



HAL
open science

Disparités géographiques et convergence des régions européennes : une approche par l'économétrie spatiale

Julie Le Gallo

► **To cite this version:**

Julie Le Gallo. Disparités géographiques et convergence des régions européennes : une approche par l'économétrie spatiale. Economies et finances. Université de Bourgogne, 2002. Français. NNT : . tel-00006432

HAL Id: tel-00006432

<https://theses.hal.science/tel-00006432>

Submitted on 9 Jul 2004

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

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

UNIVERSITÉ DE BOURGOGNE
FACULTE DE SCIENCE ÉCONOMIQUE ET DE GESTION

**DISPARITES GEOGRAPHIQUES ET CONVERGENCE DES
REGIONS EUROPEENNES :
UNE APPROCHE PAR L'ECONOMETRIE SPATIALE**

THESE

Présentée en vue de l'obtention
du Doctorat en Analyse et Politique Économiques

par

Julie Le Gallo

Soutenue publiquement
le 21 mai 2002

Membres du jury :

Mme Catherine BAUMONT	Professeur à l'Université de Bourgogne, Directeur de thèse
M. Gilbert BENHAYOUN	Professeur, Université à Aix-Marseille III
M. Cem ERTUR	Maître de Conférences à l'Université de Bourgogne
M. Bernard FINGLETON	Lecturer, University of Cambridge
M. Raymond J.G.M. FLORAX	Professeur à Vrije Universiteit, Amsterdam, Rapporteur
M. Pierre MORIN	Professeur, INSEE et Université de Metz, Rapporteur
Mme Marie-Claude PICHERY	Professeur à l'Université de Bourgogne, Directeur de thèse

« La faculté n'entend donner aucune approbation, ni improbation aux opinions émises dans les thèses. Ces opinions doivent être considérées comme propres à leurs auteurs »

Remerciements

Mes remerciements s'adressent en premier lieu à Catherine Baumont et Marie-Claude Pichery qui ont accepté de diriger mon travail de thèse.

Je remercie Catherine Baumont pour sa grande disponibilité tout au long de mes recherches, pour ses conseils judicieux, sa confiance, ses encouragements et son aide lors de l'élaboration de ce document.

Je remercie Marie-Claude Pichery pour m'avoir donné envie d'étudier plus avant l'économétrie, pour ses conseils et pour la grande confiance qu'elle m'a accordée dans mes travaux.

Je remercie Cem Ertur pour son suivi régulier de mes recherches, sa disponibilité et ses précieux conseils.

Je remercie Dominique Peeters pour m'avoir accueillie à l'Université de Louvain-la-Neuve où j'ai pu approfondir certains aspects de mes recherches.

Je remercie Raymond Florax et James LeSage pour les différents commentaires et remarques qu'ils ont pu apporter à mon travail et qui ont éclairé ma pensée.

Je remercie Geoffrey Hewings et Luc Anselin pour m'avoir invitée à l'Université d'Urbana en Illinois-Champaign lors d'un séminaire et pour les discussions fructueuses que nous avons eues à cette occasion.

Je remercie Jean-Marie Huriot pour ses encouragements et Françoise Bourdon qui n'a jamais hésité à me venir en aide en de nombreuses occasions de ma vie de doctorante.

Un grand merci à Rachel Guillain pour son soutien inestimable dans les moments difficiles, ses encouragements, son aide dans la correction de ce document.

Mes remerciements vont aussi à Luisito Bertinelli pour son accueil lors de mon séjour à Louvain-La-Neuve et ses conseils avisés lors de la relecture de mon travail.

Merci également à Xavier Brierre pour son aide technique à l'occasion de la finition de ma thèse, à Elisabeth Penez pour son aide lors des démarches administratives et à Anne-Marie Picketty pour la recherche d'articles et d'ouvrages.

Je tiens enfin à remercier mes parents qui, par leur soutien, leur affection et leurs encouragements, ont grandement contribué à l'aboutissement de cette thèse.

SOMMAIRE

INTRODUCTION	1
CHAPITRE 1 : DISPARITES GEOGRAPHIQUES ET CONVERGENCE - UN NOUVEAU DEBAT	12
CHAPITRE 2 : ECONOMETRIE SPATIALE	95
CHAPITRE 3 : ANALYSE EXPLORATOIRE DES DISPARITES REGIONALES DE PIB PAR TETE EN EUROPE	201
CHAPITRE 4 : ANALYSE ECONOMETRIQUE SPATIALE DE LA β-CONVERGENCE ENTRE LES REGIONS EUROPEENNES	289
CHAPITRE 5 : CONVERGENCE DES REGIONS EUROPEENNES ET CHAINES DE MARKOV	355
CONCLUSION	418
BIBLIOGRAPHIE	426
LISTE DES TABLEAUX ET DES FIGURES	457
TABLE DES MATIERES	464

Introduction

Le rapprochement des niveaux de développement économique, des performances en termes de croissance et la réduction du taux de chômage constituent une préoccupation majeure des politiques régionales européennes depuis l'intégration, dans les années 80, de la Grèce, de l'Espagne et du Portugal. En effet, dans un contexte général de ralentissement de la croissance économique et de difficultés de restructuration pour de nombreuses régions, l'intégration de ces nouveaux pays membres a mis en lumière le problème de l'importance des disparités régionales au sein de la Communauté Européenne et a provoqué une refonte de la politique régionale. Des écarts importants de niveaux de vie de ses citoyens ou entre leurs potentiels économiques apparaissent effectivement intolérables dans une véritable communauté.

La réforme des fonds structurels, ayant pour objectif de renforcer la cohésion économique et sociale de la Communauté Européenne, a été décidée en 1988 et mise en œuvre à partir de 1989. Cette notion de cohésion possède d'ailleurs une dimension géographique importante puisqu'elle a pour but de « réduire l'écart entre les niveaux de développement des diverses régions et le retard des régions les moins favorisées, y compris les zones rurales » (Commission Européenne, 1996, p. 13). Cette réforme vise à augmenter l'efficacité des trois fonds structurels (Le FEDER¹, créé en 1975, le FSE² et le FEOGA³ dans sa section « orientation ») et à les doter de moyens financiers accrus. Un des éléments les plus importants a porté sur la fixation de cinq objectifs prioritaires auxquels doivent se rattacher les différents Fonds : l'objectif 1 vise à promouvoir le développement et l'ajustement des régions en retard de développement, l'objectif 2 doit permettre la reconversion des régions frontalières ou parties de régions (bassin d'emplois ou communautés urbaines) gravement affectées par le déclin industriel, l'objectif 3 a pour but de combattre le chômage de longue durée, l'objectif 4 vise à faciliter l'insertion professionnelle des jeunes et l'objectif 5 doit permettre d'accélérer l'adaptation des structures agricoles (objectif 5a) et de promouvoir le développement des zones rurales (objectif 5b)⁴.

Malgré ces différentes initiatives, il subsiste toujours, depuis deux décennies, un « problème régional » au sein de la Communauté Européenne (Dignan, 1995). Ce problème se manifeste en particulier par l'existence d'un grand nombre de régions pauvres dont le PIB par tête est inférieur à 75% de la moyenne communautaire et par l'hétérogénéité du processus de rattrapage des régions les plus riches par les régions les plus pauvres. Il convient également de souligner que les régions européennes sont caractérisées par la persistance d'une structure duale entre un centre développé et une périphérie en difficulté : « c'est au niveau régional et, en particulier, entre le centre et la périphérie, que les disparités économiques de l'Europe sont les plus manifestes » (Commission Européenne, 1996, p. 21). Les disparités économiques semblent ainsi être associées à une dimension géographique affirmée sous la forme d'une polarisation Nord-Sud et d'une concentration géographique de la richesse et de la pauvreté.

¹ Fonds Européen de Développement Régional.

² Fonds Social Européen.

³ Fonds Européen d'Orientation et de Garanti Agricole.

⁴ La première période d'application de la réforme concerne 1989-1993. Depuis, d'autres périodes de programmation ont été mises en œuvre : 1994-1999 et 2000-2006 avec une redéfinition de certains des objectifs.

Face à ces observations, le processus de convergence régionale en Europe apparaîtrait comme faible, nuancé et géographiquement marqué. Il semble alors intéressant d'étudier plus précisément les relations entre convergence économique et répartition spatiale des activités.

Sur un plan théorique, les réponses apportées sont ambiguës et mobilisent à la fois les théories de la croissance et celles de l'économie géographique.

D'un côté, une lecture rapide du modèle de croissance de Solow-Swan tend à prédire la convergence des régions vers un même niveau de PIB. Dans les conditions décrites par le modèle, si les économies sont similaires à tous points de vue excepté pour leur dotation initiale de produit par tête, alors les économies caractérisées par un niveau de produit par tête moins élevé ont tendance à croître plus rapidement que les économies plus riches initialement. Ainsi, toute disparité initiale, en termes de produit par tête, génère, dans ce modèle, son propre mécanisme de résorption. Ces conclusions, optimistes quant au rattrapage des économies riches par les économies pauvres, ont été remises en cause par les théories de la croissance endogène dans lesquelles les avantages initiaux des régions les favorisées sont renforcés.

D'un autre côté, se développent au début des années 1990, les modèles de l'économie géographique, qui étudient les mécanismes économiques de formation des agglomérations géographiques d'activités économiques. Ces modèles montrent que l'intégration économique et la croissance des régions s'accompagnent le plus souvent d'une polarisation croissante des territoires. Ils permettent de comprendre la formation des disparités régionales, c'est-à-dire la façon dont l'espace s'homogénéise et se différencie, et de montrer comment ces disparités régionales interagissent avec la croissance économique. Ainsi, le processus de convergence et les schémas de répartition spatiale semblent donc être liés.

L'objectif de cette thèse est de préciser les liens existant entre localisation géographique et disparités économiques en Europe. Plus précisément, il s'agit de montrer que la prise en compte des disparités géographiques renouvelle le débat sur la convergence des économies. Ces disparités ne peuvent pas être négligées et elles permettent une amélioration de la compréhension du processus de convergence entre les économies. Dans cette optique, notre réflexion s'articule autour des questions suivantes :

- Quelles sont les caractéristiques des disparités géographiques ?
- Comment doit-on prendre en compte ces disparités géographiques dans l'analyse de la convergence des économies ?
- Quelles sont les conséquences de ces disparités sur le processus de convergence des économies ?

Afin de traiter ces questions, deux champs théoriques principaux sont mobilisés. D'une part, le problème de la convergence peut être relié au modèle néoclassique de croissance et aux modèles de croissance endogène. D'autre part, les phénomènes d'agglomération et de polarisation spatiale des activités économiques peuvent être reliés aux modèles d'économie géographique, visant à expliquer ces phénomènes et à montrer que la croissance économique est reliée aux schémas de répartition des activités sur le territoire.

Dans ce cadre, notre démarche est conceptuelle et empirique. Nous montrons pourquoi *une méthodologie particulière doit être construite et utilisée pour l'analyse empirique des disparités régionales*. En effet, la prise en compte de ces disparités impose la mise au point de tests de convergence et de modélisations économétriques appropriés. Les méthodes employées permettent d'éclairer et d'illustrer les différents débats théoriques sur le phénomène de convergence à l'aide d'une étude empirique portant sur l'analyse des PIB par tête de 138 régions européennes sur la période 1980-1995. Cette période est particulièrement intéressante pour l'analyse de la convergence car elle correspond à une intégration économique croissante entre les régions européennes, marquée à la fois par l'accélération de la libéralisation des échanges et par l'élargissement du marché à la suite de l'entrée de la Grèce, de l'Espagne et du Portugal. Elle correspond également, suite à la réforme des fonds structurels, à la mise en place d'importantes politiques régionales visant à harmoniser les potentialités de développement des territoires.

Notre objectif n'est pas de tester une équation de convergence qui serait déduite d'un modèle d'économie géographique de la croissance, comme l'ont fait, par exemple, Hanson (1998a), Combes et Lafourcade (2001) et Teixeira (2001). Nous montrons néanmoins que la prise en compte de la dimension géographique des données, c'est-à-dire la localisation des observations les unes par rapport aux autres, dans les différents concepts de convergence permet d'intégrer certains enseignements de ces modèles et d'évaluer dans quelle mesure les résultats habituellement obtenus sont modifiés. Les modèles de l'économie géographique servent à justifier, par leurs enseignements en matière d'analyse des disparités économiques et géographiques, pourquoi cette dimension spatiale devrait être prise en compte dans le cadre de ces modèles de convergence.

Les principaux résultats empiriques indiquent que les disparités économiques et leurs évolutions dans le temps sont fortement conditionnées par la localisation des régions au sein de l'espace européen. Tout d'abord, il apparaît que la répartition des PIB par tête est caractérisée par une forte concentration spatiale des richesses et de la pauvreté et par un schéma de polarisation Nord-Sud. Ces caractéristiques sont persistantes sur toute la période, malgré l'intégration économique croissante des régions européennes. Ensuite, la dépendance des régions à leur environnement géographique et la polarisation Nord-Sud sont confirmées par des régressions économétriques liant le taux de croissance annuel moyen entre 1980 et 1995 au PIB par tête en 1980. Ces équations de β -convergence spatialisées permettent alors de montrer, d'une part, l'existence d'effets de débordement géographiques significatifs entre le taux de croissance d'une région et celui de ses voisines et d'autre part, de différencier le processus de convergence entre les régions du Nord et les régions du Sud. Enfin, la mobilité d'une région au sein de la distribution des PIB par tête est très faible et fortement conditionnée par la richesse de ses voisines, ce qui met alors en évidence le phénomène de « verrouillage géographique » suggéré dans les théories de l'économie géographique.

La mise en évidence de ces résultats passe par une argumentation développée en cinq chapitres.

Le premier chapitre a pour objet de montrer pourquoi les disparités géographiques doivent être prises en compte dans les modèles ayant pour but de tester l'hypothèse de convergence.

Tester l'hypothèse de convergence impose au préalable de donner un contenu formel à cette notion. La notion de convergence et les tests associés ont déjà fait l'objet de nombreuses synthèses. Ainsi, notre lecture et notre présentation des notions de convergence est principalement axée sur le principe suivant : chaque notion de convergence est bien spécifique et elle fournit en ce sens une information particulière sur la façon dont les disparités économiques sont théoriquement conceptualisées, sur la façon dont la réduction des disparités économiques est envisagée et enfin sur l'évolution temporelle de ces disparités. Le but des deux premières sections de ce premier chapitre consiste donc à centrer notre présentation sur les liens entre les concepts de convergence et les modalités économétriques de leur étude.

Plusieurs notions de convergence ont été proposées, conduisant alors à une grande variété d'hypothèses testées et de méthodes économétriques employées. Une première notion de convergence des économies est déduite du modèle néoclassique de Solow-Swan (1956) et a donné lieu à des interprétations en termes de rattrapage des économies. Ce modèle repose sur des hypothèses de rendements d'échelle constants et de décroissance de la productivité marginale du capital. A long terme, le taux de croissance du capital et du produit par tête est égal au taux de progrès technique, supposé exogène. Puisque le progrès technique est considéré comme un bien public librement accessible à tous, ce modèle aboutit alors à une convergence en taux : à long terme, toutes les économies possèdent le même de taux de croissance. En outre, si les paramètres caractérisant les économies (taux d'épargne, taux de dépréciation et taux de croissance du capital humain) sont identiques, alors le modèle aboutit également à un résultat de convergence en niveau : toutes les économies convergent vers le même niveau de produit par tête.

Ces deux aspects recourent les notions de β -convergence absolue et de β -convergence conditionnelle. Dans le premier cas, on cherche à vérifier si le taux de croissance du produit par tête d'une économie est d'autant plus élevé que cette économie a un produit par tête initial faible. En d'autres termes, on teste l'existence d'une corrélation négative entre le taux de croissance du produit par tête et le produit par tête initial lorsque les économies sont toutes supposées atteindre un même état régulier. En revanche, dans le dernier cas, on suppose que le taux de croissance d'une économie est d'autant plus fort que la distance qui la sépare de son propre état régulier est élevée. Il convient de souligner que les régressions permettant de tester ces hypothèses sont dérivées du modèle néoclassique de croissance (ou de son extension avec capital humain proposée par Mankiw *et al.*, 1992). Cependant, le test de l'hypothèse de β -convergence ne permet pas de discriminer le modèle néoclassique des modèles de croissance endogène pour lesquels la divergence des économies est souvent admise. De ce point de vue, l'interprétation des régressions de β -convergence est relativement ambiguë.

D'autres tests de convergence ont été proposés dans la littérature. Ainsi, l'hypothèse de clubs de convergence implique que, selon les conditions initiales des économies, ces dernières convergent vers différents niveaux d'équilibre de produit par tête. Ces modèles sont

issus des modèles de croissance néoclassiques avec hétérogénéité des agents (Galor, 1996) ou des modèles de croissance avec équilibres multiples (Azariadis et Drazen, 1990). Le test de σ -convergence porte sur l'évolution dans le temps d'un indicateur de dispersion de la distribution des produits par tête de l'échantillon étudié. Quah (1993a, 1993b, 1996) propose quant à lui d'étudier l'évolution de la distribution des produits par tête et de la mobilité des économies dans cette distribution. Enfin, Bernard et Durlauf (1995, 1996) proposent le concept de convergence stochastique, basé sur l'étude des propriétés des séries temporelles.

