

STATISTIQUE, ÉCONOMÉTRIE ET ESPACE : PROGRÈS ET DÉFIS EN SCIENCE RÉGIONALE

Julie Le Gallo, Isabelle Thomas

Armand Colin / Dunod | « *Revue d'Économie Régionale & Urbaine* »

2015/1 mai | pages 9 à 26

ISSN 0180-7307

ISBN 9782200929817

Article disponible en ligne à l'adresse :

<http://www.cairn.info/revue-d-economie-regionale-et-urbaine-2015-1-page-9.htm>

!Pour citer cet article :

Julie Le Gallo, Isabelle Thomas, « Statistique, économétrie et espace : progrès et défis en science régionale », *Revue d'Économie Régionale & Urbaine* 2015/1 (mai), p. 9-26.

DOI 10.3917/reru.151.0009

Distribution électronique Cairn.info pour Armand Colin / Dunod.

© Armand Colin / Dunod. Tous droits réservés pour tous pays.

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

Statistique, économétrie et espace : progrès et défis en science régionale

Statistics, econometrics and space: advances and challenges in regional science

Julie LE GALLO

CRESE, Université de Franche-Comté - Besançon (France)
jlegallo@univ-fcomte.fr

Isabelle THOMAS

CORE/UCL - Louvain-la-Neuve (Belgique)
isabelle.thomas@uclouvain.be

- 1 -

Introduction

« L'espace n'est pas économiquement neutre. Son introduction n'apporte pas des raffinements de détail : *elle change tout* » (PONSARD, 1988). De fait, si le rôle de l'espace et des territoires dans la structuration des échanges et des interactions économiques constitue l'objet d'étude des géographes et des auteurs de sciences régionales depuis toujours, ce rôle a été assez peu considéré en économie jusqu'à l'article séminal de KRUGMAN (1991). L'une des explications avancées est celle des contraintes de modélisation (THISSE, 1997) : les externalités, les rendements croissants et la concurrence stratégique sont en effet incontournables dès lors qu'on souhaite parler d'espace de façon pertinente.

Ce qui est vrai pour la théorie économique l'est tout autant pour l'analyse économétrique. Ainsi, l'autocorrélation spatiale, phénomène bien connu des géographes quantitativistes depuis la formalisation du concept par CLIFF et ORD (1969, 1973), a été longtemps ignorée en économétrie malgré les nombreux travaux et revues de littérature (voir par exemple ceux de FOTHERINGHAM *et al.*, 2000 ; FOTHERINGHAM, 1997, 1998, 1999). Les données géolocalisées, c'est-à-dire les données pour lesquelles chaque observation est associée à une localisation repérée par ses coordonnées (latitude et longitude par exemple), n'étaient pas considérées comme nécessitant un traitement particulier. L'impact de l'autocorrélation spatiale sur l'estimation et l'inférence statistique était ignoré et les ouvrages fondateurs du champ de l'économétrie spatiale (PAELINCK et KLAASSEN, 1979 ; UPTON et FINGLETON, 1985 ; ARBIA, 1988 ; ANSELIN, 1988) inconnus hors de la science régionale. Des contraintes de modélisation peuvent être avancées ici également : contrairement à l'autocorrélation temporelle, le traitement approprié de l'autocorrélation spatiale est délicat. Il n'existe pas d'ordre naturel dans l'espace, comme il en existe un dans le temps. L'autocorrélation spatiale est multidirectionnelle et non unidirectionnelle comme l'autocorrélation temporelle. Les modélisations et les résultats classiques en économétrie des séries temporelles ne peuvent donc pas être transposés directement sur données géo-localisées. Des ajustements non triviaux sont nécessaires.

Aujourd'hui, les travaux mobilisant les techniques de la statistique et de l'économétrie spatiales foisonnent. Si les revues de science régionale ou de géographie quantitative constituent toujours un réceptacle privilégié pour ces travaux, ils apparaissent également régulièrement dans les revues d'économie d'autres champs, dans des revues d'économie généralistes mais aussi dans les revues en écologie, sociologie, sciences politiques, *etc.* Cette explosion peut s'expliquer tout d'abord par la reconnaissance du rôle de l'espace en économétrie et le développement d'outils de modélisation appropriés. De ce point de vue, un effort important a été entrepris ces dernières années afin de donner des fondements statistiques rigoureux aux modèles d'économétrie spatiale (LEE, 2004). Elle reflète ensuite une adéquation des outils techniques, que ce soit la généralisation des systèmes d'informations géographiques ou des logiciels spécifiquement dédiés à l'analyse des données spatiales, comme GeoDa (ANSELIN *et al.*, 2006) ou des librairies pour le logiciel statistique R (BIVAND,

2004). La disponibilité de bases de données à des échelles spatiales de plus en plus fines a enfin incité les chercheurs à développer les méthodes d'analyse appropriées.

C'est dans ce contexte que se situe ce numéro spécial de la *Revue d'Économie Régionale et Urbaine*. Face à la difficulté d'une pluridisciplinarité réussie, mais pourtant indispensable dès lors que l'objet d'étude est l'espace, il vise à faire un point d'étape sur la statistique et l'économétrie spatiales en regroupant des contributions d'auteurs de différentes disciplines : économistes, géographes, *regional scientists*. Avant de détailler les contributions de ce numéro dans la section 3, nous présentons rapidement dans la section 2 quelques éléments qui nous semblent caractériser la statistique et l'économétrie spatiales aujourd'hui et les défis à prévoir.

