'état stationnaire, c'est-à-dire : $\gamma_y = \frac{d \ln y(t)}{dt} = -\lambda \ln\left[\frac{y(t)}{\bar{y}}\right]$, avec : $\lambda = (1-\alpha-\beta)(n+g+\delta)$. L'équation (7) s'obtient alors de façon assez simple de la solution de ce système d'équation différentielle du premier ordre.

[2] Les estimations des parts du capital, soit α et β, sont d'ailleurs conformes aux prédictions du modèle.

d’ailleurs très courant dans la littérature) tandis que l’on aurait obtenu un taux de 4% dans le cadre du modèle de Solow. De plus, les estimations de α et β résultant de la version contrainte du modèle correspondent sensiblement à ce qu’en suggère la version libre de ce dernier.

Dans la suite de cet article, notre analyse reposera donc sur ces deux équations estimables qui constituent en réalité le socle du modèle MRW. Celui-ci constitue en effet un cadre d’analyse relativement satisfaisant du processus de croissance, même s’il n’en fournit pas d’explication à long terme comme ne manquent pas de le rappeler les théoriciens de la croissance endogène. Mankiw (1995) souligne à ce propos que l’enjeu pour le modèle de croissance néoclassique peut-être plutôt d’être apte à rendre compte de la grande diversité des expériences en matière de croissance économique.

2. L’autocorrélation Spatiale

Selon Anselin et Bera (1998), l’autocorrélation spatiale peut être définie de manière générale comme la correspondance entre la similarité des valeurs prises par une variable d’intérêt et la proximité des unités spatiales où ces mêmes valeurs sont observées. Plus précisément, elle traduit l’existence d’une relation fonctionnelle entre les observations faites au niveau des différentes localisations de l’espace étudié.

La présence d’autocorrélation spatiale est largement liée à la nature bidimensionnelle des données spatiales ou à référence géographique, et au caractère multidirectionnel des relations dans l’espace (par opposition à l’axe unidirectionnel du temps). En effet, pour chaque unité d’observation, on dispose d’une information aussi bien sur la valeur prise par la variable d’intérêt que sur la localisation de ladite unité. Dès lors, deux unités d’observation différentes peuvent être corrélées uniquement en raison de leur position géographique, et ce phénomène peut intervenir dans toutes les directions.

D’un point de vue formel, la présence d’autocorrélation spatiale entre deux localisations i et j quelconques s’exprime par une covariance non nulle des valeurs prises par la variable d’intérêt en ces deux localisations[3].On a ainsi :

$$Cov\ \left(y_i y_j\right) = E\left(y_i y_j\right) - E\left(y_i\right)E\left(y_j\right) \neq 0 \quad ; \ avec \quad i \neq j \tag{8}$$

où y_i *et* y_j représentent les valeurs de la variable d’intérêt, respectivement en i et j. Il faut cependant noter que cette covariance ne revêt un caractère véritablement spatial que lorsque

[3] Le coefficient estimé associé au logarithme du niveau de revenu par tête initial permet d’obtenir une estimation de la vitesse de convergence des économies vers leurs états stationnaires respectifs. On montre que cette vitesse s’écrit de la manière suivante : $\hat{\lambda} = -(1/T)\ \left[\log\left(1-\hat{\theta}\right)\right]$.

la distribution des unités considérées peut s'interpréter en termes de structure, d'interaction ou d'arrangement spatial des observations. Les unités en question peuvent correspondre à des points, des régions ou des zones géographiques.

Sur le plan strictement économétrique, la non nullité de la covariance remet en question une hypothèse centrale du principe des moindres carrés ordinaires (largement mis en œuvre dans les études empiriques sur la croissance), à savoir la sphéricité des erreurs. Il s'ensuit qu'en présence d'autocorrélation spatiale des erreurs, les estimateurs des MCO se révèlent inefficients selon la structure des interactions spatiale existantes entre les différentes localisations.

De fait, dans l'approche retenue par l'économétrie spatiale, cette structure repose sur la définition en amont d'un processus rendant compte de la distribution des unités spatiales, et qui conditionne subséquemment la forme fonctionnelle des covariances inter–individuelles. Il en est autrement dans le cadre de la géostatistique où la structure des covariances est imposée a priori.

Cependant, du point de vue de l'estimation, il se pose un problème d'identification dans la mesure où les n observations disponibles de l'échantillon ne permettent pas d'estimer les n variances individuelles et les $n(n-1)/2$ covariances inter–individuelles. Dès lors, l'économétrie spatiale offre des outils tels que les matrices de poids spatiaux et les variables spatiales décalées qui permettent de prendre en charge ce type de problèmes.

2.1 Matrices de poids et variables décalées

Le concept de matrice de poids constitue un élément fondamental en Econométrie Spatiale car non seulement, il permet de résoudre les problèmes d'estimation que pose la nature bidimensionnelle et multidirectionnelle des données spatiales, mais encore il permet de définir la topologie de l'espace étudié (en permettant de situer les unités spatiales les unes par rapport aux autres) et le poids relatif de chacune des unités spatiales qui le composent.

D'après la définition de l'autocorrélation spatiale que donnent Anselin et Bera (1991), on peut aisément imaginer le rôle important de notions telles que celles de « proximité » et de « voisinage » dans la schématisation de l'espace et plus largement dans la modélisation de l'autocorrélation spatiale; cependant ces notions peuvent revêtir plusieurs interprétations.

De manière générale, deux principales conceptions ont été retenues pour la détermination des éléments de la matrice de poids, respectivement fondées sur le principe de contiguïté et sur celui de distance.

Les premières mesures de dépendance spatiale sont dues à Moran (1948) et Geary (1954), et reposent sur le concept de contiguïté binaire. Elles ont d'ailleurs conduit à l'utilisation des matrices dites de contiguïté qui reposent sur le partage d'une frontière commune entre unités spatiales. Autrement dit, deux localisations sont dites contiguës dès lors qu'elles sont voisines.

De manière formelle, une matrice de contiguïté représente chacune des localisations du système spatial en ligne et en colonne. Les « poids spatiaux » (éléments de la matrice de poids) sont obtenus en affectant la valeur 1 à chaque fois que l'on est en présence de deux localisations contiguës et la valeur 0 dans le cas contraire, une localisation donnée ne pouvant être contiguë avec elle-même. Pour une région i quelconque et l'ensemble J de ses voisines, les poids w_{ij} de la matrice de contiguïté $W = (w_{ij})$ sont donc définis de la manière suivante :

$$w_{ij} = \begin{cases} 1 \, pour \quad j \in J \\ 0 \, pour \quad j \notin J \end{cases} ; \quad w_{ii} = 0 \quad , \quad \forall i \qquad (9)$$

L'idée d'un voisinage fondé exclusivement sur la notion de contiguïté présente cependant un certain nombre de limites. En effet, la contiguïté binaire ne présente l'organisation des unités du système spatial que d'une façon relativement sommaire, et par conséquent, ne permet pas de capter la réelle intensité des relations de dépendance pouvant exister entre unités spatiales. La notion de contiguïté cesse d'ailleurs d'être immédiate dès que l'on se trouve devant une disposition spatiale régulière.[4]

En revanche, le principe des matrices de distance repose sur l'idée générale selon laquelle deux unités spatiales connaissent une interaction d'autant plus forte (*resp*. faible) que la distance qui les sépare est courte (*resp*. longue). Cliff et Ord (1973, 1981) ont été à l'origine de ce type de spécification des poids spatiaux en combinant notamment, une fonction de l'inverse de la distance séparant deux localisations et la longueur relative de leur frontière commune. Les éléments de la matrice correspondante s'écrivent alors comme suit :

$$w_{ij} = \begin{cases} (d_{ij})^{-a} (\beta_{ij})^{b} ; & \forall i \neq j \\ 0 ; & \forall i = j \end{cases} \qquad (10)$$

où d_{ij} représente la distance entre les unités spatiales i et j, β_{ij} la part relative de la frontière commune entre les localisations i et j dans le périmètre de la région i. a et b sont des paramètres fixés a priori.[5]

Cependant, les spécifications les plus courantes aujourd'hui dans les études empiriques mettent en œuvre des expressions plus simples pour les poids spatiaux. En effet, les matrices de poids utilisées reposent très souvent sur une fonction exponentielle négative ou sur l'inverse de

[4] En guise d'illustration, on fait souvent référence aux cas du *fou*, de la *tour* et de la *dame* par analogie au Jeu d'Echecs, et au déplacement de ces pièces sur l'échiquier. (Voir Anselin, 1988).
[5] Contrairement aux matrices de contiguïté binaire, la matrice ainsi obtenue n'est pas symétrique.

la distance séparant deux unités spatiales i et j quelconques. De manière formelle, on a alors respectivement :

$$\text{(a)} \quad \begin{cases} w_{ij} = e^{-\alpha d_{ij}}, & \forall i \neq j \\ w_{ii} = 0 \end{cases}$$

$$\text{(b)} \quad \begin{cases} w_{ij} = \left(d_{ij}\right)^{-\beta}, & \forall i \neq j \\ w_{ii} = 0 \end{cases} \qquad (11)$$

avec α et β des paramètres fixés a priori.

Ces différents types de poids spatiaux peuvent par ailleurs être généralisés soit en fixant une distance-seuil au-delà de laquelle toute interaction disparaît, soit en fixant a priori un nombre k de localisations appartenant au voisinage de chaque unité spatiale (l'interaction avec les localisations hors voisinage étant supposée nulle) et on parle alors de « matrice des k plus proches voisins ».