Tous ces concepts de convergence posent des problèmes d'interprétation et d'estimation que nous détaillons dans la deuxième section du premier chapitre. Cependant, nous soulignons qu'un problème supplémentaire, rarement évoqué, existe. En effet, tous ces concepts de convergence supposent que les économies étudiées sont des « îles isolées » (Quah, 1996b), n'interagissant pas entre elles. Pourtant, entre pays, et plus encore entre les régions, les interactions spatiales sont nombreuses et la position absolue et relative des régions dans les espaces étudiés n'est pas neutre. Plus précisément, on observe le plus souvent une concentration spatiale des économies pauvres et des économies riches. Tous ces éléments ne peuvent pas être négligés car, sur un plan économétrique, la répartition particulière des données qui résulte de ces phénomènes d'agglomération peut impliquer une corrélation des erreurs qui devrait être prise en compte. Par ailleurs, les phénomènes de polarisation dans l'espace (comme la polarisation de type Nord-Sud) sont susceptibles de correspondre à des phénomènes de convergence en clubs et devraient également être testés.

Ces « oublis » sont d'autant plus étranges que les phénomènes d'agglomération et de polarisation spatiale des activités sont bien connus et qu'ils peuvent être reliés aux modèles d'économie géographique. La troisième section détaille quelques uns de ces modèles afin de montrer pourquoi les activités économiques ne sont pas réparties aléatoirement dans l'espace, comment des phénomènes de polarisation peuvent apparaître dans des espaces économiques intégrés et enfin, comment les schémas de répartition géographique interagissent avec les mécanismes de croissance des économies.

Une fois définies les raisons permettant de justifier pourquoi la dimension spatiale doit être intégrée dans l'analyse de la convergence des économies, le deuxième chapitre a pour objectif de présenter les méthodes statistiques et économétriques permettant de réaliser cet objectif : il s'agit des techniques de la statistique et de l'économétrie spatiales.

Les phénomènes précédemment décrits peuvent être reliés à deux concepts. Ainsi, la concentration géographique de valeurs similaires fait référence à la notion d'autocorrélation spatiale et la polarisation des activités en un schéma centre-périphérie peut être reliée à la notion d'hétérogénéité spatiale. Dans une première section, nous montrons déjà que la topologie du système est spécifiée de manière exogène par les matrices de poids. Ces dernières permettent la construction de variables spécifiques, appelées variables décalées, qui relient une observation à ses voisines. La notion de voisinage peut être définie sur la base de l'existence d'une frontière commune entre les observations ou en fonction de la distance qui sépare les centroïdes des unités spatiales. Nous présentons alors les méthodes statistiques permettant, dans une série spatiale, de détecter les deux effets spatiaux. Ces techniques,

dénommées « analyse exploratoire des données spatiales », sont récentes (Getis et Ord, 1992 ; Anselin, 1995, 1996) et elles permettent de caractériser la forme de l'autocorrélation et de l'hétérogénéité spatiales. Plus précisément, deux formes d'autocorrélation spatiale peuvent être détectées. D'une part, la forme globale indique si, au sein de l'espace étudié, la concentration géographique de données similaires est significative ou non (statistique I de Moran ou statistique c de Geary). D'autre part, une forme locale précise, au sein du schéma global précédent, où se localisent les concentrations géographiques significatives de valeurs similaires (statistiques de Getis-Ord et statistiques LISA) et si certaines observations et leurs voisines se détachent de cette tendance (diagramme de Moran et statistiques LISA).

Dans une deuxième section, les méthodes économétriques permettant de prendre en compte l'autocorrélation spatiale sont détaillées. La présentation est principalement axée sur les modèles linéaires en coupe transversale même si quelques extensions, portant sur les modèles de données de panel, les modèles à variables qualitatives ou les modèles d'équations simultanées sont rapidement présentés. La prise en compte de l'autocorrélation spatiale peut s'effectuer à l'aide de plusieurs spécifications conduisant à des interprétations différentes des coefficients associés aux effets de débordement géographiques. Ainsi, dans le modèle autorégressif, l'autocorrélation spatiale porte sur la variable expliquée alors que dans le modèle régressif croisé, l'autocorrélation spatiale porte sur une ou plusieurs variables explicatives, ce qui permet de capter directement les effets de débordement géographiques associés à ces variables. Enfin, dans le modèle avec autocorrélation spatiale des erreurs, des effets de débordement géographiques peuvent être indirectement associés aux différentes variables de la régression. Des problèmes d'estimation se posent dans le modèle autorégressif et le modèle à erreurs spatialement autocorrélées : les estimateurs obtenus à l'aide de la méthode des Moindres Carrés Ordinaires ne sont pas convergents et/ou sont inefficients. Nous détaillons alors les méthodes d'estimation adaptées à ces modèles (méthode du maximum de vraisemblance, des variables instrumentales et des moments généralisés) ainsi que les tests d'inférence statistique permettant de déterminer la spécification la plus adaptée pour l'autocorrélation spatiale.

Dans la troisième section, nous abordons le phénomène de l'hétérogénéité spatiale qui se traduit par une différenciation des comportements dans l'espace. Dans une régression économétrique, ce phénomène est modélisé à l'aide de coefficients variables et/ou d'une hétéroscédasticité variable dans l'espace. Les méthodes présentées sont relativement « standards », même si leur adaptation dans le cadre spatial permet de mettre en évidence des interprétations intéressantes, comme par exemple une instabilité structurelle liée à la localisation des observations ou une variation continue des paramètres dans tout l'espace étudié. La présence simultanée d'autocorrélation et d'hétérogénéité spatiales est plus complexe à traiter et les méthodes d'estimation et d'inférence doivent être adaptées. Nous présentons alors un panorama de ces méthodes. Ce problème est d'autant plus important que les liens unissant ces deux effets spatiaux sont complexes et il convient alors de mener, au préalable, des analyses exploratoires des données spatiales pour approcher les formes prises par ces effets.

Les trois chapitres suivants ont pour objet d'appliquer ces méthodes en conduisant différentes études empiriques du processus de convergence au sein d'un échantillon de 138 régions européennes sur la période 1980-1995. D'une part, nous mettons en pratique les méthodologies exposées dans le chapitre 2 et d'autre part, nous mettons en évidence la façon dont l'environnement géographique des régions influence les phénomènes de convergence.

Le troisième chapitre présente comment l'analyse exploratoire des données spatiales permet de caractériser les phénomènes d'autocorrélation et d'hétérogénéité spatiales dans la répartition des PIB par tête en niveau et en taux de croissance.

La base de données Eurostat-Regio, l'échantillon utilisé et les problèmes méthodologiques associés à l'utilisation de cette base de données sont d'abord détaillés. Pour étudier le processus de convergence, il est souhaitable de disposer d'une période la plus longue possible. Pour atteindre cet objectif, tout en travaillant sur des données homogènes, tant au niveau des modalités de calcul que des unités statistiques, nous n'avons pu retenir, sur cette base de données Regio, que 138 régions européennes sur toutes les années de la période 1980 à 1995 : l'Allemagne (30), la Belgique (11), le Danemark (1), l'Espagne (16), la France (21), la Grèce (13), l'Italie (20), le Luxembourg (1), les Pays-Bas (9), le Portugal (5) et le Royaume-Uni (11). Outre ce problème de disponibilité des données, deux autres problèmes sont discutés. Ils concernent le découpage régional utilisé dans les statistiques européennes et les méthodes mises en œuvre pour corriger les PIB des pouvoirs d'achat. Nous détaillons ainsi les modes de découpage en NUTS (Nomenclature des Unités Territoriales Statistiques) et rappelons les avantages et les inconvénients associés aux séries des PIB régionaux par tête en Ecus et en SPA (Standard de Pouvoir d'Achat). Aucune série n'étant « meilleure » que l'autre pour nos études empiriques, il convient d'utiliser les deux séries même si l'utilisation des séries en SPA produit un certain lissage des disparités régionales. Nous prenons soin de mentionner les différences les plus importantes de résultats obtenus selon la série utilisée et adoptons, si nécessaire, les méthodes statistiques appropriées à l'étude de chaque série. Enfin, le problème de la robustesse des résultats associé au choix de la matrice de poids, est soulevé. Compte tenu de la configuration géographique propre à l'Union Européenne et à notre échantillon, nous ne choisissons pas de travailler avec des matrices de contiguïté, mais optons pour des matrices de distance et des matrices dites « des plus proches voisins ». Ainsi, chaque étude empirique effectuée dans la thèse sera conduite avec 6 matrices de poids. Les résultats obtenus restent similaires quelles que soient les matrices de poids utilisées, ce qui atteste de leur robustesse, et une méthode d'étude de cette robustesse est proposée pour certaines de nos analyses empiriques.

Dans la deuxième section, nous étudions les phénomènes d'autocorrélation spatiale sur les niveaux de PIB par tête pour chaque année de la période 1980-1995. Dans ce cadre particulier, l'autocorrélation spatiale, mesurée par la statistique I de Moran, est globalement positive : elle se caractérise par une concentration de régions possédant des PIB par tête similaires. Les méthodes de l'analyse exploratoire locale des données spatiales permettent notamment de repérer les concentrations significatives de richesse et de pauvreté. Les premières se situent au « cœur » de l'Europe alors que les secondes concernent les régions

périphériques, la plupart étant éligibles à l'objectif 1. Peu de régions se détachent significativement de ce schéma global au cours de la période. Une première conclusion, portant sur les PIB par tête en niveaux, est donc la détection d'une polarisation Nord-Sud persistante, ce qui traduit une forme particulière d'hétérogénéité spatiale. Les statistiques locales posent un problème de comparaisons multiples et nous suggérons, dans le cadre des matrices des plus proches voisins, une approximation permettant de rendre les procédures de tests moins restrictives que celles utilisées habituellement. La robustesse de ces résultats aux différentes matrices de poids est également évaluée, à l'aide d'une application originale des matrices de transition.

La même analyse, conduite sur les taux de croissance annuels moyens entre 1980 et 1995, permet de mettre en évidence les liens entretenus par le dynamisme d'une région avec ses régions voisines. Cette section permet d'aboutir à une première évaluation de l'évolution des disparités géographiques en termes de taux de croissance en relation avec le PIB par tête. L'hétérogénéité des processus de croissance est confirmée en observant l'enrichissement continu de certaines régions malgré un niveau de départ relativement élevé alors que d'autres régions stagnent, voire régressent.

Dans le quatrième chapitre, les résultats exploratoires fournis dans le chapitre précédent sont repris afin d'évaluer, dans le cadre de modèle de β -convergence, le rôle joué par les facteurs géographiques : la localisation des régions et les effets de débordement géographiques.

Dans un premier temps, nous considérons uniquement le phénomène de l'autocorrélation spatiale, ce qui permet de modéliser de façon particulière l'équation de β -convergence et d'interpréter les coefficients estimés en termes d'effets de débordement géographiques entre les régions. En particulier, nous montrons que différents processus spatiaux peuvent être mis en évidence à l'aide de ces spécifications : un processus de diffusion spatiale de choc aléatoire affectant les régions et/ou un processus de multiplicateur spatial sur la variable expliquée. Le point de départ est l'équation de β -convergence absolue, estimée par les Moindres Carrés Ordinaires, sur laquelle nous procédons à différents tests afin de détecter l'autocorrélation spatiale. Les spécifications les plus appropriées intégrant l'autocorrélation spatiale sont ensuite estimées en vérifiant leur pertinence à l'aide des différents tests présentés dans le deuxième chapitre. Le modèle final auquel nous aboutissons intègre un effet de diffusion spatiale : un choc aléatoire dans une région se diffuse à l'ensemble du système régional, l'amplitude de l'effet de choc diminuant avec la distance à la région émettrice. Tout d'abord, nous évaluons par simulation, dans le cadre de notre échantillon, les effets d'un choc sur la région Ile-de-France et estimons le processus de diffusion spatiale associé à ce choc. Ensuite, nous généralisons ces simulations à l'ensemble du système régional en proposant une méthodologie originale permettant d'étudier et d'évaluer globalement l'impact reçu et l'impact émis par une région lorsqu'un choc aléatoire est affecté à chaque région du système régional européen étudié. Nous mettons en évidence, toujours par simulation, qu'en matière de réception, les régions occupant une position géographique centrale sont avantagées, tandis que l'impact émis est en accord avec le

dynamisme économique et la localisation des régions. Ces résultats soulignent une forme de verrouillage géographique du schéma centre-périphérie.

Dans la troisième section, le phénomène de polarisation spatiale est intégré sous la forme d'un schéma Nord-Sud afin de vérifier si cette polarisation reflète l'existence de deux clubs de convergence au sein des régions européennes. Une méthode de constitution des clubs sur critère spatial est proposée. Ainsi, la composition des groupes Nord et Sud est déduite de l'analyse exploratoire des données spatiales, effectuée dans le deuxième chapitre, sur la série des PIB régionaux par tête en 1980, pour éviter le problème du biais de sélection d'échantillon soulevé par De Long (1988). Plus précisément, nous nous servons des indicateurs locaux d'autocorrélation spatiale les plus appropriés compte tenu des spécificités des séries lorsqu'elles sont évaluées en Ecus ou en SPA. Ainsi, pour les données en Ecus, nous utilisons les indicateurs locaux d'association spatiale d'Anselin (1995). Le groupe Nord comprend 90 régions, caractérisées par une association de type HH (« Haut-Haut »), tandis que le groupe Sud est constitué de 45 régions, caractérisées par une association de type BB (« Bas-Bas »). Seules 3 régions sur les 138 régions de l'échantillon ne peuvent pas être affectées à l'un ou l'autre groupe. En revanche, pour la série des PIB par tête évalués en SPA, la discrimination en deux groupes est meilleure lorsque l'on utilise les statistiques d'Ord et Getis (1995). Le groupe Nord est constitué des 93 régions associées à une valeur positive de la statistique et le groupe Sud des 40 régions associées à une valeur positive de la statistique. Les régimes spatiaux étant ainsi définis de manière appropriée suivant la variable étudiée, l'équation de β -convergence en présence d'hétérogénéité peut être estimée. Il apparaît qu'il n'est pas possible de rejeter l'hypothèse d'absence d'hétéroscédasticité en groupes. En revanche, il apparaît une instabilité structurelle des coefficients sous la forme de deux régimes spatiaux Nord/Sud. Nous soulignons, à travers nos études empiriques, les problèmes supplémentaires d'estimation et d'inférence que pose l'intégration simultanée de l'autocorrélation et de l'hétérogénéité spatiales. Les résultats obtenus indiquent l'existence de deux clubs de convergence spatiaux, les régions du groupe Sud convergent entre elles alors que l'hypothèse de convergence est systématiquement rejetée pour le groupe des régions Nord. Ces estimations supposent cependant que le coefficient d'autocorrélation spatiale est identique dans les deux régimes et entre les deux régimes. Pour répondre à cette limite, nous proposons enfin une spécification de l'équation de β -convergence en clubs permettant de tester l'existence d'effets propres à chaque club et d'un effet entre les clubs par l'estimation et le test de la significativité de coefficients d'autocorrélation spatiale intra- et inter-club. L'hypothèse d'égalité de ces coefficients ne peut cependant pas être rejetée.

Compte tenu des nombreux problèmes théoriques et empiriques auxquels est confrontée l'hypothèse de β -convergence, le cinquième et dernier chapitre a pour objectif de montrer comment les facteurs géographiques peuvent être combinés à la dimension temporelle en utilisant un concept alternatif de convergence. Le point de départ est l'analyse de la convergence proposée par Quah (1993a, 1993b, 1996a, 1996b). Dans cette approche, la convergence est mesurée à partir de l'évolution de la forme des distributions des PIB par tête

sur la période étudiée et à partir de l'évaluation de la mobilité des régions au sein de cette distribution.

Dans un premier temps, nous appliquons ces méthodes à notre échantillon sans intégrer la dimension spatiale. Pour cela, nous estimons, par des méthodes non-paramétriques, des fonctions de densité des PIB par tête relatifs à la moyenne européenne afin de détecter les principaux changements de forme de cette densité. Ensuite, les trajectoires individuelles des régions et la mobilité de ces dernières au sein de la distribution des PIB par tête sont étudiées en utilisant des matrices de transition associées à des chaînes de Markov. Ces dernières passent par la discrétisation de la distribution des PIB par tête en un certain nombre de classes de PIB et par l'estimation de probabilités de transition entre ces différentes classes. La mobilité des régions et les schémas de convergence obtenus à l'aide de cette méthode sont analysés par le calcul de divers indices de mobilité et le calcul de la distribution de long terme ou distribution ergodique. Nous montrons ainsi une très faible mobilité inter-classes, une forte persistance des régions à l'intérieur de chaque classe et un biais progressif de la distribution des PIB vers les régions relativement pauvres.

Dans un second temps, l'espace est explicitement intégré au sein de l'analyse précédente. Ce choix est d'abord justifié en montrant que les régions subissant une hausse ou une baisse dans la distribution des PIB sont fortement spatialement autocorrélées. Diverses méthodes sont alors appliquées permettant d'étudier comment les transitions des régions sont affectées par leur environnement géographique. L'une de ces solutions conduit à la perte de la dimension temporelle dans l'analyse alors qu'une autre méthode ne permet d'évaluer que la convergence intra-régionale, c'est-à-dire la façon dont les régions se comportent par rapport à leurs voisines. Nous appliquons finalement la méthode proposée par Rey (2001) permettant d'évaluer une forme de convergence interrégionale tout en captant la façon dont les trajectoires des régions dans la distribution des PIB par tête sont influencées par le comportement des régions voisines. Par ailleurs, nous étendons cette méthodologie en proposant un test de stationnarité spatiale des probabilités de transition. Les résultats obtenus permettent de souligner la dépendance à l'environnement géographique : la probabilité qu'une région améliore son classement dans la distribution des PIB par tête est d'autant plus faible qu'elle est entourée de régions pauvres, ce qui peut encore une fois être rapproché de l'idée du verrouillage géographique.