- 2 -

Progrès et défis de la statistique et de l'économétrie spatiales

Ce qui marque tout d'abord, c'est l'explosion actuelle des travaux à la fois théoriques et appliqués de statistique et d'économétrie spatiales. Les requêtes « *spatial statistics* » et « *spatial econometrics* » renvoient ainsi respectivement à 2 540 000 et 69 000 résultats sur *Google Scholar*, ce qui doit être mis en perspective avec les 603 000 entrées pour le terme générique « *econometrics* », les 11 600 entrées pour « *microeconometrics* » et les 195 000 entrées pour « *panel econometrics* ». Les numéros spéciaux sont aujourd'hui nombreux et surtout leur fréquence s'accroît au cours du temps. Afin de l'illustrer, nous avons recherché dans *Google Scholar* et *Econlit* tous les numéros spéciaux de statistique et d'économétrie spatiales. Ils sont regroupés dans le Tableau 1.

Au-delà de ces chiffres, force est de constater que ces champs sont maintenant reconnus dans le *mainstream*. Une revue intégralement consacrée à la statistique spatiale a été créée en 2012 : *Spatial Statistics*, contenant notamment un point d'étape sur la discipline (GRIFFITH, 2012). De son côté, ANSELIN (2010) note que l'économétrie spatiale est passée au stade de la maturité, couronné par la création en 2007 de la *Spatial Econometrics Association* (ARBIA, 2011). Les articles se basant sur données spatiales et utilisant pleinement les informations apportées par la dimension géographique sont monnaie courante dans les revues d'économie généralistes ou les revues de champ tandis que les revues phares d'économétrie, *Econometric Theory* ou *Journal of Econometrics*, publient très régulièrement des articles d'économétrie spatiale théorique. L'économétrie spatiale fait également son apparition dans les manuels d'économétrie. Quelques textes mentionnent son existence (KENNEDY, 2003 ; CAMERON et TRIVEDI, 2005 ; WOOLDRIDGE, 2010) alors que d'autres y consacrent des chapitres entiers (BALTAGI, 2008 ; PIROTTE, 2010). Des chapitres spécifiques de revue de littérature existent également dans des *Handbook* généraux : *Handbook of Applied Economic Statistics* (ULLAH et GILES, 1998), *Companion to Theoretical Econometrics* (BALTAGI, 2001), *Palgrave Handbook of Econometrics* (MILLS et PATTERSON, 2006), *Econometrics of Panel Data* (MATYAS et SEVESTRE, 2008). Il est permis de penser que la

citation de WOOLDRIDGE (2010) sur l'autocorrélation spatiale (« *For better or worse, spatial correlation is often ignored in applied work because correcting the problem can be difficult* », 2010) n'est de fait plus d'actualité.

Sur le fond, la présence de chapitres consacrés au problème de l'autocorrélation spatiale dans deux ouvrages sur les données de panel montre à l'évidence que l'économétrie spatiale ne se concentre plus seulement sur le traitement de l'autocorrélation spatiale (estimation et inférence statistique) dans les modèles de coupe transversale. Après ce passage obligé, elle irrigue maintenant d'autres champs de l'économétrie : économétrie des données de panel, économétrie des variables qualitatives et microéconométrie en général, économétrie de l'évaluation, modèles simultanés, *etc.* Les champs de recherche actuels s'attachent dès lors à analyser comment le traitement de l'autocorrélation spatiale doit se combiner avec les problématiques économétriques traditionnelles, telles que le problème de l'endogénéité des variables explicatives, l'hétérogénéité inobservable, l'hétéroscédasticité, l'autocorrélation temporelle, les problèmes de spécifications, *etc.*

Cette observation nous amène à pointer un défi qui nous paraît important pour les années à venir : celui de la réflexion à mener sur la matrice de poids. Rappelons brièvement que cette matrice carrée, contenant un nombre de lignes et de colonnes égal au nombre de lieux, contient la structure de connectivité supposée entre les observations. Les matrices purement géographiques, comme les matrices de contiguïté ou les matrices basées sur la distance à vol d'oiseau, présentent l'avantage d'être exogènes au modèle tout en reflétant le rôle de l'espace comme support des interactions économiques. L'argument d'exogénéité devient moins évident dès lors qu'on se tourne vers des matrices basées sur des distances-temps ou des accessibilités, dans la mesure où l'accessibilité entre lieux peut être manipulée par l'homme. Il ne tient plus lorsqu'on spécifie les matrices de poids avec des mesures de proximité économique ou politique, des flux commerciaux ou encore des liens sur des réseaux, sachant que la présence ou l'absence de ces liens découlent de choix stratégiques des agents économiques. Certains articles très récents s'attachent à explorer les conséquences de l'endogénéité des matrices de poids (KELEJIAN et PIRAS, 2014 ; QU et LEE, 2015 ; HSIEH et LEE, 2015). D'autres auteurs approchent le problème *via* l'estimation directe de la matrice de poids en utilisant la structure de panel des données (BEENSTOCK et FELSENSTEIN, 2012) ou des contraintes structurelles (BHATTACHARJEE et JENSEN-BUTLER, 2013). Compte tenu, d'une part, de l'importance de la spécification de la matrice de poids dans les modèles d'économétrie spatiale et d'autre part, de l'incertitude sur la structure de connectivité des données, nul doute que ce champ de recherche constituera l'un des axes forts des recherches futures, étant donnée la disponibilité croissante de données relationnelles dans les villes dites intelligentes et hyperconnectées.