L'utilisation des matrices de distance nécessite précisément de définir a priori un critère de distance. La distance géographique est sans doute la première qui vient à l'esprit (distance euclidienne, distance sphérique) mais d'autres critères, fondés notamment sur des variables économiques ou sociales, ont été proposés dans la littérature avec le souci de mieux appréhender les relations interindividuelles dans l'espace (Cf. Case [1993], Conley et Ligon [2001]). Ce type de distance pose cependant de manière cruciale la question de l'exogénéité des poids spatiaux.

Par ailleurs, les matrices de poids sont habituellement standardisées en lignes afin de faciliter l'interprétation des paramètres spatiaux après estimation. Ainsi, chaque ligne i d'une matrice de poids W quelconque y est divisée par la somme des éléments w_{ij} qui la composent, et les poids spatiaux qui en résultent s'écrivent :

$$w_{ij}^{\ s} = \frac{w_{ij}}{\sum_j w_{ij}} \quad , \ \forall i \qquad (12)$$

La standardisation de la matrice de poids présente également l'avantage de permettre la comparaison des paramètres spatiaux issus de modèles différents dans la mesure où les poids d'une matrice standardisée expriment des mesures relatives et non plus absolues.

Le rôle essentiel de la matrice de poids peut également être apprécié du point de vue de la notion de « variable décalée » qui, pour une localisation i et une variable d'intérêt y quelconques, renvoie la moyenne pondérée (par les poids spatiaux) des observations associées aux localisations voisines j. Elle synthétise ainsi l'information relative au voisinage de chaque localisation, et s'obtient en prémultipliant y (vecteur des valeurs prises par la variable d'intérêt au niveau de chaque localisation) par la matrice de poids W.

Ses éléments peuvent s'écrire de la manière suivante :

$$[Wy]_i = \sum_{j \in J} w_{ij} y_j , \qquad (13)$$

w_{ij} représentant le $j^{ème}$ élément de la ligne i dans la matrice W. Les décalages spatiaux peuvent être définis pour des ordres supérieurs de la matrice de poids et correspondent alors à des moyennes pondérées des valeurs observées pour les « voisins des voisins ».

2.2 Le test *I* de Moran

Le test I d'absence d'autocorrélation spatiale de Moran (1948,1950) est le premier test de spécification ayant été proposé en économétrie spatiale et constitue de manière générale la première étape dans la recherche du processus le plus à même de représenter les données. En effet, ce test permet de se prononcer sur l'opportunité d'une modélisation intégrant les effets spatiaux même si, en lui-même, il ne donne pas d'indication sur la forme potentielle de l'autocorrélation spatiale.[6] La statistique du test se présente comme suit[7] :

$$I = \frac{n}{s}\left(\frac{\varepsilon' W \varepsilon}{\varepsilon' \varepsilon}\right) \qquad (14)$$

avec ε le vecteur des résidus de l'estimation du modèle non spatial (modèle statistique de base) par les moindres carrés ordinaires, W la matrice de poids spatiaux, n la taille de l'échantillon, et s un facteur de standardisation correspondant à la somme des éléments de W. Cliff et Ord (1981) ont montré que sous l'hypothèse d'une distribution normale pour le vecteur ε, et sous l'hypothèse nulle, l'espérance mathématique et la variance de la statistique I s'écrivent[8] :

$$E(I) = \frac{n}{s} tr \frac{(MW)}{n-K} \qquad (15)$$

Et :

$$V(I) = \left(\frac{n}{s}\right)^2 \frac{\left\{ tr(MWMW') + tr(MW)^2 + [tr(MW)]^2 \right\}}{(n-K)(N-K+2)} \qquad (16)$$

Dès lors, ils en dérivent la distribution asymptotique de la statistique I, et le test repose en définitive sur la variable centrée et réduite Z_I telle que :

---

[6] Le test de Moran ne spécifie pas d'hypothèse alternative à l'absence d'autocorrélation spatiale.

[7] Pour une matrice de poids standardisée en lignes, on a : $s = n$ et la statistique de test se récrit : $I = (\varepsilon' W \varepsilon / \varepsilon' \varepsilon)$. De par sa construction, cette statistique est très proche de celle du test d'autocorrélation temporelle de Durbin et Watson (1950).

[8] D'autres approches telles respectivement dites de « permutation » et de « randomisation » ont également été proposées en vue de l'inférence statistique sur le I de Moran (*Cf.* Cliff et Ord [1981]).

$$Z_I = \frac{I - E(I)}{\sqrt{V(I)}} \rightarrow \mathrm{N}(0,1) \qquad (17)$$

L'hypothèse d'absence d'autocorrélation spatiale globale est rejetée lorsque les résidus des MCO conduisent à une valeur de Z_I supérieure à la valeur seuil de la loi normale standard.[9]

2.3 Les spécifications économétriques

2.3.1 Modèle autorégressif spatial

Le modèle autorégressif spatial résulte de l'introduction d'une variable endogène décalée parmi les régresseurs du modèle linéaire standard. Il est le plus souvent mis en œuvre lorsqu'on est en présence d'un schéma d'interaction spatiale résultant d'un modèle théorique (Cf . Case et $al.$ [1993], Moreno et Trehan [1997]). D'un point de vue formel, on a la spécification suivante :

$$y = \rho Wy + X\beta + \varepsilon \qquad (18)$$

Soit, sous une forme réduite :

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \qquad (19)$$

Cette écriture implique que l'autocorrélation spatiale intervient à travers la corrélation existant entre le décalage spatial Wy et le terme d'erreur ε. En fait, contrairement au cas temporel où le retard y_{t-1} n'est corrélé avec ε_t que lorsque ce dernier est lui-même autocorrélé, la corrélation entre Wy et ε est indépendante de la distribution de ε. En effet, lorsque le vecteur des erreurs ε est tel que ses éléments $\varepsilon_i \rightarrow i.i.d(0, \sigma^2)$, l'espérance mathématique de y est donnée par :

$$E(y) = (I - \rho W)^{-1} X\beta \qquad (20)$$

On montre alors que la matrice de variances-covariances de y s'écrit:

$$V(y) = \sigma^2 \left[(I - \rho W)' (I - \rho W) \right]^{-1} \qquad (21)$$

Or cette matrice est pleine, ce qui traduit le fait que toutes les localisations sont reliées entre elles, et donc l'existence d'une autocorrélation spatiale entre les observations[10]. Par ailleurs, le schéma d'interaction totale que traduit la matrice de variances-covariances peut être décomposé

[9] Burridge (1980) a montré que le test de Moran est asymptotiquement équivalent au test du multiplicateur de Lagrange, tandis que King (1981) montre qu'il s'agit d'un test « localement meilleur invariant ».

[10] En effet, l'écriture du modèle sous sa forme réduite (19) n'est possible que lorsque la matrice $(I - \rho W)$ est non singulière $i.e$ lorsque $|I - \rho W| \neq 0$. Or on montre que cette condition est vérifiée pour $|\rho| \neq 0$ et lorsque $1/\rho$ n'est pas une valeur propre de la matrice de poids W.

en deux effets distincts, cela en récrivant la matrice $(I-\rho W)^{-1}$ sous la forme d'une somme infinie dans la forme réduite (19). On a alors :

$$y=\left(I+\rho W+\rho^2W^2+...\right)X\beta+\left(I+\rho W+\rho^2W^2+...\right)\varepsilon \qquad (22)$$

Le premier élément de la somme dans le membre de droite décrit un *effet multiplicateur spatial* qui signifie qu'au niveau de chaque localisation, (y) dépend des variables explicatives en ladite localisation, mais aussi de celles associées à toutes les autres localisations du système spatial.
En revanche, le second élément décrit un *effet de diffusion spatial* tel que tout choc exogène (ε) issu d'une unité spatiale donnée y affecte la variable dépendante, mais s'étend également à l'ensemble des unités spatiales. Ces deux effets s'atténuent au fur et à mesure que l'ordre de voisinage augmente.

2.3.2 Modèle à autocorrélation spatiale des erreurs

Cette spécification repose sur le rejet de l'hypothèse d'erreurs sphériques dans le modèle linéaire standard, et sur l'adoption d'un processus spatial pour le terme d'erreur ε. Plusieurs types de processus peuvent alors être utilisés, mais la spécification autorégressive spatiale est la plus usitée[11]. On ainsi :

$$\begin{cases} y=X\beta+\varepsilon \\ \varepsilon=\lambda W\varepsilon+u \end{cases} \qquad (23)$$

où λ représente le coefficient autorégressif spatial associé au décalage spatial du terme d'erreur ε, et u un vecteur d'erreurs homoscédastiques. La spécification peut être récrite sous sa forme réduite, et l'on obtient alors[12] :

$$y=X\beta+(I-\lambda W)^{-1}u \qquad (24)$$

Pour un vecteur $u \to i.i.d\left(0,\sigma^2 I\right)$, l'espérance mathématique de y s'écrit : $E(y)=X\beta$, et on en déduit l'expression de la matrice de variances-covariances :

$$V(y)=V(\varepsilon)=\sigma^2\left[(I-\lambda W)'(I-\lambda W)\right]^{-1} \qquad (25)$$

[11] Des spécifications alternatives telles le processus moyenne mobile (Cliff et Ord [1981], Haining [1978, 1990]) et le processus de Kelejian et Robinson [1993,1995] ont également été proposées pour les erreurs. Leur usage demeure cependant peu courant dans les études empiriques. Elles reposent de manière générale sur une décomposition du terme d'erreur en une partie liée aux chocs spécifiques à chaque localisation et en une partie correspondant à une moyenne pondérée des erreurs voisines.