Cette thèse vise donc à apporter des éléments de compréhension sur les liens qu'entretiennent localisation géographique, disparités régionales et convergence des régions européennes. Tout d'abord, elle apporte une synthèse critique des fondements théoriques et économétriques de l'hypothèse de convergence. Ensuite, elle propose les méthodes appropriées permettant de renouveler le débat sur la convergence lorsque les disparités géographiques sont prises en compte. Enfin, elle présente des résultats empiriques sur les conséquences en termes de polarisation et de concentration géographiques que génère la prise en compte de ces disparités.

CHAPITRE 1

*Disparités géographiques et convergence :
un nouveau débat*

Chapitre 1

Disparités géographiques et convergence :

Un nouveau débat

L'objectif de ce chapitre est de présenter deux revues de littérature sur les modèles de croissance et d'économie géographique, afin de positionner le cadre théorique dans lequel nous nous inscrivons et de montrer comment la prise en compte des disparités géographiques renouvelle le débat sur la convergence des économies.

En première approche, on peut se contenter d'observer les disparités spatiales sans pour autant chercher à les expliquer. La localisation des pays ou des régions riches, de même que celle des pays ou des régions pauvres, ne présente pas d'intérêt en soi. Cependant, cette approche peut s'avérer insatisfaisante car elle conduit à négliger les liens entre le fonctionnement des systèmes économiques et les schémas de répartition spatiale des activités économiques : quels sont-ils et comment évoluent-ils ? Les théories de l'économie géographique peuvent être mobilisées dans ce but.

L'économie géographique a deux origines. D'un côté, elle provient d'une remise en cause des modèles néoclassiques de commerce international. Krugman (1991a, 1991b, 1993b) a ainsi proposé des modèles d'équilibre interrégionaux visant à expliquer les formations d'agglomérations sans faire appel aux avantages comparatifs. D'un autre côté, l'économie géographique est issue d'un dépassement de la Nouvelle Economie Urbaine et se rattache aux théories de la formation de l'espace urbain (Ogawa et Fujita, 1980, 1989 ; Fujita, 1989) ou aux modèles d'équilibres inter-urbains (Henderson, 1987 ; Abdel-Rahman, 1994, 1996). Sur un plan méthodologique, ce courant a pour objectif d'expliquer la formation et le développement des agglomérations, c'est-à-dire la localisation simultanée des agents au sein d'espaces régionaux ou urbains.

Pour notre argumentation, nous nous basons sur les modèles interrégionaux⁵. Trois résultats théoriques importants sont mis en évidence pour les économies contemporaines. Tout d'abord, elles sont caractérisées par une situation de concurrence monopolistique, par des systèmes de production à rendement croissants, par l'importance des innovations et des externalités de connaissance... Dans ces conditions, les calculs économiques des agents les

⁵ De nombreuses revues de la littérature de ces modèles sont actuellement disponibles : Duranton (1997) ; Ottaviano et Puga (1997, 1998) ; Fujita et Thisse (1997, 2000a, 2000b) ; Baumont et Huriot (1999) ; Ghio (1999a, 1999b) ; Schmutzler (1999), Neary (2001). On parle alors parfois de « Nouvelle Economie Géographique ».

incitent à se regrouper en un petit nombre de lieux et les disparités spatiales sont expliquées par le fonctionnement des économies. Ensuite, les politiques d'intégration régionale renforcent ces tendances. Enfin, il existe des interactions entre l'évolution des disparités économiques et les disparités spatiales comme l'illustre, par exemple, la persistance des regroupements géographiques de régions riches ou de régions pauvres.

Il existe un cadre théorique et empirique important consacré à l'étude des disparités économiques et de leurs évolutions. Pourtant, les théories de la croissance et les analyses de la convergence des économies intègrent rarement les disparités spatiales dans leur analyse. Ainsi, ce n'est que récemment, à la faveur du renouvellement des théories de la croissance au sein du courant de la croissance endogène, qu'un cadre théorique, intégrant les mécanismes de formation des agglomérations (c'est-à-dire de formation des disparités spatiales) et les processus de croissance, s'est construit et développé (Baumont, 1997).

Nous abordons ainsi l'analyse des disparités géographiques en suivant la logique suivante : tout d'abord, nous examinons les concepts de convergence associés aux théories de la croissance néoclassique. Ensuite, nous comparons l'idée de la convergence-rattrapage, qui en est issue, avec d'autres concepts de convergence. Enfin, nous montrons comment la prise en compte des disparités géographiques est de nature à renouveler le débat sur la convergence. Les différentes synthèses de la littérature présentées au cours de chapitre sont alors principalement effectuées dans le but d'enrichir ce débat.

Notre positionnement est essentiellement empirique et nous développons nos arguments sur la base des méthodes d'estimations habituellement employées pour évaluer l'évolution des disparités économiques. Nous discutons des limites de ces méthodes et souhaitons montrer que la prise en compte des disparités géographiques nécessite l'utilisation de méthodes appropriées.

Pour aborder ces différentes questions, nous précisons dans une première partie les différents concepts de convergence dérivés du modèle de croissance néoclassique (section 11). Le point de départ est le modèle de Solow et nous insistons sur l'idée de convergence déduite de ce modèle, afin de positionner clairement les remises en causes proposées par les théories de la croissance endogène. En particulier, nous souhaitons éclairer le débat qui oppose souvent trop rapidement l'association « modèle néoclassique de croissance et convergence » à l'association « modèles de croissance endogène et divergence ».

Dans une deuxième partie, nous proposons une synthèse des différents concepts de convergence (section 12). Les principaux enseignements qui peuvent être déduits de chaque concept de convergence sont ainsi mis en évidence. Nous soulignons également leurs limites et discutons des méthodes d'estimation de l'évolution des disparités qui leur sont associées. Nous sommes alors amenés à reconnaître « l'oubli », dans toutes ces approches, de la dimension spatiale des données. Il s'agit d'une limite supplémentaire pour deux raisons. D'une part, des problèmes économétriques d'estimation des équations de convergence se posent effectivement dans ce cas. D'autre part, des fondements théoriques existent pour souligner le rôle de l'espace dans l'analyse des phénomènes économiques.

La dernière partie vise alors à montrer comment le cadre théorique des modèles d'économie géographique et des modèles de la « synthèse géographie-croissance » peuvent soutenir notre argumentation (section 13). Les premiers permettent d'expliquer la tendance à la polarisation des activités et à la concentration de régions similaires. Les seconds permettent de montrer qu'il existe des interactions entre croissance et schémas géographiques. Ces modèles offrent, par ailleurs, un cadre propice à l'étude de l'impact de politiques d'intégration économique sur l'évolution des disparités régionales. Ceci est particulièrement intéressant dans le contexte de l'Union Européenne que nous avons choisi d'étudier dans les travaux empiriques de la thèse.

SECTION 11 : CONVERGENCE ET MODELES DE CROISSANCE

Le but de cette section est de présenter les différentes implications, en termes de convergence des économies, du modèle de croissance néoclassique et des modèles de croissance endogène.

Le modèle de croissance néoclassique, ou modèle de Solow-Swan, est basé sur des rendements d'échelle constants et une productivité marginale du capital décroissante. Il conduit à un résultat de convergence au sens où toutes les économies auront le même taux de croissance à long terme. En outre, si l'on suppose que les économies sont semblables en termes de technologie et de préférence, elles convergeront également vers un même niveau de produit par tête (paragraphe 111).

Ces prévisions, optimistes quant aux perspectives de développement des économies les plus pauvres, ont cependant été souvent invalidées. Les hypothèses et les conclusions du modèle de croissance néoclassique ont notamment été remises en cause par les théoriciens de la croissance endogène (Romer, 1986 ; Lucas, 1988 ; Aghion et Howitt, 1992). Ces modèles ont pour but d'expliquer la formation du taux de croissance de long terme (exogène dans le cadre néoclassique) et de montrer qu'il dépend des caractéristiques des économies. Les disparités de produit par tête auront le plus souvent tendance à persister voire à s'aggraver dans le long terme (paragraphe 112).

Alors que le modèle néoclassique prédit une convergence des économies en termes des taux de croissance, voire en termes des niveaux de production par tête pour des économies similaires, la plupart des modèles de croissance endogène concluent à une persistance des niveaux d'inégalités internationales et interrégionales. Le test de l'hypothèse de convergence constituerait alors un moyen relativement simple permettant de discriminer entre les deux familles de modèles de croissance. Nous montrons que cette conclusion est erronée (paragraphe 113).

111 La convergence dans le modèle néoclassique

Le concept de convergence, inscrit au cœur du modèle néoclassique de croissance de Solow-Swan, traduit l'idée qu'une économie converge dans le long terme vers un état de croissance dit régulier. Tout d'abord, nous présentons rapidement cette notion de base (paragraphe 111.1). Ensuite, nous abordons la façon dont les concepts de convergence absolue et de convergence conditionnelle se déduisent de ce modèle (paragraphe 111.2). Enfin, nous montrons comment le capital humain a été récemment intégré par Mankiw *et al.* (1992) dans le cadre du modèle de croissance néoclassique afin de mieux rendre compte des disparités internationales de croissance (paragraphe 111.3).

111.1 Les mécanismes de la croissance dans le modèle de Solow-Swan

Nous considérons une version simplifiée du modèle de Solow-Swan en économie fermée (Solow, 1956 ; Swan, 1956). Du côté de la demande, le comportement de consommation des ménages est de type keynésien (la propension à épargner est une part constante du revenu). Du côté de l'offre, le secteur des entreprises produit un bien unique composite pouvant alternativement servir à l'investissement ou à la consommation. Il y a concurrence parfaite sur le marché du bien comme sur celui des facteurs de production, qui sont au nombre de deux : le capital et le travail.

Formellement, supposons une fonction de production de type Cobb-Douglas à facteurs substituables et à rendements constants de la forme suivante :

$$Y_t = K_t^\alpha (A_t L_t)^{1-\alpha} \quad 0 < \alpha < 1 \quad [1.1]$$

où Y_t représente le niveau de production à la date t , K_t le stock de capital physique, L_t la quantité de facteur travail et A_t le niveau de la technologie. Le progrès technique est supposé neutre au sens de Harrod : il augmente l'efficacité du facteur travail. $A_t L_t$ est donc considéré comme la force de travail mesurée en unités effectives, incorporant à la fois le montant du travail ainsi que la productivité du travail déterminée par la technologie disponible.

Cette fonction de production néoclassique satisfait les conditions traditionnelles suivantes :

1. Les productivités marginales sont positives et décroissantes par rapport à chaque facteur de production. La fonction de production est donc concave.
2. La production se fait à rendements d'échelle constants. On peut donc réécrire la fonction de production sous forme intensive :

$$\hat{y}_t = \hat{k}_t^\alpha \quad [1.2]$$

où $\hat{y}_t = Y_t / A_t L_t$ et $\hat{k}_t = K_t / A_t L_t$ sont respectivement le produit et le capital par unités de travail effectif.

3. Les conditions d'Inada (1963) sont respectées : la productivité marginale du capital (respectivement du travail) tend vers l'infini lorsque le capital (respectivement le travail) tend vers 0, et tend vers 0 lorsque le capital (respectivement le travail) tend vers l'infini.

On suppose que le travail et le progrès technique croissent respectivement aux taux exogènes et constants n et g :

$$L_t = L_0 e^{nt} \quad [1.3]$$

$$A_t = A_0 e^{gt} \quad [1.4]$$

Le nombre d'unités de travail effectif $A_t L_t$ croît donc au taux $n + g$.

Le stock de capital est supposé croître avec l'investissement et décroître avec la dépréciation : $\dot{K}_t = I_t - \delta \cdot K_t$ où \dot{K}_t représente la dérivée de K_t par rapport au temps, δ est le taux de dépréciation du capital.

Soit s_k la fraction constante de la production qui est investie dans la formation de capital physique. Dans ces conditions, le stock de capital par unités de travail effectif évolue de la façon suivante :

$$\dot{\hat{k}}_t = s_k \hat{y}_t - (n + g + \delta) \hat{k}_t \quad [1.5a]$$

$$\text{soit : } \dot{\hat{k}}_t = s_k \hat{k}_t^\alpha - (n + g + \delta) \hat{k}_t \quad [1.5b]$$

où \hat{k} désigne la dérivée de \hat{k} par rapport au temps ⁶.

Les hypothèses faites sur la fonction de production (concavité, rendement marginal décroissant du capital) assurent généralement l'existence, la stabilité et l'unicité de la solution de long terme (état régulier) \hat{k}^* qui est définie pour $\dot{\hat{k}} = 0$ dans [1.5b].

Résoudre l'équation : $s_k \hat{k}^{*\alpha} = (n + g + \delta) \hat{k}^*$ permet de déterminer le niveau de capital par unités de travail effectif à l'état régulier :

$$\hat{k}^* = \left(\frac{s_k}{n + g + \delta} \right)^{\frac{1}{1-\alpha}} \quad [1.6]$$

$$\text{soit : } \hat{y}^* = \left(\frac{s_k}{n + g + \delta} \right)^{\frac{\alpha}{1-\alpha}} \quad [1.7]$$

⁶ En effet, $\dot{K}_t = I_t - \delta \cdot K_t \Leftrightarrow \frac{\dot{K}_t}{A_t L_t} = s \hat{y}_t - \delta \cdot \hat{k}_t$ et $\frac{\dot{K}_t}{A_t L_t} = \frac{\dot{K}_t}{K_t} \cdot \frac{K_t}{A_t L_t} = \left(\frac{\dot{\hat{k}}_t}{\hat{k}_t} + n + g \right) \hat{k}_t = \dot{\hat{k}}_t + (n + g) \hat{k}_t$.

La résolution du modèle permet donc de montrer qu'à l'état régulier le niveau du produit par unité de travail effectif atteint la valeur constante \hat{y}^* définie par :

$$\ln \hat{y}^* = \frac{\alpha}{1-\alpha} \ln s_k - \frac{\alpha}{1-\alpha} \ln(n + g + \delta) \quad [1.8]$$

On en déduit l'expression du niveau du produit par tête à l'état régulier :

$$\ln y^* = \ln A_0 + gt + \frac{\alpha}{1-\alpha} \ln s_k - \frac{\alpha}{1-\alpha} \ln(n + g + \delta) \quad [1.9]$$

Le niveau du produit par tête à l'état régulier dépend donc positivement du taux d'épargne et négativement du taux de croissance de la population.

Pour obtenir le taux de croissance de long terme de l'économie, il suffit de prendre la dérivée par rapport au temps de cette dernière expression :

$$\frac{d \ln y^*}{dt} = g \quad [1.10]$$

On obtient le résultat bien connu selon lequel le taux de croissance du produit par tête est égal au taux de progrès technique g . Le taux de croissance du produit est égal à $n + g$. Le taux de croissance est donc exogène et ne peut être modifié par des mesures de politique économique. Ainsi, le modèle de Solow conduit par hypothèse à une égalisation des taux de croissance entre les différentes économies à long terme puisque le progrès technique est supposé être un bien non rival accessible à tous.

En revanche, ce résultat montre qu'en l'absence de progrès technique ($g = 0$), le modèle néoclassique aboutit à un résultat d'étouffement de la croissance à long terme : le capital et le produit par tête sont constants dans le long terme. Ce résultat est dû à l'hypothèse centrale de la décroissance de la productivité marginale du capital.

En conclusion, dans le modèle néoclassique, la croissance a deux origines. D'une part, la croissance à l'état régulier est associée à la croissance de la productivité et à celle de la population. D'autre part, il y a une croissance transitoire associée au déplacement de l'économie à partir d'un stock de capital initial vers le stock de capital à l'état régulier. Nous nous intéressons maintenant à ce deuxième aspect et au processus de convergence qui en découle.

111.2 Le processus de convergence vers l'état régulier : présentation et tests

Le modèle néoclassique simple présenté dans le paragraphe précédent possède des implications optimistes du point de vue de la dynamique de la transition vers l'état régulier, c'est-à-dire de la convergence. En effet, les économies les plus pauvres ont une plus grande incitation à épargner et croissent donc plus vite pour une part donnée d'investissement. Ainsi, elles vont graduellement réduire la distance qui les sépare des économies les plus riches, dans lesquelles le stock de capital par tête est plus élevé. En conséquence, le rendement du capital y est plus faible, ce qui explique que l'accumulation du capital y soit plus lente que dans les économies pauvres. Dans ces conditions, on observe à la fois une *convergence en niveau* (toutes les économies atteignent le même niveau de produit et de capital par tête à l'état régulier) et une *convergence en taux* (à l'état régulier toutes les économies croissent au même rythme que celui du progrès technique). Les différences de taux de croissance observées aux niveaux international et régional ne sont ainsi que transitoires. C'est l'hypothèse de *convergence absolue* (paragraphe 111.21).

Cette conclusion n'est cependant valide que si toutes les économies sont caractérisées par les mêmes paramètres structurels. Dans le cas contraire, chaque économie est caractérisée par son propre état régulier et seule une convergence en taux est observée. C'est l'hypothèse de *convergence conditionnelle* (paragraphe 111.22).