Tableau 1 – Liste des numéros spéciaux de revues en statistique et économétrie spatiales			
Date	Journal	Intitulé	Éditeurs
1992	<i>Regional Science and Urban Economics</i>	Space and applied econometrics	ANSELIN L.
1997	<i>International Regional Science Review</i>	Introduction to the special issue of spatial econometrics	ANSELIN L. et REY S.
1998	<i>Journal of Real Estate Finance and Economics</i>	Spatial statistics and real estate	PACE K.P., BARRY R. et SIRMANS C.F.
2003	<i>International Regional Science Review</i>	Spatial externalities	ANSELIN L.
2003	<i>International Regional Science Review</i>	Spatial econometric data analysis: moving beyond traditional models	FLORAX R.J.G.M. et VAN DER VLIST A.J.
2004	<i>Geographical Analysis</i>	Methodological developments in spatial econometrics and statistics	LESAGE J.P., PACE K.P. et TIEFELS-DORF M.
2004	<i>Journal of Real Estate Finance and Economics</i>	Spatial statistics and real estate	PACE K.P. et LESAGE J.P.
2005	<i>Région et Développement</i>	Croissance, convergence et interactions régionales : outils récents de l'analyse spatiale quantitative	LE GALLO J. et DALL'ERBA S.,
2006	<i>Geographical Analysis</i>	Recent advances in software for spatial analysis in the social sciences	REY S. et ANSELIN L.
2006	<i>Papers in Regional Science</i>	The spatial dimension of economic growth and convergence	BODE E. et REY S.
2007	<i>International Regional Science Review</i>	Recent developments of regional labor market analysis using spatial econometrics	MÖLLER J. et SOLTWEDEL R.

Tableau 1 – (suite)

2007	<i>Journal of Econometrics</i>	Analysis of spatially dependent data	BALTAGI B.
2008	<i>Empirical Economics*</i>	Introduction	BALTAGI B. et ARBIA G.
2008	<i>Papers in Regional Science</i>	New spatial econometric techniques and applications in regional science	ARBIA G. et FINGLETON B.
2008	<i>Revue d'Économie industrielle</i>	Introduction	LE GALLO J., MASSART N. et PLUNKET A.
2009	<i>Computational Statistics and Data Analysis</i>	Spatial statistics: methods, models and computation	LESAGE J., BANERJEE S., FISCHER M.M. et CONGDON P.
2009	<i>Geographical Analysis</i>	Celebrating 40 years of scientific impacts by Cliff and Ord	GRIFFITH D.A.
2009	<i>Journal of Geographical Systems</i>	Recent research in spatial real estate hedonic analysis	PAEZ A.
2009	<i>Papers in Regional Science</i>	Spatial analysis of economic systems and land use	PAEZ A.
2010	<i>Journal of Geographical Systems</i>	Modeling spatiotemporal data: an issue in honor of Dr. Jean Paelinck	MUR F. et LOPEZ A.
2010	<i>Regional Science and Urban Economics*</i>	Advances in spatial econometrics	ARBIA G. et KELEJIAN H.H.
2011	<i>Annals of Regional Science</i>	Advanced methods and applications in regional science	MATILLA-GARCIA M. et RUIZ MARIN M.
2011	<i>International Regional Science Review</i>	In honor of Manfred Fisher	REY S.J.
2012	<i>Economic Modeling*</i>	Frontiers in spatial econometrics modelling	ARBIA G., LOPEZ-BAZO E. et MOSCOONE F.
2012	<i>Journal of Geographical Systems</i>	In honor of James P. LESAGE	FISHER M.M.
2012	<i>Revista Estadística Espanola</i>	Numéro especial sobre Econometria Espacial	MONTERO J.M. et PAELINCK J.H.P.

Tableau 1 – (suite)			
2012	<i>Journal of Regional Science</i>	Whither spatial econometrics?	PARTRIDGE M.D., BOARNET M., BRAKMAN S. et OTTAVIANO G.
2012	<i>Statistical Methodology</i>	Editorial of the special issue on spatial statistics	BANERJEE S. et DEY D.K.
2012	<i>Spatial Economic Analysis</i>	Editorial	FINGLETON B. et ERTUR C.
2013	<i>Spatial Economic Analysis</i>	Editorial	ARBIA B. et PRUCHA I.R.
2013	<i>Econometric Reviews</i>	An overview of dependence in cross-section, time-series and panel data	BALTAGI B.H. et MAASOUMI E.
2013	<i>Journal of Geographical Systems</i>	Advances in the statistical modeling of spatial interaction data	PATUELLI R. et ARBIA G
2014	<i>Geographical Analysis*</i>	Advances in cross-sectional and panel data spatial econometric modeling	ARBIA G. et THOMAS-AGNAN C.
2014	<i>International Regional Science Review*</i>	Contributions to spatial econometrics	BAO Y., FLORAX R.J.G.M. et LE GALLO J
Note : * numéros spéciaux suivant la conférence annuelle mondiale de la Spatial Econometrics Association.			