[12] De façon analogue au modèle à variable endogène décalée, cette forme réduite est définie lorsque la matrice $(I-\lambda W)$ est non singulière, et λ doit alors vérifier les mêmes propriétés que ρ dans le modèle autorégressif spatial.

De la même manière que pour le modèle à variable endogène décalée, cette matrice est pleine et le schéma d'interaction y est global de sorte que toutes les localisations sont reliées les unes aux autres. Cependant, ce dernier est dû uniquement ici à un *effet de diffusion spatial* puisque la forme réduite se récrit :

$$y = X\beta + \left(I + \lambda W + \lambda^2 W^2 + ...\right) \varepsilon \qquad (26)$$

Un choc exogène ε dans une unité spatiale se répercute donc sur les valeurs prises par y dans toutes les localisations, cependant avec une intensité moindre à mesure que l'on s'éloigne de ladite unité.

2.3.3 Spécifications alternatives

Par analogie à la différenciation du premier ordre utilisée en Séries Temporelles, la variable dépendante peut être *spatialement filtrée* dans les deux modèles ci-dessus. Cette opération revient alors à isoler l'autocorrélation spatiale et conduit à des estimateurs des MCO centrés et efficients. Dans le cas du modèle autorégressif spatial, la transposition de ρWy dans le membre de gauche donne l'écriture suivante :

$$\left(I - \rho W\right)y = X\beta + \varepsilon \qquad (27)$$

Dans cette expression, $\left(I - \rho W\right)y$ représente la variable spatiale filtrée, tandis que le membre de droite se présente de la même manière que dans le modèle de régression linéaire standard. Par contre, dans le modèle à autocorrélation spatiale des erreurs, le filtre spatial s'applique aussi bien à la variable dépendante qu'aux variables explicatives contenues dans la matrice X . En effet, en prémultipliant les deux membres de la forme réduite du modèle (24) par $\left(I - \lambda W\right)$, on obtient l'équation ci-dessous :

$$\left(I - \lambda W\right)y = \left(I - \lambda W\right)X\beta + u \qquad (28)$$

On en déduit que le modèle à autocorrélation spatiale des erreurs est équivalent à un modèle de régression standard dont les variables dépendante et explicatives sont spatialement filtrées.

Cette dernière expression du modèle à autocorrélation spatiale des erreurs peut par ailleurs être récrite sous la forme d'une spécification « Durbin spatial » (Anselin, 1988). En effet, en développant l'équation ci-dessus et en y transposant le terme autorégressif spatial dans le membre de droite, on obtient[13] :

$$y = \lambda Wy + X\beta - \lambda WX\beta + u \qquad (29)$$

[13] Le modèle « Durbin spatial » est ainsi appelé par analogie au modèle de Durbin utilisé en Séries Temporelles. On parle également de « modèle du facteur commun ».

Ou encore[14] :

$$y = \lambda Wy + X\beta - WX\gamma + u \tag{30}$$

Le modèle « Durbin spatial » représente une forme réduite du modèle à autocorrélation spatiale des erreurs, et l'équation précédente montre qu'il s'agit d'une extension du modèle à variable endogène décalée, obtenue en introduisant dans ce dernier des variables exogènes décalées (WX).

De manière générale, le modèle à autocorrélation spatiale des erreurs est mis en œuvre lorsqu'il s'agit de rendre compte de la diffusion des chocs ou perturbations à travers un espace donné. Il permet également de prendre en charge les problèmes liés à l'omission de variables déterminantes dans l'évolution du phénomène étudié, en particulier lorsque ces dernières sont spatialement corrélées.[15]

Les spécifications à variable endogène décalée et à autocorrélation spatiale des erreurs peuvent également être combinées en un *modèle spatial général* dont l'expression formelle est la suivante :

$$\begin{cases} y = \rho W_1 y + X\beta + \varepsilon \\ \varepsilon = \lambda W_2 \varepsilon + u \end{cases} \tag{31}$$

Tout comme dans les cas précédents, ce modèle peut être récrit sous une forme réduite et l'on montre que celle-ci se présente comme suit :

$$y = (I - \rho W_1)^{-1} X\beta + (I - \rho W_1)^{-1}(I - \lambda W_2)^{-1} u \tag{32}$$

et la matrice de variances-covariances correspondante s'ensuit :

$$V(y) = \sigma^2 \left[\left(I - \rho W_1' \right)^{-1} \left(I - \lambda W_2' \right) (I - \lambda W_2)(I - \rho W_1) \right] \tag{33}$$

Ce modèle peut aussi se présenter sous une forme étendue de la spécification « Durbin spatial ». En effet, on montre que :

$$y = \rho W_1 y + \lambda W_2 y - \rho\lambda W_2 W_1 y + X\beta - \lambda W_2 X\beta + u \tag{34}$$

Pour $W_1 = W_2 = W$, cette expression devient[16] :

---

[14] L'équivalence de ces deux écritures n'est pas immédiate et le test d'un certain nombre de contraintes non linéaires est nécessaire. Ces dernières se résument à travers l'égalité suivante : $\lambda\beta = -\gamma$,et le test correspondant est désigné sous le nom de « test du facteur commun ».

[15] Cette spécification a notamment été utilisée pour l'étude des modèles hédoniques de prix immobiliers (Pace et Gilley [1998], Dubin [1998]) ou de la convergence conditionnelle (Fingleton [1999]).

[16] Dans ce cas précis, les paramètres ρ et λ ne sont identifiés que lorsque la matrice des variables explicatives X contient au moins une variable explicative, en dehors de la constante. Par ailleurs, afin de garantir l'identification et l'unicité des estimations des paramètres spatiaux, des contraintes non linéaires doivent être imposées à ces deux spécifications.

$$y = (\rho + \lambda)Wy - \rho\lambda W^2 + X\beta - \lambda WX\beta + u \qquad (35)$$

Quoiqu' ayant été utilisé dans certains travaux empiriques tels ceux de Case [1991, 1992] portant sur l'analyse de la demande et sur la diffusion de l'innovation, le modèle spatial général est encore cependant peu usité par rapport aux spécifications qu'il généralise.[17]

Comme on vient de le voir, diverses spécifications peuvent être utilisées afin de modéliser l'autocorrélation spatiale présente dans les données. Des processus encore plus complexes tels le modèle *SARMA* (Spatial Auto Regressive Moving Average) de Huang [1984] ont également été proposés dans la littérature, mais le modèle à variable endogène décalée et le modèle à autocorrélation spatiale des erreurs présentent l'avantage de la simplicité du point de vue de l'inférence statistique.

3. La méthode du maximum de vraisemblance

Le recours au principe du maximum de vraisemblance comme méthode d'estimation des modèles spatiaux est dû aux travaux de Cliff et Ord (1973) et Ord (1975) qui l'ont appliqué au modèle autorégressif spatial (variable endogène décalée) et au modèle à erreurs spatialement autocorrélées. Lorsqu'elle est appliquée aux modèles spatiaux, la méthode du maximum de vraisemblance nécessite cependant que soient vérifiées certaines conditions particulières assurant la convergence, l'efficience et la normalité asymptotiques des estimateurs.[18]

Comme d'habitude, la méthode repose sur l'expression de la fonction de log-vraisemblance et sur l'hypothèse de normalité pour les résidus du modèle. Son application au modèle spatial général permet d'en donner une présentation globale dont on déduit aisément les résultats pour le modèle à variable endogène décalée et le modèle à autocorrélation spatiale des erreurs. En effet, en partant de ce modèle et en faisant l'hypothèse que le vecteur des termes d'erreurs u a pour distribution jointe la loi normale ($u \to N(0, \sigma^2 I)$) la fonction de vraisemblance s'écrit :

$$L(u) = (2\pi\sigma^2)^{-n/2} e^{\left(-\frac{1}{2\sigma^2}u'u\right)} \qquad (36)$$

[17] Anselin et Bera [1998] soutiennent que la mise en œuvre de ce type de processus résulte souvent d'une mauvaise spécification de la matrice de poids conduisant à trouver, par exemple, de l'autocorrélation spatiale résiduelle dans un modèle à variable endogène décalée.

[18] Ces conditions de régularité ont été définies par Heijmans et Magnus (1986a,1986b,1986c) et Magnus (1978) et tiennent essentiellement à l'existence de la fonction de log-vraisemblance, à la continuité et à la différentiabilité des éléments du score et de la matrice hessienne associés. Pour les modèles spatiaux les plus courants, ces conditions se ramènent à des restrictions sur les poids spatiaux et sur l'espace des paramètres associé aux coefficients spatiaux (*Cf*. Anselin [1988]) .

Cependant, le terme d'erreurs u n'étant pas observé, la fonction de vraisemblance doit être exprimée en fonction des observations de la variable dépendante (y). On utilise alors le jacobien de la transformation (J), soit :

$$J = \det\left(\frac{\partial u}{\partial y}\right) = |I - \rho W_1||I - \lambda W_2| \tag{37}$$

Où $u = (I - \lambda W_2)(I - \rho W_1 y - X\beta)$, et l'on peut alors écrire la log-vraisemblance en fonction de la variable endogène (y)[19] :

$$\ln L(y|\rho, \lambda, \beta) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) + \ln|I - \rho W_1| + \ln|I - \lambda W_2| - \frac{1}{2\sigma^2}u'u \tag{38}$$

L'existence de la fonction de log-vraisemblance nécessite par ailleurs que J soit strictement positif, or on montre que les déterminants $|I - \rho W_1|$ et $|I - \lambda W_2|$ sont strictement positifs lorsque l'on a respectivement :

$$\frac{1}{\omega_{1\ \min}} < \rho < \frac{1}{\omega_{1\ \max}} \quad \text{et} \quad \frac{1}{\omega_{2\ \min}} < \lambda < \frac{1}{\omega_{2\ \max}} \tag{39}$$

Les valeurs $\omega_{\min}$ et $\omega_{\max}$ représentent respectivement la valeur propre négative et la valeur propre positive les plus grandes en valeur absolue de la matrice de poids (W_1 pour ρ, W_2 pour λ)[20].