Le test de l'hypothèse de convergence s'effectue formellement sur la base d'une approximation linéaire autour de l'état régulier, ce qui permet d'obtenir une estimation quantitative de la vitesse de convergence vers cet état régulier (paragraphe 111.23).

111.21 Convergence pour des économies identiques

Formellement, considérons à nouveau l'équation dynamique fondamentale [1.5b] du modèle de croissance néoclassique. Le taux de croissance de \hat{k}_t , le capital par unité de travail effectif, s'obtient en divisant [1.5b] par \hat{k} :

$$\frac{\dot{\hat{k}}_t}{\hat{k}_t} = g_{\hat{k}} = s_k \hat{k}_t^{\alpha-1} - (n + g + \delta) \quad [1.11]$$

L'équation [1.11] implique que le stock de capital d'une économie croît si $s_k \hat{k}_t^{\alpha-1} > (n + g + \delta)$ et diminue sinon.

Par exemple, nous représentons dans la figure 1.1 deux régions possédant la même fonction de production néoclassique ainsi que les mêmes caractéristiques structurelles n , δ et s . On suppose que le capital par unité de travail effectif initial de la première région est inférieur à celui de la deuxième région. Pourtant, puisque les deux économies possèdent les mêmes paramètres de production, elles vont converger vers le même point d'équilibre \hat{k}^* : à

gauche de l'état régulier, le taux de croissance de \hat{k} est positif (la courbe $s_k \hat{k}_t^{\alpha-1}$ est située au dessus de $n + g + \delta$) alors qu'à droite de l'état régulier, le taux de croissance de \hat{k} est négatif.

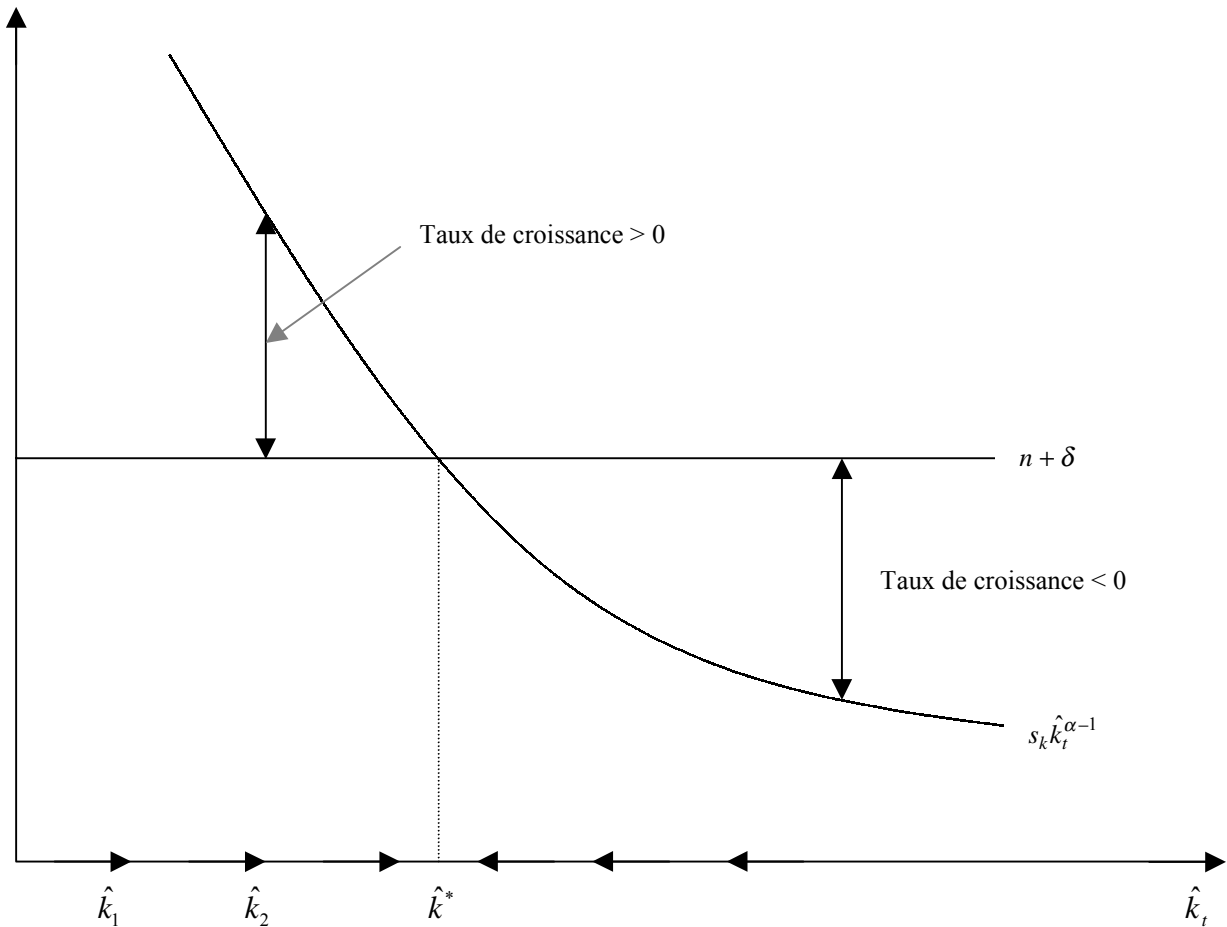


Figure 1.1 : Dynamique dans le modèle de Solow-Swan – économies identiques

De ce modèle, on déduit que la dotation en capital des économies pauvres croît plus vite que celle des économies riches. Les économies pauvres rattrapent les économies riches, c'est-à-dire qu'elles *convergent en termes de dotation de capital par tête*. En effet, de [1.11], il est aisé de montrer que la dérivée du taux de croissance de \hat{k} par rapport au temps est négative. Par conséquent, g_k décroît avec k . Ainsi, sur la figure 1.1 où $\hat{k}_1 < \hat{k}_2$, le taux de croissance de la région 1 est supérieur à celui de la région 2 mais les deux régions convergent vers \hat{k}^* . Cette propriété de convergence peut facilement être étendue à \hat{y} , ainsi qu'au capital et au produit par tête k et y .

Par conséquent, le modèle de croissance néoclassique implique la convergence en termes de capital et de production par tête. Si toutes les économies ont accès à la même technique (la fonction de production est la même pour toutes les économies, de même que le taux de dépréciation δ), alors toutes les économies, structurellement similaires en termes des

paramètres s , n et δ , et ne différant entre elles que par la quantité initiale de capital par tête, vont converger vers le même état régulier et vers un même taux de croissance, quelle que soit la valeur initiale du capital par tête. Durant la transition, les économies les plus pauvres connaîtront donc des taux de croissance de k et de y plus élevés que ceux des économies riches.

Cette hypothèse selon laquelle les économies pauvres connaissent une croissance par tête plus rapide que les économies riches, sans que cela dépende d'autres caractéristiques de ces économies, est appelée *convergence absolue*.

11.22 Convergence pour des économies différentes

Si les économies ne possèdent pas toutes les mêmes paramètres structurels, leurs états réguliers diffèrent également. Par exemple, la figure 1.2 représente deux économies qui diffèrent entre elles par deux aspects : elles ont des stocks initiaux de capital par unité de travail effectif différents $\hat{k}_1 < \hat{k}_2$ ainsi que des taux d'épargne différents ($s_{k1} < s_{k2}$). En conséquence, les quantités de capital par unités de travail à l'état régulier sont différentes ($\hat{k}_1^* < \hat{k}_2^*$).

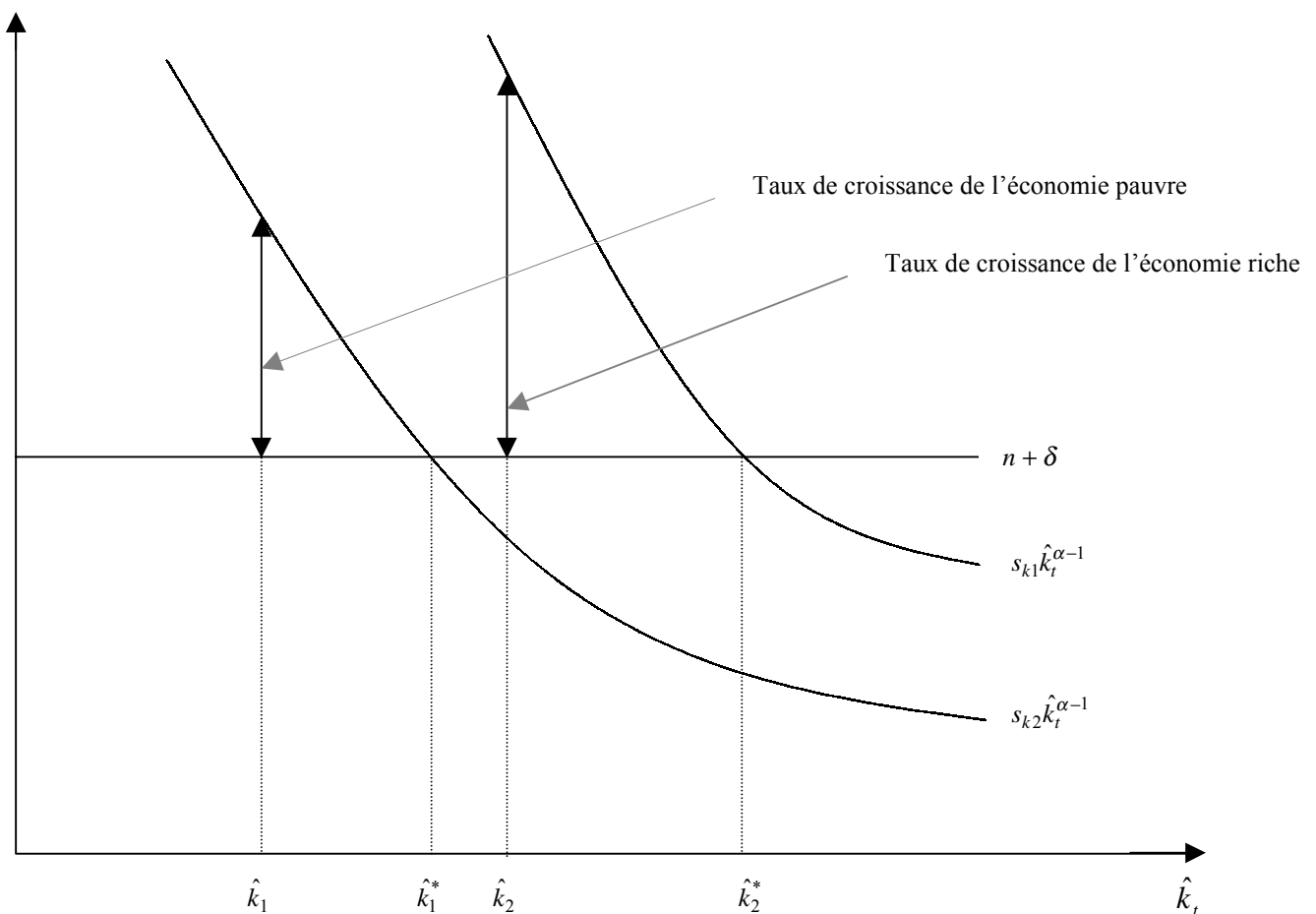


Figure 1.2 : Dynamique dans le modèle de Solow-Swan – économies différentes

Dans les conditions décrites par le graphique, la distance entre la courbe $s_k \hat{k}_t^{\alpha-1}$ et la ligne $n + g + \delta$ est plus grande pour l'économie 2 (économie riche). Le taux de croissance de l'économie 2 est donc plus élevé que celui de l'économie 1 (économie pauvre). Une économie pauvre peut donc croître moins vite qu'une économie riche. Chaque économie converge vers son propre état régulier et la vitesse de convergence vers cet état régulier est inversement reliée à la distance qui la sépare de cet état régulier. L'hypothèse selon laquelle les économies convergent vers différents états réguliers (mais à un même taux de croissance à l'état régulier) est appelée *convergence conditionnelle*.

11.23 Le test de l'hypothèse de convergence

Dans une étude empirique, on ne peut pas supposer que toutes les économies étudiées sont déjà à l'état régulier. L'estimation directe de l'équation [1.9] n'est donc pas possible. La démarche consiste alors à supposer que les économies ne se situent pas à leur état régulier, mais qu'elles en sont suffisamment proches pour pouvoir effectuer une approximation linéaire autour de cet état régulier.

Reprenons l'expression du taux de croissance du capital par unités de travail effectif (équation [1.11]) :

$$\frac{\dot{\hat{k}}_t}{\hat{k}_t} = s_k \hat{k}_t^{\alpha-1} - (n + g + \delta)$$

Lorsqu'on effectue un développement limité de Taylor à l'ordre 1 de cette expression autour de $\ln \hat{k}^*$, on obtient :

$$\frac{\dot{\hat{k}}_t}{\hat{k}_t} = (s_k (\alpha - 1) k^{*\alpha-1}) (\ln \hat{k}_t - \ln \hat{k}^*) \quad [1.12]$$

On remplace s_k par son expression en fonction de \hat{k}^* en utilisant [1.6]. Ceci permet d'obtenir le comportement de l'économie autour de sa trajectoire d'état régulier :

$$\frac{\dot{\hat{k}}_t}{\hat{k}_t} = \frac{d \ln \hat{k}_t}{dt} = -\gamma (\ln \hat{k}_t - \ln \hat{k}^*) \quad [1.13]$$

où $\gamma = (1-\alpha)(n+g+\delta)$ mesure la vitesse de convergence de l'économie vers sa trajectoire d'état régulier⁷. Par ailleurs, puisque $\hat{y}_t = \hat{k}_t^\alpha$, on obtient la même vitesse de convergence pour le produit par unité de travail effectif :

$$\frac{d \ln \hat{y}_t}{dt} = -\gamma(\ln \hat{y}_t - \ln \hat{y}^*) \quad [1.14]$$

La résolution de l'équation différentielle précédente nous permet d'obtenir l'évolution du produit par unités de travail effectif :

$$\ln \hat{y}_t = (1 - e^{-\gamma t}) \ln \hat{y}^* + e^{-\gamma t} \ln \hat{y}_0 \quad [1.15]$$

où \hat{y}_0 représente le niveau du produit par travailleur effectif à une date initiale donnée. En remplaçant \hat{y}^* par son expression [1.7], on obtient :

$$\ln \hat{y}_t - \ln \hat{y}_0 = \theta \frac{\alpha}{1-\alpha} \ln s_k - \theta \frac{\alpha}{1-\alpha} \ln(n+g+\delta) - \theta \ln \hat{y}_0 \quad [1.16]$$

avec $\theta = (1 - e^{-\gamma t})$.

On en déduit l'évolution du produit par tête y_t :

$$\ln y_t - \ln y_0 = \theta \ln A_0 + gt + \theta \frac{\alpha}{1-\alpha} \ln s_k - \theta \frac{\alpha}{1-\alpha} \ln(n+g+\delta) - \theta \ln y_0 \quad [1.17]$$

Sous sa forme estimable (Barro et Sala-I-Martin, 1991, 1995), cette équation est à la base du test de β -convergence où β est le coefficient de régression reliant le niveau du produit par tête initial au taux de croissance (cf. section 12, p. 36). Elle montre que le taux de croissance de l'économie dépend négativement du niveau du produit par tête initial.

A l'issue de cette présentation, nous pouvons résumer les principales caractéristiques du modèle néoclassique (Mankiw, 1995) :

1. Dans le long terme, une économie approche de son état régulier qui est indépendant des conditions initiales.
2. Le niveau de produit par tête à l'état régulier dépend du taux d'épargne et du taux de croissance de la population. Plus le taux d'épargne est important, plus le niveau du produit par tête à l'état régulier est important. A l'inverse, plus le taux de croissance de la population est important, plus le niveau du produit par tête à l'état régulier est faible.

⁷ En réalité, la vitesse de convergence n'est pas constante mais elle dépend de la distance à l'état régulier (Rappaport, 2000).

3. Le taux de croissance du produit par tête à l'état régulier dépend uniquement du taux de progrès technique, il ne dépend ni du taux d'épargne, ni du taux de croissance de la population.
4. A l'état régulier, le stock de capital croît au même taux que la production, ainsi le ratio capital sur production est constant.
5. A l'état régulier, la productivité marginale du capital est constante et la productivité marginale du travail croît au taux du progrès technique.

Comme nous le verrons dans la section 12, portant sur les principaux résultats empiriques obtenus à l'aide de cette méthodologie, la vitesse de convergence conditionnelle obtenue est environ égale à 2% par an, quels que soient les échantillons utilisés. D'après le modèle théorique, cette dernière est égale à $\gamma = (1 - \alpha)(n + g + \delta)$. Supposons que $(g + \delta)$ soit de 5% par an et que le taux de croissance de la population soit de 1%. Dans ces conditions, une vitesse de convergence de 2% implique que la part du capital dans le revenu national est égale à 2/3. Pourtant, la part du capital habituellement admise est de 1/3. Ce chiffre aurait pour conséquence d'aboutir à une vitesse de convergence de 4%. Le modèle semble donc prédire une vitesse de convergence plus rapide que celle qui est observée dans la plupart des études empiriques (Barro et Sala-I-Martin, 1995 ; Mankiw, 1995). Une première méthode permettant de rendre compatible ces deux éléments est de considérer que le capital incorpore non seulement le capital physique mais également le capital humain. Mankiw *et al.* (1992) proposent quant à eux d'incorporer directement le capital humain en plus du capital physique dans la fonction de production. L'analyse qu'ils proposent sert aujourd'hui de cadre de référence à la plupart des études empiriques sur la convergence.