- 3 -

Le numéro spécial

Dans la tradition des numéros spéciaux recensés dans le Tableau 1, ce numéro spécial de la *Revue d'Économie Régionale et Urbaine* a pour objectif d'illustrer par quelques exemples l'actualité et la diversité de la statistique et de l'économétrie spatiales en sciences régionales : diversité des thèmes abordés, diversité des enjeux et diversité des méthodes mobilisées. Dans ce but, il combine à la fois des articles de revue de littérature sur des thèmes choisis (DO, THOMAS-AGNAN et VANHEMS ; CHAKIR), des articles à vocation méthodologique (REY ; BA et BAUMONT ; MAYOR et PATUELLI) et des articles d'application (GUEVARA et AUTANT-BERNARD ; SARI ; DELBECQ, GUILLAIN et LEGROS ; GÉNIAUX, NAPOLÉONE et LEROUX ; ABILDTRUP, GARCIA et QUERÉ ; CHASCO et SANCHEZ) sur des données spatialement agrégées ou désagrégées, avec des approches structurelles ou non, sur des données de panel, des données microéconométriques ou en combinant économétrie spatiale et régressions quantiles.

3.1. Revues de littérature et articles méthodologiques

Le premier article, écrit par DO, THOMAS-AGNAN et VANHEMS (« Spatial reallocation of areal data – another look at basic methods »), consiste en une revue des problèmes posés par la combinaison de données spatiales mesurées en différentes échelles ou différentes partitions possibles de l'espace. En effet, lorsqu'on travaille sur données socio-économiques localisées, il est courant de disposer de ces données à des échelles différentes. Ainsi, dans les modèles hédoniques de prix immobiliers, il est usuel de régresser le prix d'une transaction sur les caractéristiques individuelles de ces transactions et sur des variables de contexte socio-économique mesurées à l'échelle du quartier ou à une échelle supérieure. Dans ce cas, une première possibilité est de contrôler la présence simultanée d'observations à des échelles différentes à travers, par exemple, des régressions hiérarchiques et/ou à travers des écarts-types robustes à l'hétéroscédasticité en groupe. Une seconde possibilité, détaillée dans cet article, est d'interpoler au niveau individuel des données disponibles à une échelle agrégée. Les données localisées peuvent également être disponibles pour des partitions différentes de l'espace car elles proviennent de sources administratives différentes. Il convient alors d'aligner toutes les données sur une partition unique : c'est le carroyage. Le problème général est celui du changement de support des données spatiales décrit par GOTWAY et YOUNG (2002). L'article proposé ici se concentre sur le **changement de support de surface à surface** et présente les trois grands types de techniques disponibles pour ce problème : les méthodes dasymétriques, les méthodes de lissage et les méthodes à base de régression. Un tableau récapitulatif et un exemple sont présentés qui permettront au praticien de mettre en œuvre ces méthodes.

Le deuxième article, proposé par CHAKIR (« L'espace dans les modèles économétriques d'utilisation des sols : enjeux méthodologiques et applications empiriques »), montre comment la prise en compte de la dimension spatiale dans les **modèles d'usage des sols** a évolué durant les trois dernières décennies. Ces modèles s'intéressent aux déterminants de l'utilisation des sols et des changements d'usage des sols ainsi qu'à leurs impacts sur la qualité de l'eau, la fragmentation des habitats ou la

déforestation. Dans cette littérature, la pratique empirique a suivi la disponibilité des données et l'évolution des techniques. En effet, les premiers travaux portaient sur des données agrégées et se basaient sur des modèles de part des usages. Les travaux sur données désagrégées à la parcelle sont plus récents et ne prenaient pas en compte l'autocorrélation spatiale ou cherchaient à évacuer le problème à travers des méthodes d'échantillonnage spatial. De fait, les modèles à la parcelle sont fondamentalement des modèles de choix discrets entre différents usages alors que le traitement de l'autocorrélation spatiale dans le cadre de ces modèles de choix discrets pose des problèmes techniques difficiles (FLEMING et MAE, 2004 ; SMIRNOV, 2010). L'article conclut par des pistes de recherche. L'une d'entre elle concerne précisément le problème de la disponibilité des données à différentes échelles administratives (départements, communes, petites régions agricoles), dont il est question dans l'article précédent.