Les estimateurs du maximum de vraisemblance sont alors obtenus par la résolution du système : $S(\theta) = \frac{\partial \ln L}{\partial \theta} = 0$; où $S(\theta)$ désigne le vecteur score (dérivées partielles premières de la log-vraisemblance) et $\theta = [\beta, \rho, \lambda]$ le vecteur des paramètres. On montre que ces estimateurs sont asymptotiquement efficients (sous réserve que les conditions de régularité soient vérifiées), puisque la matrice de variances-covariances égale l'inverse de la matrice d'information de Fisher (Borne inférieure de Cramer-Rao). En effet, on vérifie :

$$V(\theta) = [I(\theta)]^{-1}; \text{ avec : } I(\theta) = -E\left[\frac{\partial^2 \ln L}{\partial\theta\partial\theta'}\right]$$

---

[19] La fonction de log-vraisemblance n'est alors définie que lorsque les matrices $(I - \rho W_1)$ et $(I - \lambda W_2)$ sont régulières. Par ailleurs, les matrices $(I - \rho W_1)$ et $(I - \lambda W_2)$ ne sont pas ici triangulaires (comme cela peut être le cas en Séries Temporelles), ce qui complique considérablement les calculs liés à l'évaluation de la fonction de log-vraisemblance. Ord (1975) a cependant montré que les jacobiens spatiaux pouvaient être écrits en fonction des valeurs propres de la matrice de poids correspondante. On a ainsi : $|I - \rho W_1| = \prod_{i=1}^{n}(1 - \rho\omega_{1i})$ et $|I - \lambda W_2| = \prod_{i=1}^{n}(1 - \lambda\omega_{2i})$.

[20] Lorsqu'un processus moyenne mobile est retenu pour les erreurs ε, l'espace des paramètres associé au coefficient autorégressif spatial est donné par l'intervalle $\left]-1/w_{\max} ; -1/w_{\min}\right[$.

Des auteurs tels Anselin (1980,1988) ont par ailleurs montré que la méthode des Variables Instrumentales peut être utilisée afin de prendre en charge la question de la non convergence des estimateurs résultant de la corrélation entre un décalage spatial de la variable dépendante Wy et un terme d'erreur ε. Le recours à la méthode des moments généralisés a également été suggéré en vue de l'estimation de modèles présentant une autocorrélation spatiale des erreurs dans la mesure où elle permet d'y obtenir des estimations convergentes pour le coefficient autorégressif spatial.

4. Tests de Spécification

Dans la partie précédente, nous avons montré que l'autocorrélation spatiale présente dans les résidus du modèle de régression linéaire standard peut être modélisée de diverses manières. Cependant, le choix entre les différentes spécifications présentées nécessite la mise en œuvre d'une série de tests tels que l'hypothèse alternative offre une spécification spatiale explicite (ce qui n'est pas le cas du test de Moran). Ces procédures de test peuvent être fondées sur les principes du ratio de vraisemblance, de Wald ou du multiplicateur de Lagrange, ce dernier présentant cependant l'intérêt de ne nécessiter que l'estimation du modèle sous hypothèse nulle, lequel se ramène le plus souvent au modèle de régression linéaire classique, estimé par les MCO, facilitant ainsi considérablement l'inférence statistique.

4.1 Test d'omission d'une autocorrélation spatiale des erreurs

Ce test repose sur l'hypothèse de l'omission d'un processus autorégressif spatial pour le terme d'erreur ε, soit $\varepsilon = \lambda W\varepsilon + u$, dans le modèle de régression standard et porte sur la nullité du coefficient $(H_0 : \lambda = 0)$. La statistique du test est alors la suivante (Cf. Burridge [1980]) :

$$LM_\lambda = \frac{\left[\hat{\varepsilon}' W \hat{\varepsilon} / \hat{\sigma}^2\right]^2}{T} \qquad (40)$$

T désigne la trace de la matrice $(W'W + W^2)$, tandis que $\hat{\varepsilon}$ et $\hat{\sigma}^2$ correspondent aux estimations de ε et σ^2 dans le modèle contraint. Sous H_0, on a : $LM_\lambda \to \chi^2(1)$.[21]

4.2 Test d'omission d'une variable endogène décalée

La statistique associée à ce test a été proposée par Anselin [1988] et se présente de la manière suivante, sous l'hypothèse nulle $\rho = 0$:

[21] On montre que la statistique de ce test est identique à celle correspondant à un processus moyenne mobile pour le terme d'erreur ε.

$$LM_{\rho} = \frac{\left[\hat{\varepsilon}'Wy \big/ \hat{\sigma}^2\right]^2}{\hat{T}_1} \tag{41}$$

Dans cette expression, on a : $\hat{T}_1 = \left[\left(WX\hat{\beta}\right)'\left(I - X(X'X)^{-1}X'\right)\left(WX\hat{\beta}\right) + T\hat{\sigma}^2\right] \Big/ \hat{\sigma}^2$. Sous H_0, LM_{ρ} est également distribuée suivant une loi du $\chi^2(1)$.

Anselin et Bera [1998] montrent cependant qu'en cas de présence locale de ρ (*resp.* de λ) lorsqu'on procède au test LM_{λ} (*resp.* LM_{ρ}), les statistiques de test correspondantes ne sont plus distribuées suivant un $\chi^2(1)$[22]. Deux approches sont alors possibles : la première revient à tester l'hypothèse jointe $H_0 : \lambda = \rho = 0$ dans le modèle spatial général, tandis que la seconde est conditionnelle dans la mesure où elle consiste à tester l'omission d'une autocorrélation spatiale des erreurs dans un modèle à variable endogène décalée, et vice-versa.

4.3 Test joint d'une variable endogène décalée et d'une autocorrélation spatiale des erreurs

Ce test repose sur le modèle spatial général avec pour hypothèse nulle $H_0 : \lambda = \rho = 0$. Lorsque H_0 est acceptée, on retrouve le modèle de régression linéaire standard tandis que dans le cas contraire, on ne dispose pas d'indication quant à la nature du processus qui sous-tend la dépendance spatiale. La statistique du test est la suivante :

$$LM_{\lambda\rho} = \hat{E}^{-1}\left[\left(\hat{d}_{\lambda}\right)^2 \frac{\hat{D}}{\hat{\sigma}^2} + \left(\hat{d}_{\rho}\right)^2 T_{22} - 2\hat{d}_{\lambda}\hat{d}_{\rho}T_{12}\right] \tag{42}$$

Où $E = \left(D \big/ \sigma^2\right)T_{22} - \left(T_{12}\right)^2$, avec : $T_{ij} = tr\left[W_iW_j + W_i'W_j\right]$; $D = \left(W_1X\beta\right)' M\left(W_1X\beta\right) + T_{11}\sigma^2$ et $T = \left[\left(W' + W\right)W\right]$. d_{ρ} et d_{λ} représentent les scores, respectivement par rapport ρ et λ. Sous H_0, cette statistique est distribuée suivant une loi du χ^2 à deux degrés de liberté $\left(LM_{\rho\lambda} \to \chi^2(2)\right)$.

4.4 Tests conditionnels

Cette approche consiste donc à tester l'une des spécifications de base (variable endogène décalée ou autocorrélation spatiale des erreurs) en supposant la présence de l'autre. Dans un

[22] Par exemple, dans le cas du test LM_{λ}, l'hypothèse nulle risque d'être rejetée même lorsque $\lambda = 0$.

premier temps, on peut donc tester l'omission d'une autocorrélation spatiale des erreurs dans un modèle à variable endogène décalée, estimé par le maximum de vraisemblance. Le test repose ainsi sur les résidus de ce dernier et la statistique s'écrit :

$$LM_{\lambda|\rho} = \frac{\tilde{d}_{\rho}^{\,2}}{T_{22} - (T_{21A})^2 \tilde{V}(\tilde{\rho})} \qquad (43)$$

avec : $T_{21A} = tr\left[W_2 W_1 A^{-1} + W_2' W_1 A^{-1}\right]$, et $A = I - \tilde{\rho} W_1$. Sous l'hypothèse H_0, $LM_{\lambda|\rho} \rightarrow \chi^2(1)$ centré. L'omission d'un décalage spatial de la variable endogène dans un modèle à autocorrélation spatiale des erreurs peut également être testée à l'aide de la statistique suivante :

$$LM_{\rho|\lambda} = \frac{\left[\tilde{\varepsilon}' B' B W_1 y\right]^2}{H_{\rho} - H_{\theta\rho} \tilde{V}(\tilde{\theta}) H'_{\theta\rho}} \qquad (44)$$

Avec : $\hat{\varepsilon}$, le vecteur des résidus de l'estimation du modèle à erreurs autorégressives spatiales par le maximum de vraisemblance, $\theta = \left(\beta', \lambda, \sigma^2\right)$ et $B = I - \tilde{\lambda} W_2$. Par ailleurs, on a :

$$H_{\rho} = tr\left(W_1^{\,2}\right) + tr\left(BW_1B^{-1}\right)'\left(BW_1B^{-1}\right) + \frac{\left(BW_1X\tilde{\beta}\right)'\left(BW_1X\tilde{\beta}\right)}{\tilde{\sigma}^2} \qquad (45)$$

Et :

$$\begin{bmatrix} \dfrac{(BX)' BW_1X\tilde{\beta}}{\tilde{\sigma}^2} \\ tr\left(W_2B^{-1}\right)BW_1B^{-1} + trW_2W_1B^{-1} \\ 0 \end{bmatrix} \qquad (46)$$

Sous $H_0 : \rho = 0$, la statistique $LM_{\rho|\lambda} \rightarrow \chi^2(1)$.