111.3 Le processus de convergence en présence de capital humain

Mankiw *et al.* (1992) élargissent le modèle de Solow en prenant en compte la présence d'un nouveau facteur de production accumulable dans la fonction de production : le capital humain. La fonction de production s'écrit alors :

$$Y_t = K_t^{\alpha_1} H_t^{\alpha_2} (A_t L_t)^{1 - \alpha_1 - \alpha_2} \quad 0 < \alpha_1 + \alpha_2 < 1 \quad [1.18]$$

où H_t représente le stock de capital humain à la date t .

Sous les mêmes hypothèses que le modèle de Solow, le capital physique et le capital humain suivent les dynamiques d'accumulation suivantes :

$$\dot{\hat{k}}_t = s_k \hat{y}_t - (n + g + \delta) \hat{k}_t \quad [1.19]$$

et :
$$\dot{\hat{h}}_t = s_h \hat{y}_t - (n + g + \delta) \hat{h}_t \quad [1.20]$$

où $\hat{h}_t = H_t / A_t L_t$ représente le stock de capital humain par unité de travail effectif et s_h est la part du produit investie dans la formation de capital humain. Par simplicité, on suppose que les deux types de capitaux se déprécient au même taux δ .

A l'état régulier, le niveau des deux variables par unité de travail effectif est constant. On déduit alors des équations précédentes les valeurs du capital physique et du capital humain à l'état régulier :

$$\hat{k}^* = \left[\frac{s_k^{1-\alpha_2} s_h^{\alpha_2}}{n + g + \delta} \right]^{1/1-\alpha_1-\alpha_2} \quad [1.21]$$

$$\hat{h}^* = \left[\frac{s_k^{\alpha_1} s_h^{1-\alpha_1}}{n + g + \delta} \right]^{1/1-\alpha_1-\alpha_2} \quad [1.22]$$

On peut alors déterminer la valeur du produit par tête à l'état régulier en remplaçant ces expressions dans la fonction de production exprimée sous forme intensive : $\hat{y}_t = \hat{k}_t^{\alpha_1} \hat{h}_t^{\alpha_2}$:

$$\begin{aligned} \ln y^* = \ln(A_0) + gt - \left(\frac{\alpha_1 + \alpha_2}{1 - \alpha_1 - \alpha_2} \right) \ln(n + g + \delta) + \left(\frac{\alpha_2}{1 - \alpha_1 - \alpha_2} \right) \ln(s_k) \\ + \left(\frac{\alpha_2}{1 - \alpha_1 - \alpha_2} \right) \ln(s_h) \end{aligned} \quad [1.23]$$

En dérivant cette dernière expression par rapport au temps, on trouve à nouveau que le taux de croissance de l'économie est égal au taux de croissance du progrès technique.

En utilisant un raisonnement analogue à celui mené dans le cadre du modèle de Solow initial (approximation de Taylor d'ordre 1 autour de $\ln k^*$ et de $\ln h^*$), Mankiw *et al.* (1992) établissent que la vitesse de convergence dans le modèle de Solow augmenté est :

$$\gamma = (1 - \alpha_1 - \alpha_2)(n + g + \delta) \quad [1.24]$$

Le taux de croissance de l'économie est alors gouverné par l'équation suivante :

$$\begin{aligned} \ln \hat{y}_t - \ln \hat{y}_0 = \theta \frac{\alpha_1}{1 - \alpha_1 - \alpha_2} \ln s_k + \theta \frac{\alpha_2}{1 - \alpha_1 - \alpha_2} \ln s_h \\ - \theta \frac{\alpha_1 + \alpha_2}{1 - \alpha_1 - \alpha_2} \ln(n + g + \delta) - \theta \ln \hat{y}_0 \end{aligned} \quad [1.25]$$

où $\theta = (1 - e^{-n})$. En termes du produit par tête y_t , on obtient donc :

$$\begin{aligned} \ln y_t - \ln y_0 = & \theta \ln A_0 + gt + \theta \frac{\alpha_1}{1 - \alpha_1 - \alpha_2} \ln s_k + \theta \frac{\alpha_2}{1 - \alpha_1 - \alpha_2} \ln s_h \\ & - \theta \frac{\alpha_1 + \alpha_2}{1 - \alpha_1 - \alpha_2} \ln(n + g + \delta) - \theta \ln y_0 \end{aligned} \quad [1.26]$$

De cette équation, on déduit que le taux de croissance dépend des taux d'accumulation en capital physique et en capital humain ainsi que du niveau du produit initial. En conséquence, le mécanisme de rattrapage, c'est-à-dire l'effet favorable qu'un faible niveau de produit par tête exerce sur le taux de croissance, n'est pas l'unique déterminant du taux de croissance. Le niveau des taux d'investissement en capital physique ou en capital humain est susceptible de renforcer ou de contrecarrer le mécanisme de rattrapage.

112 La convergence dans les théories de la croissance endogène

Pendant longtemps, la théorie néoclassique de la croissance a constitué le cadre d'analyse dominant pour l'étude et la compréhension de la croissance économique et de la convergence des économies. Pourtant, la pertinence théorique et empirique de ce modèle a fait l'objet dans les années 80 de nombreuses critiques. Les premiers modèles de croissance endogène, apparus dans les années 80 (Romer, 1986 ; Lucas, 1988), ont pour objet de proposer une explication de la croissance à long terme et d'expliquer les disparités de croissance observées au niveau international (paragraphe 112.1). D'un point de vue technique, les conditions d'obtention d'une croissance endogène passent par une remise en cause des rendements marginaux décroissants du capital. Le modèle AK, dans lequel les rendements du capital sont constants, conduit à une croissance auto-entretenu et une absence de convergence des économies (paragraphe 112.2). L'absence de convergence absolue et/ou conditionnelle constitue en outre une implication des premiers modèles de croissance endogène (paragraphe 112.3).

112.1 Les origines de la croissance endogène

Les modèles de croissance endogène sont issus d'une double critique. Sur le plan théorique, le modèle néoclassique ne constitue pas une théorie de la croissance. Sur le plan empirique, la comparaison des rythmes de croissance des pays et des régions a conduit certains économistes à s'interroger sur la pertinence du modèle néoclassique.

Sur le plan théorique, la critique principale concerne les sources de la croissance. En effet, le taux de croissance à l'état régulier doit nécessairement provenir d'un élément extérieur

au modèle sous la forme d'un progrès technique fixé de façon exogène et gratuit. En effet, l'hypothèse de rendements constants implique l'épuisement du produit par la rémunération des facteurs travail et capital. Le comportement des agents, notamment leur comportement d'épargne, n'a aucun effet sur le taux de croissance à long terme (le taux d'épargne détermine seulement le niveau du capital et du produit par tête et non leur croissance). Le phénomène de la croissance reste donc pour une large part inexplicé. Le modèle néoclassique ne propose pas de théorie de la croissance de long terme mais explique le mécanisme d'accumulation du capital et la convergence vers l'état stationnaire. Ainsi, comme le souligne Mankiw (1995), le modèle de croissance néoclassique permet d'expliquer les différences de taux de croissance entre les économies comme résultats de la convergence vers les différents états réguliers, mais il ne permet pas d'expliquer la persistance de la croissance. Face à ce problème, l'apport principal des modèles de croissance endogène est d'explicitier la formation du taux de croissance de long terme. Il s'agit, en particulier, de montrer que ce dernier dépend des caractéristiques de l'économie, tel que le taux d'épargne.

Sur le plan empirique, l'optimisme du modèle néoclassique de la croissance quant aux perspectives de développement à long terme des économies les moins développées a été largement remis en cause, et son aptitude à rendre compte des écarts de croissance entre les économies mise en doute. Dans le modèle néoclassique de Solow-Swan, en effet, si les économies sont structurellement similaires (c'est-à-dire caractérisées par les mêmes valeurs du taux d'épargne et la même croissance de la population) et ont accès aux mêmes technologies, il doit y avoir à la fois convergence en termes de taux de croissance et convergence en termes des niveaux de revenu par tête. Ce point a suscité d'abondantes études empiriques dont les conclusions sont contrastées (cf. *infra*, section 12, p. 36). Cependant, d'une façon générale, l'hypothèse selon laquelle les économies pauvres, caractérisées par un stock de capital par tête moins élevé, rattrapent les économies riches, est le plus souvent remise en cause lorsque les échantillons sont composés de l'ensemble des pays du monde (Romer, 1986 ; Lucas, 1988 ; Barro, 1991 ; Mankiw *et al.*, 1992) ⁸.

Romer (1994) indique que ces deux insatisfactions sont à l'origine des nouvelles théories de la croissance et ont influencé le cours initial de ces nouvelles théories. Ainsi, l'objectif principal des modèles de croissance endogène est de lever les hypothèses d'un progrès technique exogène et de rendements décroissants dans le but d'explicitier l'origine de la croissance à long terme et de pouvoir mieux rendre compte de la diversité des taux de croissance des économies ⁹.

⁸ Cette seconde critique n'est cependant pas fondée dès lors que l'on considère le concept de convergence conditionnelle, comme nous le verrons dans le paragraphe 113.2.

⁹ Par exemple, on peut se référer à Grossman et Helpman (1991), Lordon (1991), Amable et Guellec (1992), Barro et Sala-I-Martin (1995), Aghion et Howitt (1998), Amable (1999).

112.2 Le modèle AK

D'un point de vue technique, les conditions d'obtention d'une croissance endogène passent par des rendements non décroissants sur les facteurs de production accumulables. Ainsi, dans le modèle AK (Rebelo, 1991), les rendements décroissants sont remplacés par des rendements constants, l'élasticité de la production par rapport au seul facteur accumulable est égale à 1 et sa productivité marginale est constante et égale à A (elle ne s'annule donc pas comme dans le modèle de Solow sans progrès technique exogène).

Formellement, le modèle s'écrit de la façon suivante :

$$Y_t = AK_t \quad [1.27]$$

où Y_t et K_t représentent respectivement le niveau de production et le stock de capital à la date t , A est une constante reflétant le niveau de technologie. On pose $A > 0$. Il s'agit d'un modèle sans progrès technique car le but est de montrer que la croissance par tête de long terme peut se produire même en l'absence d'un progrès technique exogène.

Si les autres hypothèses du modèle néoclassique sont maintenues (taux d'épargne exogène et constant), il est aisé de montrer que l'évolution du capital et du produit par tête est donnée par l'expression suivante :

$$\frac{\dot{k}_t}{k_t} = \frac{\dot{y}_t}{y_t} = sA - (n + \delta) \quad [1.28]$$

De cette équation, on déduit que le taux de croissance du capital par tête et le taux de croissance du produit par tête sont constants et non nuls. Ils dépendent négativement des taux de croissance de la population et de dépréciation, et positivement des caractéristiques de la technologie et du taux d'épargne. La production croît ainsi indéfiniment, même sans l'hypothèse d'un progrès technique exogène.

Une représentation graphique du modèle est fournie dans la figure 1.3. La courbe $\delta + n$ est une ligne horizontale. Contrairement au modèle de croissance néoclassique, la courbe d'épargne n'est plus une courbe décroissante mais une ligne : sA . Dans la figure 1.3, on suppose que $sA > \delta + n$. Ceci implique une distance positive et constante entre les deux lignes et donc des taux de croissance du capital et du produit par tête positifs et constants.

Considérons maintenant deux économies qui possèdent les mêmes caractéristiques structurelles et qui diffèrent simplement de par leur stock du capital initial. Le modèle prédit que le taux de croissance des deux économies est le même quel que soit le stock de produit initial. Ce modèle ne prévoit donc ni convergence absolue, ni convergence conditionnelle.

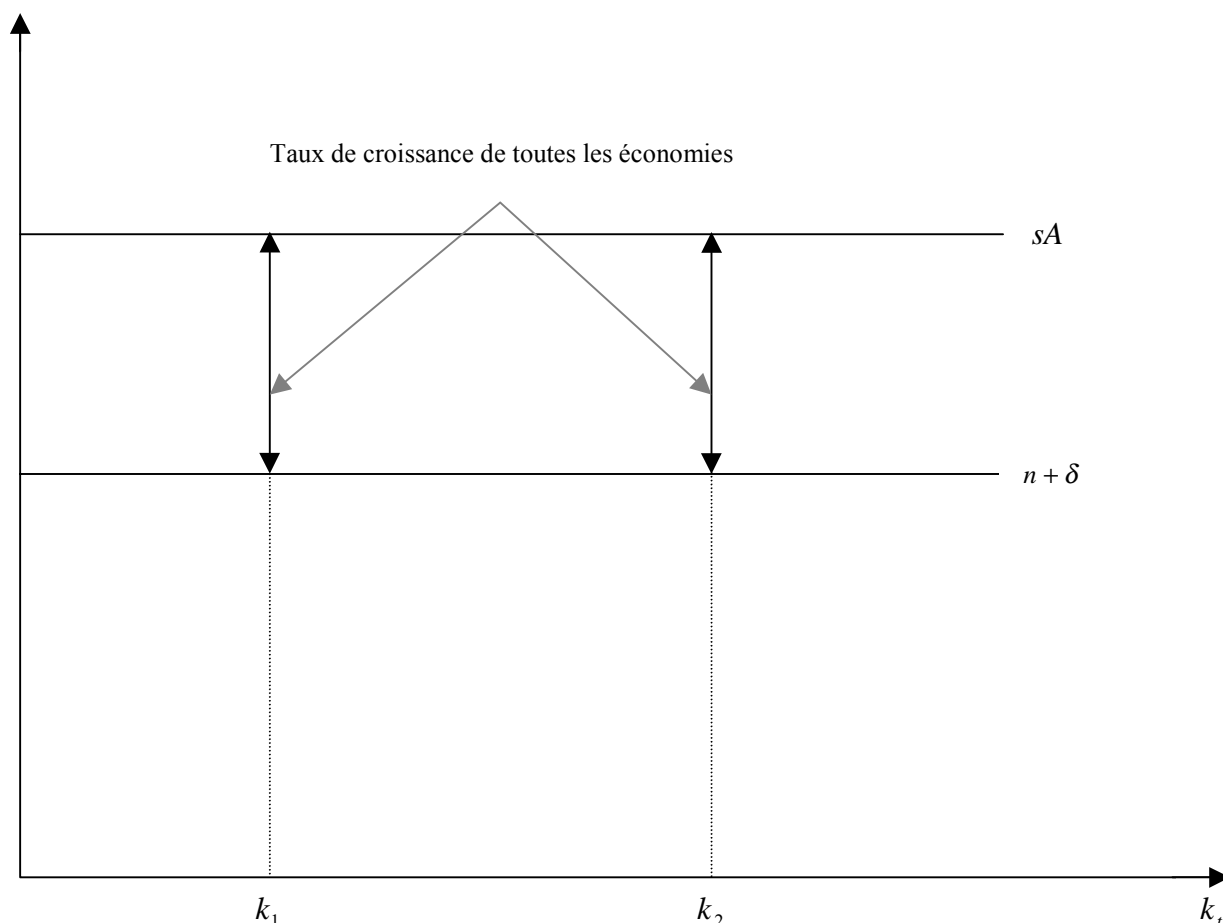


Figure 1.3 : Le modèle AK

La modification des hypothèses du modèle de Solow portant sur l'hypothèse des rendements décroissants du facteur accumulable a ainsi deux conséquences. D'une part, une économie peut connaître des taux de croissance par tête à long terme non nuls sans aucun progrès technique et d'autre part, le taux de croissance du produit par tête dépend aussi des paramètres de comportement du modèle, tel que le taux d'épargne et le taux de croissance de la population. Il peut ainsi être influencé par des mesures de politique économique, notamment par des mesures de politique fiscale.

Ce modèle de croissance endogène peut être considéré comme un cas limite du modèle néoclassique (Mankiw, 1995). Le modèle AK est en effet un modèle de Solow avec fonction Cobb-Douglas dans laquelle la part du capital tend vers 1. Dans ce cas, la vitesse de convergence vers l'état régulier dans le modèle néoclassique, déterminée dans le paragraphe 112.2, tend vers 0. Cette interprétation est rendue possible dès lors que l'on considère que la variable K dans cette fonction de production se réfère à un concept de capital élargi : le travail est assimilé au capital humain et ensuite agrégé au capital physique.

112.3 Les premiers modèles de croissance endogène

L'obtention et la justification de rendements non décroissants rendent nécessaire l'intégration d'externalités, qui constituent alors le moteur de la croissance à long terme, ou le remplacement de la concurrence parfaite par une concurrence imparfaite (concurrence monopolistique). Ces modifications entraînent que l'équilibre décentralisé est le plus souvent sous-optimal, rendant nécessaire des interventions publiques pour modifier les taux de croissance.

Les auteurs faisant appel au premier type de mécanisme ajoutent dans la fonction de production un facteur supplémentaire dont l'accumulation produit une externalité positive. Les rendements globaux sont donc croissants, alors qu'ils peuvent être non croissants pour chaque agent. Ces externalités peuvent ainsi porter sur le capital physique (Romer, 1986) ou humain (Lucas, 1988). Dans le second cas, la concurrence parfaite est remplacée par la concurrence monopolistique et les rendements croissants s'exercent sur chacun des biens appartenant à un ensemble de grande taille (Romer, 1990 ; Aghion et Howitt, 1992). Dans tous les cas, le progrès technique n'est plus exogène. Au contraire, il est devenu endogène, soit car il est induit par la croissance ou l'investissement, soit car il est le résultat de dépenses volontaires d'agents (par exemple, par le biais de dépenses de R&D).