L'article de REY (« Discrete regional distributional dynamics revisited ») se situe dans une large littérature s'intéressant à **l'évolution dans le temps et l'espace des disparités régionales de revenu par habitant** (HARRIS, 2011). L'hypothèse de convergence a fait ainsi l'objet de nombreuses analyses sur la base de différents concepts : σ -convergence, β -convergence... Cet article se situe dans la lignée de QUAH (1993) consistant à s'intéresser à l'évolution de l'ensemble de la distribution plutôt qu'à une économie représentative. Deux grandes approches existent. La première consiste à discrétiser la distribution par tête et d'estimer les probabilités de transition dans une chaîne de MARKOV discrète. La seconde consiste à analyser les densités inconditionnelles et conditionnelles des revenus par tête. L'article présenté ici se situe dans la première mouvance et présente un ensemble de considérations méthodologiques liées à cette procédure : la spécification des classes, la question de la pondération des transitions et le conditionnement géographique, afin de prendre en compte l'autocorrélation spatiale dans la distribution des revenus par habitant. Une application sur les états américains sur longue période (1929-2011) permet de mettre en perspective les résultats obtenus.

Le quatrième article, proposé par BA et BAUMONT (« Hétérogénéité des interactions et évaluation des effets de traitement : une approche par les effets de dépendance spatiale »), s'attache à former les bases d'un pont entre l'économétrie spatiale et **l'économétrie de l'évaluation des politiques publiques**. De nombreuses techniques ont été proposées afin de faire face au problème inhérent de l'évaluation, celui du biais de sélection lorsque les individus choisissent de participer ou non à une politique sur la base de caractéristiques observables ou inobservables. Dans le cadre du modèle causal de RUBIN, il est souvent fait l'hypothèse que le résultat d'une politique économique sur un individu ne dépend que des caractéristiques de cet individu et non du fait que d'autres individus aient été « traités » ou non. Il s'agit de l'hypothèse SUTVA : *Stable Unit Treatment Value Problem*. Cette hypothèse exclut donc toutes les externalités, les effets de pair ou les effets d'équilibre général. Son caractère très restrictif a fait l'objet récemment de nombreuses discussions (TCHETGEN TCHETGEN et VANDERWEELE, 2012). Dans ce contexte, on comprend bien pourquoi l'économétrie spatiale doit avoir un rôle important à jouer dans ces discussions

dans la mesure où l'un de ses objectifs majeurs est de capter les interactions inter-individuelles. Cet article s'inscrit dans cet objectif et propose un modèle d'évaluation à deux étapes. La première étape est un modèle probit avec autocorrélation spatiale des erreurs modélisant la décision des individus de participer à une politique tout en contrôlant par des facteurs inobservables spatialement autocorrélés. La deuxième étape consiste en un modèle autorégressif spatial pour prendre en compte la façon dont la réponse des individus à une politique est impactée par celle d'individus voisins. Une application est proposée traitant de la politique de revitalisation des zones rurales et montre que les effets de cette politique ne sont pas significatifs lorsque l'hypothèse SUTVA est levée. Nul doute que cet article suscitera une réflexion sur la façon dont l'économétrie de l'évaluation des politiques publiques doit s'approprier les techniques de l'économétrie spatiale.

L'article de MAYOR et PATUELLI (« *Spatial panel data forecasting over different horizons, cross-sectional and temporal dimensions* ») s'intéresse au problème de la **prévision spatio-temporelle**. Dans le cadre de données géo-localisées, ce problème peut s'appréhender de différentes façons. D'une part, on peut vouloir prévoir pour une partie de l'espace les valeurs d'une variable pour laquelle on dispose d'un modèle explicatif sur une autre partie de l'espace. D'autre part, lorsqu'on dispose de données géo-localisées à différentes périodes du temps, on peut vouloir prévoir les valeurs d'une variable pour des périodes futures sur la base d'un modèle estimé à une ou plusieurs périodes observables. C'est à ce deuxième problème que s'attaque l'article. Si le problème de la prévision de valeurs futures a une longue tradition en économétrie des séries temporelles ou en économétrie des données de panel (CLEMENTS et HENDRY, 2011), il prend une nouvelle dimension dans un contexte spatio-temporel. En effet, il faut alors prendre en compte un double aspect : d'une part, celui de l'autocorrélation temporelle entre valeurs d'une variable observée en différents points du temps et d'autre part, celui de l'autocorrélation entre valeurs d'une variable observée en différents points de l'espace. L'article compare différentes méthodes permettant d'appréhender la question de la prévision dans un cadre spatial : un modèle VAR spatial et un modèle de panel dynamique avec filtrage spatial. L'exercice de prévision sur des données de taux de chômage est mené sur trois pays, la France, l'Espagne et la Suisse et montre que le modèle VAR spatial est préférable lorsque le nombre d'observations temporelles est plus important que le nombre d'observations géolocalisées mais que cet avantage diminue pour des horizons de prévision éloignés.

3.2. Applications

Ce numéro spécial comprend également des applications d'économétrie spatiale reflétant la diversité des thématiques pouvant être abordées et par niveau décroissant d'agrégation spatiale des données soulevant des problèmes non triviaux. Ainsi, l'article de GUEVARA et AUTANT-BERNARD (« *Technological interdependence between South American countries : a spatial panel data growth model* ») représente un exemple d'un exercice d'économétrie spatiale structurelle, représentatif de l'évolution récente de la pratique en économétrie spatiale. En effet, en partant du modèle de croissance schumpétérien proposé par ERTUR et KOCH (2011) dans lequel il existe des

externalités spatiales de connaissances, on aboutit à une spécification économétrique incluant une variable autorégressive spatiale. Les auteurs évaluent alors le rôle de la productivité entre pays sud-américains sur la période 1990-2010. L'unité d'observation retenue est ici très agrégée puisqu'il s'agit de pays. Pour autant, les externalités spatiales, et donc la diffusion internationale de la technologie, sont une réalité entre pays d'Amérique du Sud avec un rôle central joué par le Brésil dans ce processus.