4.5 Tests robustes

Sur la base du cadre défini par Bera et Yoon (1993), Anselin et *al.* (1996) ont développé des tests robustes à une mauvaise spécification locale, tel que cela peut être le cas pour les tests LM_{λ} et LM_{ρ}. Il s'agit en fait de versions ajustées de ces derniers, permettant d'obtenir comme distribution asymptotique un $\chi^2(1)$ centré sous hypothèse nulle, cela même en la présence respective de ρ ou λ. La version ajustée du test LM_{λ} sous $H_0 : \lambda = 0$ s'écrit alors de la manière suivante :

$$LM_{\lambda}^{*} = \frac{\left[\hat{d}_{\lambda} - T_{12}\hat{\sigma}^2\hat{D}^{-1}\hat{d}_{\rho}\right]}{T_{22} - T_{12}^{\,2}\hat{\sigma}^2\hat{D}} \quad \text{ou pour } W_1 = W_2 = W, \quad LM_{\lambda}^{*} = \frac{\left[\hat{d}_{\lambda} - T\hat{\sigma}\hat{D}^{-1}\hat{d}_{\rho}\right]}{T\left(1 - T\hat{\sigma}^2\hat{D}\right)} \qquad (47)$$

De même, sous $H_0 : \rho = 0$, le test robuste $LM_\rho{}^*$ se présente comme suit :

$$LM_\rho{}^* = \frac{\left[\hat{d}_\rho - T_{12}T_{22}{}^{-1}d_\lambda\right]}{\hat{\sigma}^{-2}\hat{D} - T_{12}{}^2 T_{22}{}^{-1}} \text{ ; soit pour } W_1 = W_2 = W \text{ , } LM_\rho{}^* = \frac{\left[\hat{d}_\rho - \hat{d}_\lambda\right]}{\hat{\sigma}^{-2}\hat{D} - T} \qquad (48)$$

Anselin et Rey (1991) proposent une combinaison des tests présentés ci-dessus afin de choisir le modèle spatial le mieux à même de représenter les données lorsque le test de Moran conclut à la présence d'autocorrélation spatiale. Cette procédure repose à la base sur les niveaux de significativité des tests LM_λ et LM_ρ, mais Anselin et Florax (1995) lui ajoutent les tests robustes $LM^*{}_\lambda$ et $LM^*{}_\rho$ qui permettent de préciser les résultats.

Les outils que nous venons de présenter complètent le cadre d'analyse devant nous guider dans la définition du processus spatial régissant les relations entre les pays de notre échantillon. L'économétrie spatiale nous permet donc de tenir compte de la localisation relative de ces pays et d'étudier l'impact que celle-ci peut avoir sur leurs performances économiques. Cependant, le test d'autocorrélation spatiale de Moran devrait plus largement être mené à chaque fois que l'on est en présence de données en coupe transversale. Certes, la disponibilité de données sur la localisation des unités spatiales a longtemps pu faire défaut, mais le récent développement des systèmes d'information géographique (SIG) a permis de surmonter cet obstacle.

5. Effets de l'autocorrélation spatiale dans le modèle MRW

5.1 Echantillon de données

Les données que nous utilisons dans cet article sont tirées de la version 6.0 de la série des « Penn World Table » initiée par Summers et Heston en 1988[23]. Notre étude repose sur les mêmes variables que celles utilisées par Mankiw et *al.* (1992) mais observées sur la période 1960-1995 ; les pays de l'échantillon sont ainsi choisis sur la base de la disponibilité des données pour les différentes variables étudiées et pour la période considérée. Nous suivons par ailleurs le choix effectué par Mankiw et *al.* d'exclure les pays producteurs de pétrole, ce qui nous conduit en fin de compte à un échantillon composé de 89 pays[24].

Nous utilisons ainsi les variables $\ln gdp95$ et $grate6095$ comme variables endogènes dans l'« équation de revenu » et dans l'« équation de convergence » du modèle MRW. En effet, ces deux variables sont respectivement représentatives du niveau de revenu réel par tête de 1995

[23] *Cf.* Summers, R., Heston, A. (1988), "A new set of international comparisons of real product and prices: *estimates for 130 countries", Review of Income and Wealth,* 34, 1-26.

[24] Notre échantillon se distingue de celui de Florax et Nijkamp (2003) qui reprennent l'échantillon de Mankiw et *al.* (1992) portant sur 98 pays, sur la période 1960-1985.

et du taux de croissance annuel moyen du revenu par tête sur la période 1960-1995, dans chacun des 89 pays de l'échantillon. Le niveau de revenu par tête est lui-même obtenu par le rapport entre le PIB réel et le nombre de travailleurs de l'année correspondante dans chacun des pays.

Par ailleurs, nous approchons le taux d'accumulation du capital physique $(\ln iony)$ par la part moyenne de l'investissement (privé et public) dans le PIB réel sur la période. La variable $(\ln school)$ désigne la part moyenne de la population en âge de travailler et inscrite dans l'enseignement secondaire sur la période 1960-1995. La variable $(\ln pop)$ représente quant à elle le taux effectif de dépréciation du stock de capital physique et humain, avec des observations sur le taux de croissance démographique ou plus précisément sur le taux de croissance de la population en âge de travailler (population entre 15 et 64 ans), sachant que, de la même manière que Mankiw et *al.*, nous fixons la somme du taux de croissance de la technologie et du taux d'amortissement du capital à 5%.

Dans nos estimations, nous mettons en œuvre une seule et unique matrice W dont les poids sont obtenus en prenant l'inverse de la distance séparant chaque couple de pays comme mesure de l'intensité de la relation qui les unit. Ladite distance est elle-même calculée en localisant chaque pays à l'aide des coordonnées géographiques de sa capitale (latitude et longitude), cela en appliquant le critère de la distance sphérique.[25] La matrice de distance inverse est très courante en économétrie spatiale et un certain nombre d'études telles celles de Florax et Nijkamp (2003) ou Moreno et Trehan (1997), y ont recours afin de quantifier le degré d'interaction spatiale entre économies. Par ailleurs, nous standardisons la matrice de poids ainsi obtenue dans le but de faciliter l'interprétation de nos résultats.

5.2 Recherche des spécifications spatiales

Comme souligné plus haut, la première étape de notre analyse repose sur le test d'absence d'autocorrélation spatiale dans les résidus des spécifications de base du modèle MRW, estimées par la méthode des Moindres Carrés Ordinaires. Dès lors, les résultats

[25] Les longitudes et les latitudes des différentes capitales sont tirées du « Thesaurus of Geographic Names » disponibles sur le site Internet : www.getty.edu. La formule de la distance sphérique (distance du « grand cercle) calculée dans Spacestat© est donnée par :

$$d_{ij} = 3959 * arc \cos \left\{ \cos \left| Y_i - Y_j \right| * \sin X_i * \sin X_J + \cos X_i * \cos X_j \right\}.$$

Les variables X et Y correspondant respectivement aux latitudes (lat) et longitudes (lon) des capitales de pays transformées de la manière suivante : X = $(90 - lat + \pi)/180$ et $Y = (lon + \pi/180)$.

d'estimation et ceux du test de Moran, associés respectivement à l'« équation de convergence » et à l'« équation du revenu » du modèle, se présentent comme suit :

Tableau 1 : Estimation par les Moindres Carrés Ordinaires

Modèle de revenu

Variable Dépendante = lngdp95

R^2	0.8086
$\overline{R}^2$	0.8019
Ecart-type	0.2354
Nobs, Nvars	89,4

Variable	Coefficient	*t*-statistique	prob. crit.
constante	5.822	5.138	0.000
lniony	0.538	4.872	0.000
lnschool	0.647	7.205	0.000
lnpop	-2.347	-5.889	0.000

Modèle de convergence

Variable Dépendante = grate6095

R^2	0.5810
$\overline{R}^2$	0.5610
Ecart-type	0.1637
Nobs, Nvars	89,5

Variable	Coefficient	t-statistique	prob. crit.
constante	3.424	3.360	0.001
lngdp60	-0.438	-5.008	0.000
lniony	0.520	5.643	0.000
lnschool	0.351	3.982	0.000
lnpop	-1.111	-2.900	0.005

Tableau 2 : Test d'autocorrélation spatiale de Moran

	Modèle de convergence	**Modèle de revenu**
I de Moran	0.04738847	0.03652254
I standardisé	2.92649295	2.36077552
Probabilité Critique (*p*)	0.00551039	0.02458622
Moyenne	-0.01867508	-0.01749225
Ecart-type	0.02257431	0.02288011

Les résultats du test de Moran sur chacune des spécifications du modèle permettent de conclure au rejet de l'hypothèse nulle d'absence d'autocorrélation spatiale. En effet, la probabilité marginale du test (*p*) est inférieure à 5% aussi bien pour le modèle de convergence que pour le modèle de revenu.