Dans le but d'expliquer les sources de la croissance, les nouvelles théories cherchent à mobiliser d'autres facteurs de production que les facteurs traditionnels, capital et travail. Les modèles proposés sont d'une très grande variété et mobilisent un grand nombre de facteurs. Amable et Guellec (1992) proposent de classer les modèles de croissance endogène d'après quatre sources ou facteurs de croissance : l'accumulation de capital physique (Romer, 1986), l'accumulation de capital humain (Lucas, 1988), la recherche-développement (Romer, 1990 et Aghion et Howitt, 1992) et les infrastructures publiques (Barro, 1990).

Dans le modèle de Romer (1986), l'accumulation du capital (non nécessairement le capital physique) engendre deux types d'externalités technologiques positives. La première est liée à l'accumulation des connaissances et au processus d'apprentissage par la pratique (*learning by doing*) : l'expérience acquise en produisant permet à chaque travailleur d'augmenter à la fois le niveau de ses connaissances mais aussi celui de la collectivité. L'externalité provient du fait que chaque entreprise peut alors bénéficier du stock de connaissances accumulées par tous. Le second type d'externalité technologique est lié à l'existence de complémentarités entre firmes et à la diffusion de la connaissance. Différents cas de figure apparaissent selon les valeurs des paramètres de la fonction de production. L'un d'entre eux est l'obtention d'un taux de croissance par tête constant auquel cas il n'existe pas de période de transition vers un état régulier, comme dans le modèle AK.

Dans le modèle de Barro (1990), les différentes dépenses de l'état (en éducation, en infrastructures, etc.), engendrent des rendements globaux croissants qui soutiennent la croissance économique. Lucas (1988), Grossman et Helpman (1991) et Romer (1990), montrent que des taux de croissance positifs peuvent être maintenus indéfiniment dans des modèles dans lesquels le taux de progrès technique est déterminé d'une façon endogène par

des décisions privées à investir dans le capital humain ou technologique, si ces activités ne sont pas soumises à des rendements décroissants.

Compte tenu de la diversité des dynamiques d'accumulation proposée par les théories de la croissance endogène, de la multiplicité des paramètres et des variables impliquées, ces modèles sont le plus souvent caractérisés par des équilibres de long terme multiples, y compris des configurations de croissance explosive peu réalistes. Ces différents équilibres de long terme dépendent des conditions initiales de l'économie qui, au contraire du modèle néoclassique, ont un rôle déterminant sur sa trajectoire de long terme.

Les premiers modèles de croissance endogène conduisent ainsi à remettre en cause l'hypothèse de convergence : les économies qui démarrent avec des niveaux peu élevés de stock de capital par tête ne connaissent pas nécessairement des taux de croissance plus élevés que ceux des économies riches. Il n'existe pas de phénomène de convergence et des taux de croissance inégaux peuvent persister à long terme, ce qui implique une augmentation des niveaux d'inégalité internationale et interrégionale. Certains auteurs ont pu alors considérer que la vérification empirique de l'hypothèse de convergence est un moyen permettant de discriminer entre les deux familles de modèles de croissance.

113 Convergence et discrimination entre les modèles de croissance

La principale différence entre les deux familles de modèles semble porter sur le signe de la corrélation partielle entre le taux de croissance et le niveau initial de produit par tête. Cette corrélation devrait être négative dans le cadre du modèle néoclassique alors que le signe attendu devrait être positif dans la plupart des modèles de croissance endogène. Dans ces conditions, le test de l'hypothèse de convergence serait un moyen relativement simple permettant de départager les deux familles de modèles.

Cette conclusion est problématique de deux points de vue. D'une part, le modèle néoclassique de croissance n'implique pas la convergence absolue mais plutôt la convergence conditionnelle, compatible avec de fortes disparités des niveaux de revenu par tête dans le long terme (paragraphe 113.1). D'autre part, certains modèles de croissance endogène intègrent également une dynamique transitoire associée à un phénomène de convergence conditionnelle (paragraphe 113.2).

113.1 Convergence absolue *versus* convergence conditionnelle

Au milieu des années 80, le rejet de l'hypothèse de convergence absolue pour des échantillons contenant un large ensemble de pays du monde a été interprété par certains théoriciens de la croissance endogène comme une preuve permettant de rejeter les modèles

néoclassiques en faveur des nouvelles théories (Romer, 1986 ; Lucas, 1988). A la suite de ces critiques, différents auteurs, tels que Barro et Sala-I-Martin (1991, 1992) et Mankiw *et al.* (1992), ont rappelé que le modèle de croissance exogène ne prédit pas une β -convergence absolue, mais plutôt une β -convergence conditionnelle. En effet, comme nous l'avons déjà remarqué dans les paragraphes précédents, la convergence absolue, ou encore le rattrapage des économies riches par les économies pauvres, dans le modèle de Solow, repose sur une hypothèse cruciale : *la seule différence entre économies est le stock de capital initial*. Cependant, si les économies ont des caractéristiques différentes (en termes de taux d'épargne, de niveau de technologie par exemple), alors elles convergent vers des états réguliers différents. Le rejet de l'hypothèse de convergence absolue ne peut donc pas être à la base d'un rejet du modèle néoclassique (Kocherlakota et Yi, 1995 ; Mankiw, 1995 ; Pack, 1997).

En outre, le concept de convergence conditionnelle est compatible avec l'observation d'une persistance des inégalités entre pays ou entre régions si les états d'équilibres de long terme des économies sont très différents. Par conséquent, comme le soulignent Islam (1998a) et De la Fuente (1997), il y a équivalence observationnelle entre la théorie néoclassique et la théorie de la croissance endogène dès lors que l'on considère le concept de convergence conditionnelle. De ceci, on ne peut déduire ni l'un ni l'autre.

113.2 Croissance endogène et convergence

Bien que les théories de la croissance endogène aient en commun le rejet d'un taux de croissance exogène à l'état régulier, elles sont en revanche extrêmement diverses pour ce qui est de la source de la croissance et peuvent avoir différentes implications en termes de convergence. Certains modèles de croissance endogène intègrent ainsi une dynamique transitoire et un mécanisme de convergence conditionnelle.

Soulignons tout d'abord qu'il n'est pas nécessaire que la technologie soit non convexe pour qu'il existe une croissance dans le long terme. En effet, le résultat d'annulation de la croissance dans le long terme, en l'absence de progrès technique, n'est pas dû à la présence de rendements décroissants mais plutôt à la violation de la deuxième condition d'Inada :

$$\lim_{k \rightarrow \infty} f'(k) = 0 \quad [1.29]$$

En conséquence, il est possible d'obtenir une croissance auto-entretenu dans le long terme, même en présence de rendements décroissants, à la condition de modifier la technologie afin que la productivité marginale du capital ne décroisse pas indéfiniment mais soit bornée inférieurement de telle sorte que $\lim_{k \rightarrow \infty} f(k) = A > 0$.

Jones et Manuelli (1990) montrent qu'une technologie possible qui ne vérifie pas la dernière condition d'Inada est de la forme :

$$Y = F(K, L) = AK + \Omega(K, L) \quad [1.30]$$

où $\Omega(K, L)$ est une fonction de production néoclassique qui vérifie toutes les conditions d'Inada.

La partie AK de la fonction de production permet à la croissance d'être auto-entretenu dans le long terme tandis que la partie $\Omega(K, L)$ permet au modèle d'avoir une dynamique transitoire et donc d'exhiber un phénomène de convergence.

Croissance endogène et convergence conditionnelle ne sont donc pas incompatibles. Plusieurs modèles de croissance endogène ont été développés combinant ces deux aspects.

Dans le modèle de Tamura (1991), les économies possèdent des niveaux initiaux de capital humain différents. L'externalité liée au capital humain bénéficie à toutes les économies de telle sorte que celles qui possèdent un stock inférieur à la moyenne bénéficient d'un retour sur investissement supérieur à la moyenne. Le capital humain étant le seul facteur de production, il y a convergence complète par simple retour à la moyenne.

Mulligan et Sala-I-Martin (1993) proposent un modèle de croissance endogène dans la lignée d'Uzawa (1965) et de Lucas (1988) dans lequel il existe deux secteurs : les firmes produisant les biens et les services, les universités produisant le facteur de production connaissance utilisé gratuitement par les deux secteurs. La fonction de production est à rendements d'échelle constants sur les facteurs accumulables de telle sorte que le modèle puisse générer une croissance perpétuelle sans progrès technique. Les auteurs montrent que ce modèle possède une dynamique transitionnelle. Dans le modèle néoclassique, il y a convergence car le niveau de capital par tête est différent du niveau de l'état régulier. Dans les modèles bi-sectoriels de type Uzawa-Lucas, il y a convergence conditionnelle si le stock de capital est faible relativement au stock de connaissances. La production de l'économie croît alors plus rapidement.

Plus récemment, divers modèles de croissance endogène avec diffusion des technologies et convergence conditionnelle ont été proposés (Barro et Sala-I-Martin, 1997 ; Basu et Weil, 1998 ; Howitt, 2000). Par exemple, dans le modèle de Barro et Sala-I-Martin (1997), le taux de croissance mondial de long terme est déterminé par les découvertes technologiques effectuées par les économies leader. En revanche, les économies en retard sur le plan technologique sont capables d'imiter les produits inventés par les économies leader. Puisque les coûts d'imitation sont inférieurs aux coûts d'innovation, les économies en retard croissent relativement rapidement et auront alors tendance à rattraper les économies leader. Dans le modèle de Basu et Weil (1998) en revanche, il n'y a pas de coûts d'adoption des technologies mais une économie en retard ne peut utiliser la technologie de l'économie leader que si elle a atteint un niveau suffisant de développement.

Il paraît difficile dans ces conditions de discriminer les modèles de croissance endogène et de croissance exogène sur la base du seul test de convergence. D'une part, le rejet de la convergence absolue ne peut pas être à la base du rejet du modèle néoclassique. D'autre part, dans la mesure où certains modèles de croissance endogène possèdent les mêmes implications testables du point de vue de la convergence conditionnelle que le modèle néoclassique, l'acceptation de la convergence conditionnelle ne doit pas signifier le rejet de tous les modèles de croissance endogène (McCallum, 1996).

Il en résulte que le test entre croissance néoclassique et croissance endogène doit être formulé sur d'autres bases que celles de la convergence et que des tests « directs » des théories de la croissance endogène doivent être effectués (comme par exemple : Jones, 1995a, 1995b ; Kocherlakota et Yi, 1996, 1997 ; Evans, 1996, 1998).

Par exemple, Jones (1995b) propose un test des modèles de type AK sur la base du caractère transitoire ou permanent des changements de politique économique sur le taux de croissance. En effet, dans les modèles de croissance endogène, un accroissement d'une variable de politique économique entraîne un accroissement permanent sur le taux de croissance alors que l'effet n'est que transitoire dans les modèles de croissance exogène. Dans ces derniers, les variations de politique économique n'entraînent que des effets de niveau. A partir d'un échantillon de 15 pays de l'OCDE sur longue période et de l'étude des propriétés statistiques des séries, Jones montre que les modèles de type AK semblent rejetés : un accroissement permanent du taux d'investissement n'affecte la croissance que de manière transitoire durant une période de 8 à 10 ans seulement. Jones (1995b) propose également un test des modèles, basés sur la R&D, aboutissant à un effet d'échelle : la taille de l'économie est un déterminant de la croissance de long terme (Romer, 1990 ; Grossman et Helpman, 1991 ; Aghion et Howitt, 1992). Ces modèles sont également rejetés car la croissance du nombre de chercheurs engagés dans des activités de R&D ne semble pas affecter le taux de croissance. Les modèles récents de croissance endogène se sont d'ailleurs attachés à éliminer cet effet d'échelle (Kortum, 1997 ; Segerstrom, 1998 ; Young, 1998 ; Jones, 1999).

Conclusion de section

Préciser le concept de convergence au cœur du modèle de croissance néoclassique fait apparaître de nombreux débats sur l'hypothèse de convergence qui doivent être clarifiés. En effet, le processus de convergence étudié dans le modèle de Solow-Swan est celui qui conduit une économie vers un état de croissance de long terme. La notion de convergence-rattrapage, qui en est déduite, a de ce fait largement orienté les tests de l'hypothèse de convergence. Cependant, les notions de β -convergence absolue et conditionnelle doivent être distinguées car elles ne fournissent pas la même information quant au processus de rattrapage. Par ailleurs, nous avons montré que l'idée, selon laquelle le test de l'hypothèse de convergence est un moyen simple permettant de discriminer entre les modèles néoclassiques et les modèles de croissance endogène, est problématique. La notion de convergence-rattrapage déduite du modèle néoclassique de croissance peut enfin être discutée et comparée à d'autres concepts de convergence que nous présentons dans la section suivante.

SECTION 12 :

ANALYSE EMPIRIQUE DE LA CONVERGENCE

Le débat sur la convergence a donné lieu à de très nombreuses études empiriques aux résultats souvent contradictoires, comme le soulignent les synthèses récentes réalisées par Klenow et Rodríguez-Clare (1997a), Islam (1998a) et Durlauf et Quah (1999). Ces résultats contradictoires s'expliquent, en partie, parce que différentes conceptions de convergence ont été testées et que différentes approches méthodologiques et procédures de tests ont été employées (coupe transversale, données de panel, séries temporelles, etc.). Dès lors, la signification et les implications économiques de la convergence ont donné lieu à de nombreuses controverses. Le but de cette section est de clarifier les différents concepts de convergence et leurs interprétations afin de mieux situer le cadre dans lequel nous nous inscrivons et le champ des comparaisons auquel nous pourrions être confrontés.

Les premiers développements (paragraphe 121) concernent l'idée de convergence-rattrapage qui est associée au concept de β -convergence. Il existe cependant trois approches permettant de tester l'hypothèse de β -convergence : le test de la β -convergence absolue, de la β -convergence conditionnelle ou de la convergence en clubs. Ces trois approches ne sont pas substituables. En effet, elles s'appuient sur des fondements théoriques différents et le rejet de l'une ne peut appuyer la confirmation de l'autre. Ainsi, on rappelle que la β -convergence absolue et la β -convergence conditionnelle dérivent du modèle de croissance néoclassique selon que les états réguliers des économies sont similaires ou non. En revanche, l'hypothèse de β -convergence en clubs est un concept plus récent qui dérive à la fois des modèles de croissance endogène à équilibres multiples et du modèle de croissance néoclassique avec hétérogénéité des agents.

Sur la base des nombreuses critiques soulevées par la mise en œuvre de l'hypothèse de β -convergence, différents auteurs ont proposé des méthodes et des concepts alternatifs (paragraphe 122). En particulier, les méthodes portant sur l'étude de la distribution des séries de PIB par tête ont donné lieu aux approches basées sur l'analyse de l'écart-type de la coupe transversale (concept de σ -convergence) et sur l'étude des chaînes de Markov. Par ailleurs, les méthodes fondées sur les propriétés des séries temporelles ont donné lieu au concept de convergence stochastique.

121 La convergence-rattrapage ou β -convergence

La première définition de convergence est associée à la notion de rattrapage, déjà énoncée par Baumol (1986), et qui est déduite du modèle néoclassique de croissance de Solow-Swan. Plusieurs notions de β -convergence ou de « convergence-rattrapage » sont

possibles, selon que les paramètres, caractérisant l'état régulier des économies, sont supposés identiques ou différents. Le terme « rattrapage » traduit alors l'idée que l'économie converge vers son propre état régulier.

Reprenons en effet l'équation [1.7] qui détermine le niveau de produit par unités de travail effectif à l'état régulier dans le modèle de Solow. De cette équation, on déduit le niveau du produit par tête à l'état régulier :

$$y^* = A_0 e^{gt} \left(\frac{s_k}{n + g + \delta} \right)^{\frac{\alpha}{1-\alpha}} \quad [1.31]$$

Le niveau de produit par tête d'une économie à l'état régulier dépend donc des six éléments suivants : $\theta = (A_0, g, s_k, n, \delta, \alpha)$. En outre, les conditions initiales des économies sont variables.

Dans ces conditions, selon que tout ou partie des éléments contenus dans le vecteur θ sont similaires ou non, on aboutit à l'un des trois concepts de β -convergence : absolue, conditionnelle ou convergence en clubs (Islam, 1998a).

- Si tous les éléments de θ sont similaires entre les économies alors il y a β -convergence absolue (paragraphe 121.1).
- Si certains des éléments de θ diffèrent entre les économies, il y a β -convergence conditionnelle. Nous verrons que ce cas de figure se traduit par l'ajout de variables explicatives dans les équations de convergence afin de contrôler les différences à l'état régulier (paragraphe 121.2).
- Dans l'approche en termes de clubs de convergence, l'accent est davantage mis sur les conditions initiales. Les économies semblables convergent à long terme les unes vers les autres si leurs conditions initiales appartiennent au « bassin d'attraction » d'un même équilibre stationnaire (paragraphe 121.3).

121.1 La β -convergence absolue et le « rattrapage » des économies riches par les économies pauvres

Si toutes les économies sont structurellement identiques et ont accès à la même technologie, alors elles convergent toutes vers le même état régulier et possèdent le même taux de croissance à l'état régulier.