Les articles suivants mobilisent des données spatialement plus désagrégées. Ainsi, l'article de SARI (« Home-ownership and unemployment : a french test of the OSWALD hypothesis ») se penche sur le test de l'**hypothèse d'OSWALD** selon lequel l'une des origines du chômage dans les pays de l'OCDE est l'augmentation du nombre de propriétaires. En effet, l'accès à la propriété limite la mobilité et réduit ainsi les possibilités d'ajustement des marchés du travail. Cette hypothèse est testée sur les aires d'emploi en France métropolitaine en 2006. La méthode empirique présente la particularité de considérer non seulement l'autocorrélation spatiale, dans les termes d'erreurs et sous la forme d'une variable autorégressive spatiale, mais aussi l'endogénéité de l'une des variables explicatives : celui du statut résidentiel. Si l'économétrie spatiale s'est longtemps exclusivement attachée à traiter l'endogénéité impliquée par la présence d'une variable autorégressive spatiale, la question du traitement de l'endogénéité d'autres variables doit également être prise en compte (FINGLETON et LE GALLO, 2012). Les résultats obtenus ne corroborent pas l'hypothèse d'OSWALD, puisque le lien entre la part des propriétaires et le taux de chômage est négatif.

DELBECQ, GUILLAIN et LEGROS (« L'analyse de la criminalité à Chicago : de nouvelles perspectives offertes par l'économétrie spatiale à une question ancienne ») s'ancre dans une tradition importante en criminologie visant à capter **le rôle de la localisation dans les activités criminelles** (PIQUERO et WEISBURD, 2010). L'analyse économétrique proposée ici prend pour point de départ la théorie de la désorganisation sociale qui montre pourquoi la délinquance est concentrée dans l'espace. Les explications basées sur les effets de quartier ont rencontré un écho particulièrement important en économie en mettant en avant le rôle des effets de pair et la structure et la formation des réseaux dans les phénomènes de délinquance (voir par exemple PATACCHINI et ZÉNOU, 2012). Sur la base de données socio-économiques disponibles au niveau du *census tract* dans la ville de Chicago pour la période 2009-2011, les auteurs lient la proportion de crimes violents aux variables mises à en avant par la théorie de la désorganisation sociale et justifient l'utilisation de l'économétrie spatiale par la réduction du biais de variables omises, lorsque ces dernières sont spatialement autocorrélées et par les effets de débordement liés au découpage géographique administratif. L'économétrie est couplée aux méthodes de données de panel pour capter l'hétérogénéité individuelle inobservable.

L'article de GENIAUX, NAPOLÉONE et LEROUX (« Les effets prix de l'offre foncière ») présente l'intérêt de comparer les résultats obtenus par les modèles hédoniques spatiaux et ceux obtenus par une expérience quasi-naturelle pour évaluer les **effets prix de l'offre foncière** pour 306 communes en Ile-de-France. En effet, dans un

contexte de forte croissance des prix immobiliers, il s'agit d'évaluer la pertinence de l'intervention publique dans le fonctionnement des marchés fonciers et immobiliers à travers la régulation de l'offre foncière. Pour autant, l'analyse de la relation entre offre foncière et prix doit gérer à la fois l'hétérogénéité des contextes locaux, la présence d'effets de débordement et l'endogénéité de l'offre foncière. Sur la base de données fortement désagrégées (les parcelles), les auteurs estiment d'abord un modèle hédonique spatial en contrôlant le fait que l'offre foncière est endogène *via* une approche en termes de variables instrumentales, à l'instar de SARI (dans ce numéro). Ensuite, ils s'attaquent directement au problème de l'endogénéité *via* une approche contrefactuelle : celle des méthodes de type *matching*. Le lien entre les deux méthodes est assuré par la focalisation sur les communes ayant une offre foncière forte. La relation est positive dans les deux cas. En comparant une approche contrefactuelle à une approche spatiale, cet article constitue un complément à celui de BA et BAUMONT (dans ce numéro), qui plaide pour une approche directement spatiale des méthodes d'évaluation.

La question de l'endogénéité est également au cœur de l'article d'ABILDTRUP, GARCIA et KERE (« Land use and drinking water supply: a spatial switching regression model with spatial endogenous switching »). Ils s'intéressent en effet à **l'impact des usages du sol sur l'offre d'eau potable** dans le département du Rhône. L'économétrie spatiale est ici combinée avec un modèle à sélection d'échantillon afin de prendre en compte simultanément les interactions spatiales entre les services d'offre en eau et l'endogénéité du mode de gestion (public ou délégué). L'approche utilisée est celle de FLORES-LAGUNA et SCHNIER (2012) pour l'estimation d'un modèle à sélection d'échantillon avec autocorrélation spatiale des erreurs. L'équation de sélection consiste en modèle probit spatial visant à estimer les probabilités de choix du mode de gestion alors que l'équation d'intérêt est une équation expliquant les prix de l'eau potable. Les paramètres des deux équations sont estimés simultanément. Les résultats montrent que les prix de l'eau potable sont significativement moins élevés dans les communes où la part des forêts est plus importante. L'approche permet aussi de capter précisément l'influence du mode de gestion.