Autrement dit, les équations fournies par le modèle de Mankiw et *al.* (1992) sont mal spécifiées (pour notre échantillon) car l'hypothèse d'indépendance des observations qui sous-tend la méthode des moindres carrés ordinaires est rejetée et les estimations qui en résultent sont non convergentes et/ou inefficientes.[26]

5.3 Tests de spécification

Compte tenu de la présence d'autocorrélation spatiale dans les données, l'étape suivante consiste à s'interroger sur la forme fonctionnelle des processus spatiaux générant ces dernières, et dans cette perspective, on a recours à la procédure suggérée par Anselin et Rey (1991) et reprise par Florax, Folmer et Rey (2003).

Dans un premier temps, on procède donc à deux tests d'hypothèse simple (fondés sur les résidus des MCO) permettant de discriminer entre une spécification à variable endogène décalée et une spécification à autocorrélation spatiale des erreurs pour la prise en compte explicite de la dépendance spatiale dans les équations de convergence et du revenu. On a alors :

Tableau 3 : Tests simples d'omission d'une autocorrélation spatiale des erreurs ou d'une variable endogène décalée dans les modèles de convergence et de revenu.

Test simple d'omission d'une variable endogène décalée $\left(LM_{\rho}\right)$

	Modèle de convergence	**Modèle de revenu**
$\tilde{LM}_{\rho}$	5.50203643	6.73366539
Probabilité Critique (*p*)	0.01899434	0.00946100
Khi-deux (1) à 1%	6.63500000	6.63500000

Test simple d'autocorrélation spatiale des résidus $\left(LM_{\lambda}\right)$

	Modèle de convergence	**Modèle de revenu**
$\tilde{LM}_{\lambda}$	2.61073135	1.55073935
Probabilité Critique (*p*)	0.10614280	0.21302633
Khi-deux (1) à 1%	6.63500000	6.63500000

[26] Il faut rappeler que le test de Moran ne donne pas d'indication quant à la forme explicite de l'autocorrélation spatiale.

Les résultats du tableau 3 montrent donc a priori que la spécification autorégressive spatiale (à variable endogène décalée) correspond mieux au processus générateur des données, et cela aussi bien pour le modèle de convergence que pour le modèle de revenu. En effet, l'hypothèse d'omission d'une variable endogène décalée n'est pas rejeté pour chacun de ces deux modèles (la probabilité marginale (*p*) est inférieure à 5% dans les deux cas), tandis que l'hypothèse d'une autocorrélation spatiale des erreurs est rejetée aussi bien pour le modèle de convergence que pour le modèle de revenu.

Cependant, comme on a pu le souligner précédemment, ces tests simples ne prennent pas en compte l'éventuelle présence locale d'une variable endogène décalée lorsque l'on teste l'omission d'une autocorrélation spatiale des erreurs, et vice versa. Il en résulte donc que ces tests peuvent souffrir de biais, et on doit alors mettre en œuvre les tests robustes qui permettent de les corriger. Les résultats de ces tests sont donnés dans le tableau suivant :

Tableau 4 : Tests robustes d'omission d'une autocorrélation spatiale des erreurs ou d'une variable endogène décalée dans les modèles de convergence et de revenu.

Test robuste d'omission d'une variable endogène décalée $\left(LM_{\rho}^{*}\right)$

	Modèle de convergence	Modèle de revenu
$L\widetilde{M}_{\rho}^{*}$	2.94649641	5.19171438
Probabilité Critique (*p*)	0.08606395	0.02269482
Khi-deux (1) à 1%	6.63500000	6.63500000

Test robuste d'autocorrélation spatiale des résidus $\left(LM_{\lambda}^{*}\right)$

	Modèle de convergence	Modèle de revenu
$L\widetilde{M}_{\lambda}^{*}$	0.05519133	0.00878833
Probabilité Critique (*p*)	0.81426434	0.92531085
Khi-deux (1) à 1%	6.63500000	6.63500000

Les résultats du tableau 4 montrent que l'hypothèse d'omission d'une autocorrélation spatiale des erreurs est très fortement rejetée (la probabilité marginale (*p*) est supérieure à 80% aussi bien pour l'équation de convergence que pour l'équation du revenu.

Par contre, les tests robustes suggèrent qu'une variable endogène décalée parmi les régresseurs peut permettre de capter l'autocorrélation spatiale décelée à la suite des estimations du tableau 1, ce qui corrobore les résultats du tableau 3 ; l'hypothèse nulle correspondant à l'omission d'une variable endogène décalée peut en effet être acceptée au seuil de 10% dans les modèles de convergence et du revenu.

5.4 Estimation des modèles spatiaux

La spécification SAR (Spatial Auto-Regressive) peut donc être retenue comme le processus générateur des données. On peut alors d'ores et déjà estimer, par la méthode du maximum de vraisemblance, les modèles de convergence et du revenu « spatiaux » résultant de l'introduction d'un décalage spatial, respectivement du taux de croissance annuel moyen sur la période 1960-1995 (grate6095) et du logarithme du niveau de revenu par tête de long terme (lngdp95). On obtient :

Tableau 5 : Estimation de spécifications autorégressives spatiales pour les modèles de convergence et de revenu.

Modèle de revenu

Modèle autorégressif spatial
Variable Dépendante = lngdp95

R^2	=	0.8198
$\overline{R}^2$	=	0.8134
Variance	=	0.2072
Nobs, Nvars	=	89,4
Log-vraisemblance	=	-25.811205

Variable	Coefficient	*t*-stat. asymptot.	prob. crit.
const	2.748930	1.846952	0.064754
lniony	0.505757	4.844385	0.000001
lnschool	0.536354	5.868178	0.000000
lnpop	-1.808048	-4.223398	0.000024
ρ	0.459989	2.814203	0.004890

Modèle de convergence

Modèle autorégressif spatial
Variable Dépendante = grate6095

R^2	=	0.5970
$\overline{R}^2$	=	0.5779
Variance	=	0.1443
Nobs, Nvars	=	89,5
Log-vraisemblance	=	-9.9378068

Variable	Coefficient	*t*-stat. asymptot.	prob. crit.
constante	3.477497	3.627769	0.000286
lngdp60	-0.427146	-5.209702	0.000000
lniony	0.502146	5.781010	0.000000
lnschool	0.309672	3.692904	0.000222
lnpop	-0.891775	-2.404042	0.016215
ρ	0.545988	2.674385	0.007487

L'observation de ces derniers résultats montre que les variables explicatives associées aux spécifications de base (modèles non spatiaux) demeurent toutes significatives après l'introduction d'une variable endogène décalée comme régresseur supplémentaire. Et il en est de même pour le paramètre spatial ρ qui est fortement significatif dans les deux spécifications spatiales correspondant aux équations de convergence et de revenu.

La recherche de spécification que nous venons de mener suggère donc que la spécification autorégressive spatiale, obtenue par l'introduction d'une variable endogène décalée dans les équations standard du modèle MRW, fournit la meilleure représentation des données pour l'échantillon considéré, et cela est confirmé par le test d'absence d'autocorrélation spatiale résiduelle effectué sur les spécifications SAR du tableau 5 :

Tableau 6 : Test d'absence d'autocorrélation spatiale dans les résidus des spécifications *SAR*.

Test d'autocorrélation spatiale des résidus du modèle SAR $\left(LM_{\lambda|\rho}\right)$

	Modèle de convergence	**Modèle de revenu**
$LM_{\lambda\|\rho}$	3.23582865	1.46114281
Probabilité Critique (*p*)	0.07204385	0.22674837
Khi-deux (1) à 1%	6.63500000	6.63500000

Une fois retenu le processus spatial générateur des données, l'étape suivante consiste à en étudier la pertinence, c'est-à-dire évaluer l'impact de la variable endogène décalée, introduite dans les équations de convergence et de revenu, sur les résultats économétriques et les conclusions théoriques de Mankiw et *al.* (1992).

5.5 Résultats et Discussion

Modèle de Revenu

Les résultats de l'estimation d'une spécification SAR pour l'équation du revenu montrent que le paramètre autorégressif spatial (ρ) et très significatif, notamment avec une probabilité critique de *p* de 0.005, ce qui confirme l'omission de l'autocorrélation spatiale des erreurs dans la spécification standard (modèle non spatial) et donc la non convergence et la non efficience des estimations obtenues dans le tableau 1. L'estimation pour le paramètre spatial montre par ailleurs que les effets spatiaux omis dans le modèle non standard sont relativement importants ($\tilde{\rho} = 0.46$). Ainsi, pour un pays donné, une hausse de 1% de la moyenne spatialement pondérée des revenus de ses voisins (ici, tous les autres pays de l'échantillon) entraîne une augmentation d'environ 0.45% de son propre niveau de revenu par tête de long terme.

Cependant, l'introduction d'un décalage spatial de la variable endogène ne modifie pas sensiblement l'effet du taux d'accumulation du capital physique sur le niveau de revenu de long terme. Ce dernier évolue néanmoins légèrement à la baisse et ceci est également le cas pour les variables représentatives du capital humain et du taux de croissance démographique. Ces résultats sont donc conformes à ce l'on trouve dans la littérature et cela traduit probablement le fait que la partie la plus importante de l'autocorrélation spatiale détectée est due à des variables non prises en compte par le modèle théorique, elles-mêmes spatialement corrélées, et qui jouent un rôle important dans la détermination du niveau de revenu par tête de long terme.