Lorsque les données sont disponibles en coupe transversale pour deux périodes, la période initiale 0 et la période finale T , Barro et Sala-I-Martin (1991, 1995) démontrent que, dans ce cas, on estime l'équation suivante par les Moindres Carrés Ordinaires (MCO) :

$$\frac{1}{T} \log \left(\frac{y_{iT}}{y_{i0}} \right) = a + \beta \log(y_{i0}) + u_i \quad u \sim i.i.d(0, \sigma_u^2) \quad [1.32]$$

où $i = 1, \dots, N$, N est le nombre d'économies dans l'échantillon, y_{it} est le produit par tête de l'économie i (mesuré par exemple par le PIB par tête) à la période t , $t = 0$ ou T , $(1/T) \cdot \log(y_{iT}/y_{i0})$ est le taux de croissance moyen du produit par tête entre la date initiale et la date T , a et β sont les paramètres inconnus à estimer¹⁰.

On dit qu'il y a β -convergence absolue ou inconditionnelle lorsque β est négatif et statistiquement significatif. En effet, dans ce cas, le taux de croissance moyen du PIB par tête entre les dates 0 et T est négativement corrélé avec le niveau initial du PIB par tête : les économies pauvres tendent à croître plus vite que les économies riches. L'estimation de β permet en outre de calculer la vitesse de convergence :

$$\gamma = -\ln(1+T\beta)/T \quad [1.33]$$

La durée nécessaire pour que les économies comblent la moitié de l'écart qui les sépare de leur état stationnaire, appelée la demi-vie, est définie de la façon suivante :

$$\tau = -\ln(2)/\ln(1+\beta) \quad [1.34]$$

L'étude de l'hypothèse d'une corrélation négative entre le taux de croissance du produit par tête et le niveau initial remonte à Abramovitz (1986), Baumol (1986) ou De Long (1988). Par exemple, Baumol (1986) utilise les données fournies par Maddison (1982) pour régresser le taux de croissance de la productivité du travail de 1870 à 1979 sur le logarithme de la productivité du travail en 1870 à partir d'un échantillon de 16 pays industrialisés. Il obtient un coefficient négatif, interprété comme une preuve de β -convergence absolue. La vitesse de convergence estimée est relativement faible mais le pouvoir explicatif de la régression est de plus de 80%.

Cette étude a fait l'objet de deux critiques formulées par De Long (1988). La première porte sur l'existence d'un biais de sélection ex-post sur l'échantillon. En effet, les pays choisis par Baumol sont uniquement des pays ayant atteint un certain niveau de développement en 1979. Les pays qui n'ont pas convergé sont donc exclus *a priori* de l'analyse. De Long (1988) ajoute 7 pays à l'échantillon de Baumol sur la base d'un critère de sélection ex-ante. Il montre que l'on accepte encore l'hypothèse de β -convergence mais que le coefficient associé au PIB par tête est moins élevé en valeur absolue.

La seconde critique concerne les erreurs de mesure du PIB par tête en 1870. En effet, les données ont été construites à partir de sources différentes. Ces erreurs de mesure provoquent une corrélation négative artificielle entre le taux de croissance par tête et le PIB par tête initial et biaisent les estimations (De Long, 1988). Cette critique souligne que d'une

¹⁰ On trouve parfois aussi la régression avec le taux de croissance cumulé comme variable expliquée : $\log(y_{iT}/y_{i0}) = a + \beta \log(y_{i0}) + u_i$, $i = 1, \dots, N$.

façon générale, le test de la convergence est conditionné par l'obtention de données de bonne qualité. Temple (1999a) mentionne par exemple que les *Penn World Tables* (Summers et Heston, 1988, 1991), souvent utilisées pour les études internationales, posent des problèmes de qualité de données pour les pays en développement et des problèmes d'interprétation pour les calculs des taux de croissance.

L'hypothèse de β -convergence absolue a été testée sur d'autres échantillons de pays et de régions sur différentes périodes. Il apparaît que cette hypothèse est largement rejetée lorsque l'échantillon contient un large ensemble de pays du monde (Baumol et Wolff, 1988 ; De Long, 1988 ; Barro, 1991) alors que dans la plupart des cas, elle ne peut être rejetée pour des échantillons plus restreints tels que les pays de l'OCDE (Mankiw *et al.*, 1992) ou les régions européennes (Armstrong, 1995a, 1995b ; Barro et Sala-I-Martin, 1995). Cette constatation est à la base de la recherche de l'existence de « clubs de convergence » (cf. paragraphe 121.3, p. 50) : bien que l'hypothèse de β -convergence soit rejetée pour un vaste échantillon, il existerait cependant des groupes de pays au sein desquels un processus de convergence serait à l'œuvre. La notion de β -convergence conditionnelle est en revanche différente.

121.2 La β -convergence conditionnelle ou l'individualisation de la notion de rattrapage

Le test de l'hypothèse de β -convergence conditionnelle s'effectue en rajoutant à l'équation de convergence [1.32] des variables permettant de maintenir constant l'état régulier des économies (paragraphe 121.21). Cependant, cette procédure a fait l'objet de nombreuses critiques (paragraphe 121.22) et plusieurs auteurs se sont récemment tournés vers l'estimation des équations de β -convergence conditionnelle à l'aide des méthodes de données de panel (paragraphe 121.23).

121.21 Le test de l'hypothèse de β -convergence conditionnelle

Pour tester l'hypothèse de β -convergence conditionnelle associée au modèle de Solow-Swan, il faut maintenir constants les déterminants de l'état stationnaire propres à chaque économie. Ceci peut s'effectuer en rajoutant dans l'équation [1.32] des variables permettant de contrôler l'hétérogénéité des trajectoires à long terme :

$$\frac{1}{T} \log \left(\frac{y_{iT}}{y_{i0}} \right) = a + \beta \log(y_{i0}) + \gamma X_i + u_i \quad u \sim i.i.d(0, \sigma_u^2) \quad [1.35]$$

où X_i est un vecteur de variables permettant de maintenir constant l'état stationnaire de l'économie i . Ce modèle s'estime habituellement par les MCO. Si le coefficient β est négatif et statistiquement significatif, on accepte l'hypothèse de β -convergence conditionnelle.

Les variables additionnelles se divisent en deux groupes : d'une part, les variables d'état introduites en accord avec la version élargie du modèle de Solow et d'autre part, d'autres variables de contrôle et d'environnement.

Les niveaux initiaux des variables d'état, comme le capital physique, le capital humain et le taux de croissance de la population, sont introduits dans le cadre proposé par Mankiw *et al.* (1992). Cet article a donné lieu à un « renouveau » du modèle de croissance néoclassique car il apporte en amont, c'est-à-dire dans la conceptualisation de la fonction de production, une justification à la prise en compte de variables supplémentaires dans la description et dans l'explication des processus de croissance (Klenow et Rodriguez-Clare, 1997b). Différents auteurs ont ainsi repris le cadre empirique proposé par Mankiw *et al.* (1992) pour estimer les équations de convergence. Il faut cependant souligner que la prise en compte explicite du capital humain pose deux problèmes.

Premièrement, inclure le capital humain pose de nombreux problèmes de mesure (Mankiw, 1995). Mankiw *et al.* (1992) utilisent le taux de scolarisation dans l'enseignement secondaire comme mesure du capital humain et ignorent celui de l'enseignement primaire. Dans ces conditions, les auteurs parviennent à expliquer 80% des différences internationales du revenu par tête. Klenow et Rodriguez-Clare (1997b) estiment que cette mesure introduit une variabilité excessive dans la mesure du capital humain et montrent que le modèle de Solow augmenté n'explique plus que 50% des différences internationales de revenu par tête dès lors que l'on tient compte des autres niveaux d'enseignement. D'autres critiques portent sur la prise en compte de l'enseignement supérieur, sur la définition du système de formation plutôt que d'éducation (en incluant notamment la formation professionnelle), sur les mesures en taux ou en niveaux, sur les entrées ou les sorties des différents cycles de formation... Bien souvent cependant, la disponibilité des données et les problèmes de comparabilité au niveau international l'emportent.

Deuxièmement, dans le modèle de Solow augmenté, le capital humain est un facteur de production supplémentaire dans la fonction de production. Par conséquent, le taux de croissance du capital humain doit être lié à celui de l'économie. Pourtant, dans une étude basée sur 78 pays sur la période 1965-1985, Benhabib et Spiegel (1994) montrent que le coefficient associé au taux de croissance du stock de capital humain est négatif mais non significatif, alors que le coefficient associé au niveau moyen du capital humain est positif et significatif. Il faut cependant noter que Temple (1999b) estime que ces derniers résultats sont une illusion statistique et sont dus à la présence de valeurs aberrantes dans l'échantillon. Il est donc nécessaire d'exclure tous les pays pour lesquels l'équation de croissance semble inadaptée. En utilisant comme critère de sélection les résidus de l'équation de convergence, il obtient une corrélation significative entre le taux de croissance de l'économie et la croissance

du taux de scolarisation. Cette méthode conduit toutefois Temple à exclure 14 pays de l'échantillon.

Plusieurs auteurs ont cherché à étendre encore la méthodologie proposée par Mankiw *et al.* (1992) en incluant diverses variables telles que le rôle du gouvernement, à travers le capital physique public et les transferts (Bajo-Rubio, 2000), ou le chômage (Bräuning et Panneberg, 2002). Knowles et Owen (1995) distinguent capital humain tiré de l'éducation et capital humain tiré de la santé (dans l'étude de Mankiw *et al.* (1992), seul le premier type de capital humain est considéré). Le premier est défini comme le nombre d'années passées à s'éduquer par la population âgée de 25 ans et plus. Le second est défini comme la proportion du revenu associé aux dépenses de santé. McDonald et Roberts (2002) effectuent la même distinction mais utilisent deux variables proxy pour le capital humain lié à la santé : la mortalité infantile et l'espérance de vie.

Nonneman et Vanhoudt (1996) utilisent une fonction de production avec trois facteurs de production : le capital physique, le capital humain et la technologie mesurée par les dépenses de R&D. Les auteurs parviennent alors à améliorer le pouvoir prédictif du modèle de Mankiw *et al.* (1992), qui n'était que de 30% pour les pays de l'OCDE, en aboutissant à un coefficient de détermination supérieur à 70%.

Les études empiriques incluent souvent de nombreuses variables de contrôle ou d'environnement, de nature économique ou non, telles que le ratio de la consommation publique au PIB, le ratio de l'investissement domestique au PIB, les modifications des termes de l'échange, le taux de fécondité, le degré d'instabilité politique, etc. (Barro et Sala-I-Martin, 1995 ; Barro, 1997). Le plus souvent, ces variables ne sont pas liées formellement à un modèle de croissance mais sont choisies d'une façon *ad hoc*. Les régressions « informelles » de ce type ont parfois été appelées « régressions de Barro », d'après l'article de Barro (1991).

Les variables utilisées dans la littérature sont très nombreuses et variées. Ainsi, Durlauf et Quah (1999) recensent 87 exemples de variables de contrôle alors que Sala-I-Martin (1997) étudie 63 variables. Les effets attendus de ces variables de contrôle sur le taux de croissance correspondent à leurs influences sur la position de l'état régulier.

L'analyse de l'hypothèse de la β -convergence conditionnelle a fait l'objet de nombreuses études empiriques. Elle a été testée sur différents échantillons de pays du monde (Barro et Sala-I-Martin, 1995), de régions européennes (Armstrong, 1995a, 1995b) ou de régions d'un même pays¹¹. Sala-I-Martin (1994, 1996b) observe que, quel que soit l'échantillon utilisé, pays ou régions, les études en coupe transversale aboutissent très souvent à une vitesse de convergence stable de l'ordre de 2%. Le processus de convergence semble être extrêmement faible : une vitesse de convergence de 2% par an implique en effet qu'il faut environ 35 ans pour qu'un pays réduise de moitié son écart par rapport à sa propre trajectoire

¹¹ Les exemples sont très nombreux. Parmi les plus récents on trouve : Mas *et al.* (1995) pour l'Espagne, Cárdenas et Pontón (1995) pour la Colombie, Chen et Fleisher (1996) pour la Chine, Persson (1997) pour la Suède, Siriopoulos et Asteriou (1998) et Petrakos et Saratsis (2000) pour la Grèce, Azzoni (2001) pour le Brésil...

de long terme. Le modèle néoclassique de base n'est compatible avec cette vitesse de convergence que si la part de la rémunération du capital dans le PIB est très supérieure à 30%. Dans le modèle de Mankiw *et al.* (1992) en revanche, une vitesse de convergence de 2% implique une part du capital de 40% environ, un chiffre qui est compatible avec les estimations issues des comptes nationaux.

121.22 Critiques de l'approche en termes de β -convergence conditionnelle

Bien que très souvent employée dans la littérature, l'approche en termes de β -convergence conditionnelle a fait l'objet de nombreuses critiques (Le Pen, 1997 ; Durlauf et Quah, 1999 ; Temple, 1999a).

1) Les problèmes d'endogénéité des variables explicatives

La première critique porte sur l'endogénéité des variables explicatives dans les équations de β -convergence conditionnelle : les variables explicatives ne sont pas exogènes, mais elles sont déterminées conjointement avec le taux de croissance. Ce problème a deux conséquences. D'une part, sur le plan économétrique, une modélisation erronée des relations de causalité a pour conséquence de biaiser les estimations. D'autre part, sur le plan économique, la question est alors de savoir comment interpréter les résultats des estimations.

Par exemple, considérons une équation de β -convergence dans laquelle le coefficient estimé pour le capital humain est significatif et positif. Ce résultat peut vouloir dire que les économies possédant de forts taux de scolarisation sont devenues plus riches, toutes choses égales par ailleurs. Cependant, on pourrait tout aussi bien argumenter qu'un taux de scolarisation élevé est le résultat d'un taux de croissance élevé ou encore qu'une troisième variable est la cause du taux de croissance et du taux de scolarisation (Cho, 1996). Les mêmes interrogations sont valables pour les autres variables habituellement incluses dans les équations de convergence.

En particulier, les liens de causalité entre investissement en équipement et taux de croissance ont souvent été étudiés. Ainsi, De Long et Summers (1991) ou Levine et Revelt (1992) concluent à une forte influence positive de l'investissement sur le taux de croissance. Blomström *et al.* (1996) reviennent sur ce résultat à partir d'une étude portant sur 101 pays et la période 1965-1985. Leurs tests de causalité suggèrent que la croissance induit un investissement en capital physique. La causalité s'exercerait donc dans le sens inverse que celui habituellement utilisé dans les équations de β -convergence conditionnelle. Cette conclusion se retrouve dans l'étude de Cho (1996) qui montre, à partir des données de Summers-Heston (1988) pour 109 pays sur la période 1960-1985, que le taux d'investissement et le taux de croissance de la population ne sont pas exogènes. En conséquence, l'endogénéité de ces variables induit un biais négatif sur le coefficient associé au PIB par tête initial. Ce biais est suffisamment important pour que le résultat de

convergence conditionnelle ne soit pas valable lorsque des méthodes de régressions alternatives, prenant en compte le biais de simultanéité, sont utilisées.

Face à ce problème, la solution économétrique consiste à trouver des variables exogènes comme instruments. Ces derniers doivent remplir deux conditions, d'une part, ils ne doivent pas être corrélés avec les erreurs et d'autre part, ils doivent être corrélés avec les variables explicatives. Ainsi, l'utilisation de variables instrumentales à la place du taux d'investissement a pour effet de diminuer le coefficient estimé lié à cette variable (Barro et Sala-I-Martin, 1995 ; Caselli *et al.*, 1996 ; Barro, 1997).

L'utilisation de variables instrumentales pose cependant plusieurs problèmes. Tout d'abord, les variables instrumentales disponibles sont rares (Mankiw, 1995 ; Temple, 1999a). Comme la croissance peut être expliquée par de nombreuses variables, il est difficile d'identifier des instruments qui sont fortement corrélés avec les variables endogènes et qui peuvent être légitimement éliminés de la régression. Ensuite, l'utilisation de variables retardées en données de panel comme instrument n'assure pas l'exogénéité de ces variables. En effet, une variable n'est pas nécessairement exogène lorsqu'elle est prédéterminée : les effets du capital humain sont parfois observables après de longs délais. Enfin, les auteurs motivent le plus souvent le choix des instruments sur la base de leur exogénéité, en utilisant par exemple les niveaux initiaux des variables. Cependant, les instruments choisis sont le plus souvent corrélés avec les erreurs. C'est le cas lorsqu'il y a des variables omises qui affectent à la fois la croissance et les instruments choisis pour les variables endogènes. Par exemple, Frankel et Romer (1996) étudient la relation entre croissance et commerce. L'ouverture à l'échange étant une variable endogène, ils choisissent comme instrument une variable géographique, l'aire des pays. Utiliser cette variable présuppose alors que les corrélations entre la taille et tous les déterminants de la croissance sont négligeables (Durlauf et Quah, 1999 ; Brock et Durlauf, 2001). Pourtant, il est probable que la taille des pays est corrélée avec des facteurs de croissance omis dans la régression, tels que le régime politique, les dépenses militaires, etc.