Le dernier article de ce numéro, celui de CHASCO et SANCHEZ (« Valuation of environmental pollution in the city of Madrid : an application with hedonic models and spatial quantile regressions »), s'intéresse à la façon dont les aménités environnementales (ici la pollution de l'air et le bruit ambiant) sont capitalisés dans les prix immobiliers. En mobilisant une base de données de 5 080 transactions sur la ville de Madrid, les auteurs estiment des **modèles hédoniques immobiliers**. Il s'agit donc de l'application utilisant les données les plus spatialement désagrégées de ce numéro, ici des observations ponctuelles. Les modèles hédoniques de prix immobiliers sont traditionnellement confrontés à de nombreux problèmes économétriques : hétérogénéité inobservée et hétéroscédasticité, multicollinéarité entre les variables explicatives, non-linéarité des relations, autocorrélation spatiale... L'approche de CHASCO et SANCHEZ est de considérer le double problème de l'hétérogénéité et de l'autocorrélation spatiale à travers l'estimation des régressions quantiles spatialisées. De fait, l'intérêt des régressions quantiles est de considérer non pas uniquement l'impact des variables explicatives sur une transaction correspondant au point moyen

de l'échantillon mais de considérer cet impact sur l'ensemble de la distribution des prix immobiliers. L'estimation en termes de régressions quantiles spatialisées est relativement récente en économétrie spatiale et cet article montre bien comment les aménités environnementales sont capitalisées de façon différente selon le quantile de la distribution des prix.

- 4 -

Conclusion

L'espace est complexe et fait de localisations, d'attributs de ces localisations et d'interrelations entre lieux mais aussi entre les lieux et leur(s) environnement(s). L'objectif de ce numéro spécial n'est évidemment pas de recenser tous les problèmes liés à ces complexités ni tous les défis et progrès actuels en analyse spatiale mais bien d'en présenter une sélection par le biais d'articles choisis. Nous espérons à ce titre que ce numéro sera de nature à susciter des réflexions et conduire à une systématisation de la mise en œuvre des méthodes de la statistique et de l'économétrie spatiales en sciences régionales francophones.

Depuis quelques décennies déjà, les chercheurs ont à cœur de contrôler les effets pervers de l'espace en statistique et en économétrie, pour mieux construire/vérifier les théories et – par voie de conséquence – mieux aider la prise de décision en sciences régionales. Des thèmes comme l'autocorrélation spatiale, l'échelle spatiale d'analyse et l'emboîtement d'échelles, la taille et forme de la maille d'analyse (MAUP), les dynamiques spatio-temporelles, l'erreur écologique (« ecological fallacy »), ou l'endogénéité conduisent à la publication de nombreux articles théoriques et empiriques dans des disciplines variées. Ces thématiques sont récurrentes car leur solution n'est ni simple ni unique. Elles font l'objet, directement ou indirectement, des quelques articles rassemblés dans ce numéro spécial. L'avènement rapide et récent des grandes bases de données, « big data », ouvre clairement de nouveaux horizons, particulièrement en matière d'interrelations et de connectivités, tout en rendant encore plus pressant le besoin de résoudre ces problèmes statistiques ainsi que ceux qui apparaîtront dans le futur. Le monde de demain sera hyperconnecté et livrera des masses gigantesques de données. Encore faudra-t-il comprendre ce qu'elles nous disent et mesurer correctement le rôle réel de l'espace dans les résultats. En d'autres termes, l'économétrie des interactions est devant nous.