Modèle de Convergence

L'estimation de la spécification SAR pour le modèle de convergence montre que les coefficients estimés pour les variables explicatives du modèle non spatial sont tous significatifs, et le sens de leur impact sur le taux de croissance est conforme à ce que prédit la théorie économique. En particulier, on obtient un coefficient négatif et très significatif pour le logarithme du niveau de revenu initial, ce qui corrobore l'hypothèse de convergence conditionnelle. De la même manière que pour l'équation du revenu, on peut remarquer que les coefficients résultant de l'estimation SAR sont assez proches de ceux de l'estimation par les MCO, même si l'impact des variables correspondantes est globalement (en valeur absolue) plus élevé dans cette dernière. Par ailleurs, le coefficient correspondant à la variable *lngdp60* dans la spécification SAR suggère une estimation de la vitesse de convergence égale à 1.5%, contre 1.6% dans la spécification non spatiale, pour des demi-vies correspondant respectivement à 46 et 43 ans[27]. Même si ces valeurs estimées restent proches, rappelons toutefois qu'en cas d'omission de la variable endogène décalée, les estimateurs des MCO sont biaisés et non convergents.

Les résultats du tableau 5 donnent une estimation positive et significative du coefficient autorégressif spatial d'environ 0.55, ce qui est encore une fois assez important dans la mesure où cela traduit le fait que le taux de croissance d'un pays réagira par une hausse de 0.55% à une augmentation de 1% de la moyenne pondérée (par les poids spatiaux) des taux de croissance des autres pays, toutes autres choses étant égales. Il existe donc d'importants effets de débordement géographique liés au taux de croissance entre les pays de l'échantillon.

Conclusion

Les résultats ci-dessus nous ont donc permis de montrer que l'« équation de convergence » et l'« équation de revenu » sont mal spécifiées et que ni le niveau de revenu de long terme, ni le taux de croissance n'échappent aux effets de la localisation et de l'espace. En effet, dans les deux cas, le test de Moran nous conduit à fortement rejeter l'hypothèse nulle suggérant ainsi l'omission d'une autocorrélation spatiale significative dans les spécifications du modèle de base. Le cas particulier du modèle de convergence montre que le rajout des variables explicatives standard du modèle MRW à l'équation de convergence absolue ne permet pas d'absorber la très forte autocorrélation spatiale que suggère le test de Moran effectué sur cette dernière.

Le recours aux outils de l'économétrie spatiale nous permet donc d'obtenir des estimateurs convergents et une inférence statistique fiable. En effet, la mise en œuvre de la procédure de recherche de spécifications spatiales proposée par Anselin et Rey (1991) et reprise par Florax, Folmer et Rey (2003) nous conduit notamment à retenir la spécification SAR comme processus générateur des données, or l'omission d'une variable endogène décalée dans les

[27] La demi-vie définit le nombre d'années juste nécessaire pour que les économies couvrent la moitié de la distance les séparant de leur état stationnaire. Elle peut être approximée par la formule suivante : $\ln(0.5)/\lambda$.

spécifications de base du modèle MRW se traduit par l'absence de prise en compte d'une dépendance spatiale « substantielle ». Dès lors, cette omission implique la non convergence des estimateurs des moindres carrés ordinaires et donc, les résultats obtenus dans le cadre du modèle MRW en sont altérés. Cela est d'autant plus vrai que le paramètre autorégressif spatial estimé est relativement élevé et très significatif aussi bien dans le modèle de convergence que dans le modèle de revenu.

Toutes autres choses étant égales, le niveau de revenu de long terme et le taux de croissance d'une économie seront d'autant plus élevés que la moyenne pondérée (par l'inverse des distances entre pays) des niveaux de revenu et des taux de croissance des économies voisines sera élevée. Autrement dit, plus les voisins seront riches (resp. pauvres), plus l'économie considérée sera riche (resp. pauvre) ; plus ses voisines croîtront vite, plus la croissance de son revenu par tête sera également forte. Il existe donc d'importants effets de débordement géographique entre économies traduisant probablement l'omission de variables importantes, sous-jacentes au processus de croissance, et fortement influencées par la localisation des différentes économies.

Dès lors, les effets de débordement géographique observés dans les équations de convergence et de revenu du modèle MRW peuvent être d'origines diverses. De fait, la mobilité du capital, le commerce international et les transferts de technologie qu'il peut engendrer, sont autant de facteurs pouvant expliquer la corrélation spatiale observée sur les niveaux de revenu et les taux de croissance. En particulier, les effets de la diffusion de la technologie paraissent tout à fait en phase avec le modèle MRW dans la mesure où ce dernier, de par sa construction théorique, explique la croissance à long terme par le seul biais du taux de croissance technologique ; il en est de même pour le commerce international qui joue un rôle fondamental dans la vulgarisation des technologies et confère ainsi, en grande partie, son caractère public au progrès technique.

Par ailleurs, compte tenu de la spécification de notre matrice de poids, on note que l'interaction d'un pays avec ses voisins sera d'autant plus faible que la distance qui les sépare sera importante. Des spécifications alternatives des poids spatiaux, certes plus complexes, fondées sur des variables économiques explicites pourraient permettre de mieux capter l'intensité des relations économiques dans l'espace ; cependant, l'utilisation de ce type de matrices pose la question de l'exogénéité requise des poids spatiaux.

En fait, la prise en compte explicite des effets de l'espace dans le modèle MRW est à apprécier tout d'abord sous l'angle de la fiabilité des estimations. En effet, la méthode des moindres carrés ordinaires conduit à des résultats fortement biaisés si l'on en juge par l'importance des effets spatiaux, et l'impact réel des variables explicatives standard sur le revenu par tête et le taux de croissance de long terme est alors difficilement appréciable. Dans tous les cas, l'estimation convergente fournie par la spécification SAR suggère que les MCO surestiment cet impact et sous-estiment la vitesse de convergence des économies vers leur état stationnaire. En définitive, le test de l'hypothèse d'absence d'autocorrélation spatiale dans les régressions en coupe transversale traditionnellement mise en œuvre dans les études empiriques sur la croissance paraît d'une naturelle nécessité.

Bibliographie

Agénor, P.R. (2000), "*The Economics of Adjustment and Growth*", Academic Press.

Anselin, L., (1980), "*Estimation methods for spatial autoregressive structures*", Cornell University, Regional Science Dissertation and Monograph Series #8, Ithaca, NY.

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

Anselin, L., (2001a), "Rao's score test in spatial econometrics", *Journal of Statistical Planning and Inference*, 97, 113-139.

Anselin, L. (2001b), "*Spatial Econometrics*", in: Baltagi B. (ed.), Companion to Econometrics, Basil Blackwell, Oxford.

Anselin, L., Bera, A.K. (1998), "Spatial dependence in linear regression models with an introduction to spatial econometrics", in *Handbook of applied Economic Statistics* (A. Ullah, and D. Giles, Eds.). Marcel Dekker, New York.

Anselin, L., Bera, A.K., Florax, R.J.G.M., Yoon, M. (1996), "Simple diagnostic tests for spatial dependence", Regional Science and Urban Economics, 26.

Anselin, L., Florax, R.J.G.M. (1995), "Small sample properties of tests for spatial dependence in regression models, in: Anselin, L., Florax, R.J.G.M. (eds), *New Directions in Spatial Econometrics*, Springer, Berlin.

Anselin, L., Griffith, D.A. (1988), "Do spatial effects really matter in regression analysis?",
Papers of the Regional Science Association, 65, 11-34.

Anselin, L., Rey, S. (1991), "Properties of tests for spatial dependence in linear regression models, *Geographical Analysis*, 23, 112-131.

Aten, B. (1997), "Does space matter? International comparisons of the prices of tradables and
nontradables", *International Regional Science Review*, 20, 35-52.

Azariadis, C., Drazen, A. (1990), "Threshold externalities in economic development", *Quarterly Journal of Economics*, CV, 501-526.

Baumol, W.J. (1986), "Productivity growth, convergence and welfare: what the long run data show", *American Economic Review*, 78, 1155-1159.

Baumont, C., Ertur, C., Le Gallo, J. (2002), ''Estimation des effets de proximité dans le processus de convergence régionale : une approche par l'économétrie spatiale sur 92 régions européennes (1980-1995), *Revue d'Economie Régionale et Urbaine*, 2.

Barro, R.J. (1991), "Economic growth in a cross-section of countries", *Quarterly Journal of Economics*, 106:2, 407-443.