Bibliographie

- ANSELIN L (1988) *Spatial econometrics: methods and models*. Kluwer Academic Publishers, Dordrecht.
- ANSELIN L (2010) Thirty years of spatial econometrics. *Papers in Regional Science* 89 (1): 3-25.
- ANSELIN L, SYABRI I, KHO Y (2006) GeoDa, an introduction to spatial data analysis. *Geographical Analysis* 38: 5-22.
- ARBIA G (1988) *Spatial data configuration in statistical analysis of regional economic and related problems*. Kluwer Academic Publishers, Dordrecht.
- ARBIA G (2011) A lustrum of SEA : recent research trends following the creation of the Spatial Econometrics Association (2007-2011). *Revue d'Économie Régionale et Urbaine* 2012 (1) : 3-18.
- BALTAGI BH (2001) *A companion to theoretical econometrics*, Blackwell, Oxford.
- BALTAGI BH (2013) *Econometric analysis of panel data*, Wiley, New York.
- BEENSTOCK M, FELSENSTEIN D (2012) Nonparametric estimation of the spatial connectivity matrix using spatial panel data. *Geographical Analysis* 44 (4): 386-397.
- BHATTACHARJEE A, JENSEN-BUTLER C (2013) Estimation of the spatial weights matrix under structural constraints *Regional Science and Urban Economics* 43 (4): 617-634.
- BIVAND RS, PEBESMA EJ, GOMEZ-RUBIO V (2008) *Applied spatial data analysis with R*, Springer, Berlin.
- BRUECKNER JK (1998) Testing for strategic interaction among local governments: the case of growth controls. *Journal of Urban Economics* 44 (3): 438-467.
- CAMERON AC, TRIVEDI PK (2005) *Microeconometrics: methods and applications*, Cambridge University Press, Cambridge.
- CLEMENTS MP, HENDRY DF (2011) *The Oxford handbook of economic forecasting*. Oxford University Press.
- CLIFF AD, ORD JK (1969) The problem of spatial autocorrelation. In: SCOTT AJ (eds) *Studies in Regional Science*. Pion, Londres.
- CLIFF AD, ORD JK (1973) *Spatial autocorrelation*. Pion, Londres.
- ERTUR C, KOCH W (2011) A contribution to the Schumpeterian growth theory and empirics. *Journal of Economic Growth* 16 (3): 215-255.
- FINGLETON B, LE GALLO J (2012) Endogénéité et autocorrélation spatiale : quelle utilité pour le modèle de Durbin. *Revue d'Économie Régionale et Urbaine* 1 : 3-18.
- FLEMING MM, MAE F (2004) Techniques for estimating spatially dependent discrete choice models. In: ANSELIN L, FLORAX RJGM (eds) *Advances in spatial econometrics*. Springer-Verlag, Heidelberg.
- FOTHERINGHAM A, BRUNSDON C, CHARLTON M (2000) *Quantitative geography: perspectives on spatial data analysis*, Sage, Londres.
- FOTHERINGHAM A (1997) Trends in quantitative methods I: Stressing the local. *Progress in Human Geography* 21 (1): 88-96.
- FOTHERINGHAM A (1998) Trends in quantitative methods II: Stressing the computational. *Progress in Human Geography* 22 (1): 283-292.
- FOTHERINGHAM A (1999) Trends in quantitative methods III: Stressing the visual. *Progress in Human Geography* 23 (4): 597-606.
- GOTWAY CA, YOUNG LJ (2002) Combining incompatible spatial data. *Journal of the American Statistical Association* 97 (458): 632-648.
- GRIFFITH DA (2012) Spatial statistics: a quantitative geographer's perspective. *Spatial Statistics* 1: 3-15.
- HARRIS R (2011) Models of regional growth: past, present and future. *Journal of Economic Surveys* 25 (1): 109-155.
- HSIEH C-S, LEE L-F (2015) A social interactions model with endogenous friendship formation and selectivity, à paraître.
- KENNEDY P (2003) *A guide to econometrics*. MIT Press, Cambridge, MA.

- KELEJIAN HH, PIRAS G (2014) Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Regional Science and Urban Economics* 46: 140-149.
- LEE LF (2011) Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrics* 72 (6): 1899-1925.
- LE GALLO J (2014) Cross-section spatial regression models. In: FISCHER MM, NIJKAMP P (eds) *Handbook of regional science*. Springer-Verlag, Berlin.
- LEUNG Y (2009) *Knowledge discovery of spatial data*, Springer-Verlag, Berlin.
- MANSKI C (2000) Economic analysis of social interactions. *Journal of Economic Perspectives* 14: 15-136.
- MATYAS L, SEVESTRE P (1998) *The econometrics of panel data, fundamentals and recent developments in theory and practice*, Springer-Verlag, Berlin.
- MILLS TC, PATTERSON K (2006) *Palgrave handbook of econometrics: Volume 1, Econometric theory*, Palgrave Macmillan, Basingstoke.
- PAELINCK JH, KLAASSEN L (1979) *Spatial econometrics*. Saxon House, Farnborough.
- PATACCHINI E, ZÉNOU Y (2012) Juvenile delinquency and conformism. *Journal of Law, Economics & Organization* 28 (1): 1-31.
- PIROTTE A (2011) *Économétrie des données de panel*, Economica, Paris.
- PIQUERO AR, WEISBURD D (2010) *Handbook of quantitative criminology*, Springer-Verlag, Berlin.
- PONSARD C (1988) *Analyse économique spatiale*, Presses Universitaires de France, Paris.
- SMIRNOV OA (2010) Modeling spatial discrete choice. *Regional Science and Urban Economics* 40 (5): 292-298.
- TCHETGEN TCHETGEN EJ, VANDERWEELE TJ (2012) On causal inference in the presence of interference. *Statistical Methods in Medical Research* 21: 55-75.
- THISSE JF (1997) L'oubli de l'espace dans la pensée économique. *Région et Développement* 6 : 1-29.
- ULLAH A, GILES DEA (1998) *Handbook of applied economic statistics*, Marcel Dekker, New York.
- UPTON GJ, FINGLETON B (1985) *Spatial data analysis by example, Volume 1*, Wiley, New York.
- QU X, LEE L-F (2015) Estimating a spatial autoregressive model with an endogenous spatial weight matrix. *Journal of Econometrics* 184 (2): 209-232.
- WOOLDRIDGE JM (2010) *Econometric analysis of cross section and panel data*, MIT Press.