Barro, R.J., Sala-i-Martin, X. (1995), “*Economic Growth*”, New York, McGraw-Hill.
Benhabib, J., Spiegel, M.M. (1994), “The role of human capital in economic development: evidence from cross-country data”, *Journal of Monetary Economics*, 34, 143-173.
Bera, A.K., Yoon, M. (1993), “Specification testing with locally misspecified alternatives”, *Econometric theory*, 9, 649-658.
Bernard, A.B. and Durlauf, S.N. (1996), “Interpreting tests of the convergence hypothesis”, *Journal of Econometrics*, 71, 161-173.
Burridge, P. (1980), “On the Cliff-Ord test for spatial autocorrelation”, Journal of the Royal Statistical Society, B, 42, 107-108.
Can, A. (1990), “The measurement of neighborhood dynamics in urban house prices, *Economic Geography*, 66, 254-272.
Case, A.C. (1991), “Spatial patterns in household demand”, *Econometrica*, 59, 953-965.
Case, A.C. (1992), “Neighborhood influence and technological change”, *Regional Science and Urban Economics*, 22, 491-508.
Case, A.C., Rosen, H., S., Hines, J.R. (1993), “Budget spillovers and fiscal policy interdependence: Evidence from the states, *Journal of Public Economics*, 52, 285-307.
Caselli, F., Esquivel, G., Lefort, F. (1996), “Reopening the convergence debate: A new look at the cross-country growth empirics”, *Journal of Economic Growth*, 1, 363-389.
Cass, D. (1965), “Optimum growth in a aggregative model of capital accumulation”, *Review of Economic Studies*, 32, 223-240.
Chiang, A.C. (1984), “*Fundamental Methods of Mathematical Economics*”, 2 ed. New York: McGraw-Hill Book Company.
Cho, D. (1996), “An alternative interpretation of conditional convergence results”, *Journal of Money, Credit, and Banking*, 28, 669-681.
Cressie, N. (1993), “*Spatial Statistics for spatial data”*, John Wiley, New York.
Cliff, A.D., Ord, J.K. (1972), “Testing for spatial autocorrelation among regression residuals”, *Geographical Analysis*, 4, 267-284.
Cliff, A.D., Ord, J.K. (1973), “*Spatial Autocorrelation*”, Pion, Londres.
Cliff, A.D., Ord, J.K. (1981), “*Spatial Processes: models and applications*”, Pion, Londres.
Conley, T.G. (1996), “*Econometric modelling of cross sectional dependence*”, Manuscript, Northwestern University.
Conley, T.G., Ligon, E. (2002), “Economic distance and cross-country spillovers”, *Journal of Economic Growth*, Springer, vol. 7(2), pages 157-87.
De Long, J.B. (1988), “Productivity growth, convergence and welfare: Comment”, *American Economic Review*, 78, 1138-1154.
Dubin, R.A. (1988), “Estimation of regression coefficients in the presence of spatially autocorrelated error terms, *Review of Economics and Statistics*, 70, 466-474.
Dubin, R.A. (1998), “Predicting house prices using multiple listings data”, *Journal of Real Estate Finance and Economics*, 17, 35-59.
Durbin, J., Watson, G.S. (1950), “Testing for serial correlation in least squares regression-I”, *Biometrika*, 37, 409-428.
Durlauf, S.N., Johnson, P.A. (1995), “Multiple regimes and cross-country growth behaviour”, *Journal of Applied Econometrics*, 10:4, 365-384.
Durlauf, S.N., Quah, D.T. (1999), “*The New Empirics of Economic Growth*” in Taylor, J.B., and Woodford, M., eds., Handbook of Macroeconomics Vol. 1A (North-Holland).
Easterly, W., Levine, R. (1998), “Troubles with the neighbours: Africa’s Problem, Africa’s Opportunity”, *Journal of African Economics*, 7:1, 120-142.
Ertur, C., Le Gallo, J. Baumont C. (2005), « The European Regional Convergence Process, 1980-1995: Do Spatial Regimes and Spatial Dependence Matter? », forthcoming in *International Regional Science Review*.
Ertur, C., Le Gallo, J. (2002), ‘‘An exploratory spatial data analysis of European regional disparities, 1980-1995, in: Fingleton B. (ed.), *European Regional Growth*, Springer, Berlin.
Fingleton, B. (1999), “Estimates of time to economic convergence: an analysis of regions of the European Union”, *International Regional Science Review*, 22, 5-34.

Florax R., Folmer H., Rey S. (2003) Specification searches in spatial econometrics: The relevance of Hendry's methodology, forthcoming in *Regional Science and Urban Economics*, 33.

Florax, R.J.G.M. & Nijkamp, P. (2003), "*Misspecification in linear spatial regression models*", Tinbergen Institute Discussion Papers 03-081/3, Tinbergen Institute.

Galor, O. (1996), "Convergence? Inferences from theoretical models", *Economic Journal*, 106, 1056-1081.

Gandolfo, G. (1997), "*Economic Dynamics*", 3rd Ed., Springer-Verlag, Berlin.

Geary, R.C. (1954), "The contiguity ratio and statistical mapping", *The Incorporated Statistician*, 5:115-145.

Grossman, G. M., Helpman, E. (1994), "Endogenous Innovation in the Theory of Growth", *Journal of Economic Perspectives*, 8: 23 – 44.

Haining, R. (1978), "The moving average model for spatial interaction" Transactions and Papers of the Institute of British Geographers, 3, 202-225.

Haining, R. (1990), "Spatial data analysis in the social and environmental sciences", *Cambridge University Press*, Cambridge.

Heijmans, R.D.H., Magnus, J.R. (1986a), "On the first-order efficiency and asymptotic normality of maximum likelihood estimators obtained from dependent observations", *Statistica Neerlandica*, 40, 169-188.

Heijmans, R.D.H., Magnus, J.R. (1986b), "Asymptotic normality of maximum likelihood estimators obtained from normally distributed but dependent observations", *Econometric Theory*, 2,374-412.

Heijmans, R.D.H., Magnus, J.R. (1986c), "Consistent maximum likelihood estimation with dependent observations: the general (non normal) case and the normal case, *Journal of Econometrics*, 32, 253-285.

Huang, J.S. (1984), 'The autoregressive moving average model for spatial analysis", *Australian Journal of Statistics*, 26, 169-178.

Inada, K. (1964), "Some structural characteristics of turnpike theorems", *Review of Economic Studies*, 31 (January):43-58.

Magnus, J. (1978), "Maximum likelihood of the GLS model with unknown parameters in the disturbance covariance matrix, *Journal of Econometrics*, 7, 281-312 (Corrigenda, Journal of Econometrics, 10, 261).

Kelejian, H.H., Robinson, D.P. (1993), "A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model, *Papers in Regional Science*, 72, 297-312.

Kelejian, H.H., Robinson, D.P. (1995), "Spatial correlation: A suggested alternative to the autoregressive model", in Anselin, L., Florax, R.J.G.M. (eds), *New Directions in Spatial Econometrics*, Springer, Berlin.

Kendrick, J.W. (1976), "*The formation and stocks of total capital*", New York: Columbia University for NBER.

King, M.L. (1981), "A small sample property of the Cliff-Ord test for spatial correlation", *Journal of the Royal Statistical Society*, B, 43, 263-264.

Klenow, P.J., Rodriguez-Clare, A. (1997), "The neoclassical revival in growth economics: has it gone too far?", *NBER Macroeconomics Annual*, 1997, 73-102.

Koopmans, T. (1965), "*On the concept of optimal growth, The econometric approach to development planning*", North-Holland, Amsterdam.

Lee, K., Pesaran, M., Smith, R. (1998), "Growth empirics: A panel data approach - a comment", *Quarterly Journal of Economics*, vol. 113, Feb., 319-323.

Le Gallo, J. (2002), "*Disparités géographiques et convergence des régions européennes : une approche par l'économétrie spatiale*", Thèse de doctorat, LATEC, Université de Bourgogne.

Le Gallo, J., Ertur, C. (2002), ''Exploratory spatial analysis of the distribution of regional per capita GDP in Europe, 1980-1995, *Papers in Regional Science*.

Le Gallo, J., Ertur, C., Baumont, C. (2002), "A spatial econometric analysis of convergence,
across European regions, 1980-1995, in: Fingleton B. (ed), *European Regional Growth*, Springer, Berlin.

Lucas, R.E. (1988), "On the mechanics of economic development", *Journal of Monetary Economics*, 22, 3-42.

Mankiw, N.G., Romer, D., Weil, D. (1992), “A contribution to the empirics of economic growth“, *Quarterly Journal of Econometrics*, CVII, 407-437.
Mankiw, N.G. (1995), “The growth of nations”, *Brooking Papers on Economic Activity*, 1, 275-310.
Moran, P. (1948), “The interpretation of statistical maps”, *Journal of the Royal Statistical Society*, 10B, 243-251.
Moran, P. (1950), “A test for serial independence of residuals”, *Biometrika*, 37, 178-181.
Moreno, R., Trehan, B. (1997), “Location and the growth of nations”, *Journal of Economic Growth*, 2, 399-418.
Ord, J.K. (1975), “Estimation methods for models of spatial interaction”, *Journal of the American Statistical Association*, 70, 120-126.
Pace, R. K., Gilley, Gilley, O.W.(1998), “Optimally Combining OLS and the Grid estimator,” *Real Estate Economics*, Volume 26, Number 2, pp. 331-347.
Paelinck, J.H.P., Klaassen, L.H. (1979), “Spatial Econometrics”, Saxon House, Farnborough. Pesaran, M.H., Smith, R. (1994), ‘Estimating long-run relationships from dynamic heterogeneous panels”, Working Paper, University of Cambridge and forthcoming, Journal of Econometrics.
Romer, P.M. (1986), “Increasing returns and long-run growth”, *Journal of Political Economy*, 94, 1002-1037.
Rey, S.J., Montouri, B.D. (1999), “U.S. regional income convergence: a spatial econometric perspective”, *Regional Studies*, 33, 145-156.
Solow, R. (1956), “A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70, 65-94.
Shone, R., (1997), “*Economic Dynamics: Phase Diagrams and Their Economic Applications*”, Cambridge: Cambridge University Press.
Summers, R., Heston, A. (1988), “A new set of international comparisons of real product and prices: estimates for 130 countries”, *Review of Income and Wealth*, 34, 1-26.
Summers, R., Heston, A. (1991), “The Penn World Table, Mark5: An expanded set of international comparisons, 1950-1988”, *Quarterly Journal of Economics*, 106:2, 327-368.
Temple, J. (1999), “The new growth evidence”, *Journal of Economic Literature*, 37, 112-156.