2) *Les problèmes de multicolinéarité*

L'introduction de nombreuses variables de contrôle dans les équations de convergence induit souvent des problèmes de multicolinéarité, ce qui ne permet pas une analyse précise de leurs effets individuels. La multicolinéarité a alors pour effet de fausser la significativité statistique des différents régresseurs. Comme le soulignent une fois encore Durlauf et Quah (1999), le modèle de Solow-Swan admet une telle variété d'extensions, comme l'inégalité, le régime politique ou l'ouverture à l'échange, que les différentes variables sont souvent fortement corrélées entre elles, sans qu'aucune d'entre elles ne constitue des explications prioritaires ou mutuellement exclusives d'explication de la croissance. Par conséquent, il convient d'être prudent quant à l'interprétation de la significativité statistique d'un sous-ensemble arbitrairement choisi de variables de contrôle.

3) Les problèmes de degrés de liberté dans les études internationales

Le nombre de variables explicatives pose également un problème de degrés de liberté (Mankiw, 1995 ; Brock et Durlauf, 2001). En effet, de nombreuses variables de contrôle ont été proposées dans la littérature alors que dans les études internationales basées sur les *Penn World Tables*, seuls 120 pays sont disponibles. En revanche, lorsque seul un sous-ensemble de variables est inclus dans les équations de convergence, les résultats sont dépendants des variables incluses et des variables omises. Dans ces conditions, il n'est pas possible d'établir les conditions associées à une croissance rapide et de se baser sur une inférence statistique fiable pour déterminer les diverses explications de la croissance.

Ce problème est relié à celui de la robustesse des variables explicatives.

4) La robustesse des variables explicatives

Cette critique est liée au manque de robustesse des estimations de β -convergence conditionnelle. En effet, face à la très grande variété des variables explicatives utilisées dans les régressions de convergence, Levine et Revelt (1992) et Levine et Zervos (1993) montrent que la plupart des variables utilisées ne sont plus significatives lorsqu'on ajoute des variables de conditionnement supplémentaires.

La méthode employée s'inspire de l'analyse des bornes extrêmes développée par Leamer (1978, 1983, 1985) et Leamer et Leonard (1983).

Soit la régression :

$$Y = \beta_x X + \beta_m M + \beta_z Z + \varepsilon \quad [1.36]$$

où Y représente le taux de croissance entre deux périodes, X est un ensemble de variables présentes dans toutes les régressions, M est la variable d'intérêt et Z est un sous-ensemble de variables choisies parmi un ensemble de variables potentiellement importantes pour la croissance. Les variables X sont les suivantes : part de l'investissement dans le PIB, niveau initial du PIB, niveau initial du taux de scolarisation dans le secondaire, taux de croissance annuel moyen de la population. En outre, Levine et Revelt choisissent de restreindre le nombre de variables dans Z à trois pour limiter les problèmes de multicollinéarité qui risquent de fausser l'analyse.

Le but de l'analyse est de faire varier le sous-ensemble des variables Z pour trouver l'intervalle dans lequel se trouve l'estimateur β_m . Plus précisément, on estime l'équation [1.36] pour toutes les combinaisons possibles de variables Z . Pour chaque modèle i estimé, on obtient un estimateur $\beta_{m,i}$ ainsi que son écart-type $\sigma_{m,i}$. La borne extrême inférieure est définie par la valeur minimale ($\beta_{m,i} - 2\sigma_{m,i}$) et la borne supérieure est définie par la valeur maximale de ($\beta_{m,i} + 2\sigma_{m,i}$). Si le coefficient β_m est significatif et est du même signe que les

deux bornes extrêmes, la variable M est dite robuste. Si le coefficient β_m n'est pas significatif ou que les deux bornes extrêmes ne sont pas du même signe, on dira que la variable M est fragile.

Les données employées sont extraites de Summers-Heston (1988) et couvrent 119 pays. Selon les données disponibles, les estimations sont effectuées sur la période 1970-1989 ou 1974-1989. A l'aide de cette méthodologie, Levine et Revelt (1992) montrent que presque toutes les variables d'intérêt dans les équations de convergence sont fragiles, sauf l'investissement et le PIB initial. Cette dernière variable reste négativement corrélée avec le taux de croissance si le capital humain figure également parmi les variables explicatives. La croissance de la population, les variables proxy du capital humain, la plupart des variables fiscales et monétaires ainsi que les différents indicateurs d'ouverture et d'instabilité politique ne passent pas les tests de robustesse définis par ces auteurs.

Sala-I-Martin (1997) estime que ce test est trop restrictif car il y a peu de chances que le domaine de définition de $\beta_{m,i}$ soit entièrement positif ou entièrement négatif. Il est toujours possible de trouver un ensemble de régresseurs qui rend une variable non significative dans une équation de convergence et qui la fera ainsi apparaître comme fragile. En conséquence, il propose une autre approche moins contraignante dans le but d'examiner quelles sont les variables qui sont effectivement corrélées avec le taux de croissance. Au lieu de qualifier simplement une variable robuste ou fragile en observant un changement de signe, il construit un intervalle de confiance pour celle-ci. Dans ce but, l'auteur étudie l'ensemble de la distribution de l'estimateur de $\beta_{m,i}$ et appelle robustes les variables qui sont statistiquement significatives dans 95% des régressions contenant des variables de contrôle.

L'auteur comptabilise dans la littérature empirique 62 variables qui ont été employées dans des régressions à la Barro et étudie la significativité de 59 d'entre elles. Il obtient alors des résultats opposés à l'analyse précédente dans le sens où il trouve que 21 variables peuvent être considérées comme significativement corrélées avec le taux de croissance. Ces variables se classent dans 9 catégories : 1/ Variables régionales (variables muettes pour l'Afrique subsaharienne et l'Amérique Latine qui ont un effet négatif alors que la latitude a un effet positif), 2/ Variables politiques (mesures des libertés civiles, de l'instabilité politique...), 3/ Variables religieuses, 4/ Variables mesurant les distorsions de marché, 5/ Variables mesurant les types d'investissement (investissement en équipement ou non), 6/ Variables mesurant la production du secteur primaire, 7/ Le degré d'ouverture, 8/ Le type d'organisation économique et 9/ Les anciennes colonies espagnoles.

Andrés *et al.* (1996) évaluent la robustesse des estimations des équations de β -convergence conditionnelle dans le cas des pays de l'OCDE selon une méthode différente de celle de Levine et Revelt. Ils étudient en particulier l'effet d'indicateurs macroéconomiques et montrent qu'ils ont un impact significatif et robuste sur la croissance économique : les exportations et la croissance monétaire sont positivement corrélées avec la croissance alors

que l'inflation, la variabilité de l'inflation et la variabilité de la croissance monétaire sont négativement corrélées avec le taux de croissance. En outre, les estimations des coefficients, associés aux taux d'accumulation en capital physique et en capital humain, sont peu robustes à des changements de périodes d'estimation ou d'échantillon.

5) *Le problème des observations extrêmes*

Le problème des observations extrêmes a principalement été soulevé par Temple (1998, 1999a, 1999b). Les observations extrêmes sont des observations qui sont particulièrement peu « représentatives », à cause de l'hétérogénéité des paramètres, de variables omises ou d'erreurs de mesure. Par exemple, De Long et Summers (1991) mettent en évidence l'effet important que les pays du Botswana et de la Zambie ont sur les estimations des coefficients et leurs précisions dans leurs régressions.

Temple (1998, 1999b) suggère l'utilisation des méthodes de régression robustes proposées par Rousseeuw et Leroy (1987) afin d'identifier de telles observations extrêmes. Une fois que ces dernières sont identifiées, elles sont exclues de l'échantillon et l'équation de convergence est à nouveau estimée par les MCO. Dans l'article de 1998, Temple reprend les estimations effectuées par Mankiw *et al.* (1992) sur les pays de l'OCDE. Il teste également l'extension proposée par Nonneman et Vanhoudt (1996) qui consiste à ajouter la technologie pour améliorer les performances empiriques du modèle de Solow augmenté sur les pays de l'OCDE. Les résultats obtenus montrent que le coefficient associé à la part de la R&D n'est pas significatif dès lors que le Japon, une observation extrême dans ce cas de figure, est éliminé de l'échantillon. En outre, le modèle de Mankiw *et al.* (1992) perd une grande partie de son pouvoir explicatif, dès lors qu'il est estimé par les méthodes de régressions robustes.

6) *Corrélation entre le PIB initial et le terme d'erreur*

Une autre critique porte sur une possible corrélation entre le PIB par habitant initial et le terme d'erreur de la régression : ceci invalide l'application des MCO dans l'estimation du modèle [1.32] ou [1.35] et l'inférence statistique fondée sur cette estimation. Evans et Karras (1993, 1996) et Evans (1996) montrent ainsi que l'estimation des modèles de β -convergence conditionnelle, basée sur l'utilisation de coupes transversales, n'est valide que sous des hypothèses contraignantes et improbables. Plus précisément, les écarts de PIB par tête par rapport à la moyenne en coupe doivent suivre des processus autorégressifs de premier ordre identiques. En outre, les différences permanentes entre économies doivent être totalement intégrées dans les régressions de β -convergence conditionnelle. Dans le cas contraire, si les économies possèdent des structures dynamiques autorégressives différentes, le résidu de l'équation de β -convergence conditionnelle est corrélé avec le PIB par tête initial. En conséquence, l'estimation des paramètres de cette équation est biaisée (Evans, 1996).

7) *Biais des variables omises*

Une dernière critique tient au fait qu'il est peu probable de réussir à trouver l'ensemble des variables incluses dans le vecteur X de l'équation [1.35] permettant de contrôler les différences d'états réguliers entre économies. Le phénomène de la croissance étant complexe et multidimensionnel, il est vraisemblable que le terme résiduel dans les équations de β -convergence contienne un certain nombre de variables omises corrélées avec les régresseurs présents, auquel cas les coefficients estimés sont biaisés. Par ailleurs, la solution qui consiste à augmenter le nombre de variables explicatives se heurte directement aux problèmes d'endogénéité et de multicollinéarité, mentionnés précédemment.

Ainsi, les régressions en coupes transversales ne seraient pas à même de traiter correctement les effets spécifiques ou effets fixes, par pays ou par régions, puisqu'elles contraignent les paramètres, comme le niveau de technologie initiale, à être identiques pour toutes les économies. Pour répondre à ces critiques et échapper aux nombreuses limites des procédures de test précédentes, une nouvelle voie de recherche s'est récemment orientée vers l'utilisation des méthodes de l'économétrie des données de panel.

121.23 *Estimation par les données de panel*

Différentes études s'orientent vers l'estimation des équations de convergence à l'aide des techniques de données de panel (Islam, 1995 ; Caselli *et al.*, 1996 ; De la Fuente, 1996 ; Lee *et al.*, 1997 ; Jean-Pierre 1999 ; Goddard et Wilson, 2001).

Pour illustrer les problèmes des régressions en coupe transversale et montrer comment une solution peut être apportée par les techniques de données de panel, reprenons la spécification [1.26] du modèle de Mankiw *et al.* (1992), qui exprime le taux de croissance du produit par tête en fonction du produit par tête initial et des autres paramètres du modèle :

$$\ln y_t - \ln y_0 = \theta \ln A_0 + gt + \theta \frac{\alpha_1}{1 - \alpha_1 - \alpha_2} \ln s_k + \theta \frac{\alpha_2}{1 - \alpha_1 - \alpha_2} \ln s_h - \theta \frac{\alpha_1 + \alpha_2}{1 - \alpha_1 - \alpha_2} \ln(n + g + \delta) - \theta \ln y_0$$

avec les mêmes notations que précédemment : y_t est le produit par tête à la période t , y_0 est le produit par tête à la date initiale, $\theta = (1 - e^{-\gamma})$, γ est la vitesse de convergence, A_0 est le niveau de technologie initial, g est le taux de croissance du progrès technique, s_k est la fraction de la production investie dans la formation de capital physique, s_h est la fraction de la production investie dans la formation de capital humain, n est le taux de croissance de la population, δ est le taux de dépréciation.

D'après Mankiw *et al.* (1992), le terme A_0 ne reflète pas simplement la technologie mais également les dotations en ressources, le climat, les institutions, etc. Ils admettent alors que ce terme peut varier entre les pays. Pourtant, pour aboutir à l'équation estimable, ils font l'hypothèse suivante :

$$\ln A_0 = a + \varepsilon \quad [1.37]$$

où a est une constante et ε est un choc spécifique à chaque pays. Cette hypothèse pose deux problèmes. D'une part, A_0 est supposé identique pour tous les pays. Ceci implique que la fonction de production agrégée est identique pour tous les pays et constitue alors une hypothèse contradictoire avec la définition donnée par les auteurs. D'autre part, A_0 est supposé non-corrélé avec s ou n , ce qui est difficilement acceptable.

Dans les équations de convergence en coupe transversale, le niveau initial de technologie A_0 n'est pas observé. Si ce terme est ignoré, alors qu'il est corrélé avec les variables explicatives incluses dans la régression, les estimations des paramètres seront biaisés. Finalement, ignorer le terme A_0 produit donc un biais de variables omises (Islam, 1995, 1998a).

Dans ces conditions, il faut mobiliser les méthodes de l'économétrie des panels pour pouvoir particulariser la constante et introduire de l'hétérogénéité dans la spécification des équations de β -convergence conditionnelle.

La plupart des modèles estimés dans la littérature sont des modèles de panel dynamiques à effets fixes. Compte tenu de la possible corrélation entre les effets individuels et les variables exogènes du modèle, il n'est pas possible d'utiliser une modélisation en termes d'effets aléatoires. D'une façon générale, ces équations de β -convergence en panel s'écrivent alors de la façon suivante :

$$\ln y_{it} - \ln y_{it-\tau} = \beta \ln y_{it-\tau} + \delta X_{i,t-\tau} + \eta_i + \xi_t + \varepsilon_{it} \quad [1.38]$$

où y_{it} est le produit par tête de l'économie i à la période t , X est le vecteur contenant les déterminants de la croissance, η_i est l'effet fixe individuel, ξ_t est l'effet fixe temporel et ε_{it} est le terme d'erreur.

Knight *et al.* (1993) et Islam (1995) reprennent l'étude de Mankiw *et al.* (1992) et réestiment leur équation de β -convergence avec effets fixes individuels, grâce à l'estimateur à distance minimale de Chamberlain (1982, 1983). Dans ces deux études, la vitesse de convergence trouvée est bien plus importante qu'en coupe transversale : elle se situe entre 5 à 10% par an.

Caselli *et al.* (1996) étendent les résultats précédents en tenant compte de deux problèmes rencontrés dans les études en coupe transversale : d'une part, la corrélation entre les effets fixes et les variables explicatives et, d'autre part, l'existence d'un biais de simultanéité dû à la présence de régresseurs endogènes. Sur la base d'un échantillon de 97 pays sur la période 1960-1985 avec des intervalles de 5 ans, l'équation de β -convergence conditionnelle est alors estimée en différence première à l'aide de l'estimateur de la méthode des Moments Généralisés (GMM) proposée par Arellano et Bond (1991). Les résultats font également apparaître une vitesse de convergence nettement plus élevée que dans les autres études empiriques, de l'ordre de 10% par an.

Lee *et al.* (1997) généralisent les études précédentes en permettant également l'hétérogénéité dans le taux de croissance à l'état régulier. A partir d'un échantillon de 102 pays pour la période 1960 et 1989, l'hypothèse d'égalité des taux de croissance à l'état régulier est fortement rejetée alors que la vitesse de convergence estimée atteint 30% par an. Comme le soulignent Islam (1998b) et Lee *et al.* (1998), cette dernière approche a pour conséquence de « vider » le concept de convergence entre différentes économies. En effet, chaque économie converge vers des niveaux de produit par tête différents à un rythme différent. L'interprétation des résultats est rendue encore plus difficile dans ce contexte.

Les vitesses de convergence rapides ne sont compatibles avec le modèle de croissance néoclassique de base que si l'on considère un concept restreint de capital et de fortes productivités marginales décroissantes. Une autre interprétation consiste cependant à supposer que des vitesses de convergence rapides reflètent d'autres mécanismes de convergence, en particulier, le phénomène de rattrapage technologique par comblement de l'écart technologique entre les différents pays et l'économie leader grâce aux processus de diffusion technologique (Abramovitz, 1986 ; Bernard et Jones, 1996 ; Hall et Jones, 1999). Si le progrès technique est un bien public, alors les pays les moins avancés sont favorisés et bénéficient d'un fort potentiel de croissance, à condition qu'ils aient une « capacité sociale »¹² (*social capability*) d'absorber les technologies étrangères pour les adapter à leurs propres besoins et que l'environnement politique et macroéconomique soit favorable (Abramovitz, 1986). Dans ce cas, ces pays sont dans une meilleure position que les pays leader qui doivent assumer les coûts financiers et temporels associés au développement de nouvelles technologies. Au lieu de supposer que les différences en technologie ne jouent pas de rôle systématique dans la convergence (Mankiw *et al.*, 1992) ou que ces différences existent mais sont stationnaires (Islam, 1995), diverses études empiriques différenciant les deux sources de convergence sont récemment mises en œuvre pour estimer explicitement le rôle de la diffusion technologique dans le processus de convergence (Paci et Pigliaru, 1999, 2001 ; De la Fuente, 1995, 1997, 2000).

¹² c'est-à-dire, selon Abramovitz (1986), des facteurs tels que le niveau d'éducation et la disponibilité de personnels techniques et scientifiques qualifiés.

Bien que les méthodes de données de panel présentent l'avantage de pouvoir contrôler l'hétérogénéité non observée dans les équations de β -convergence conditionnelle, plusieurs inconvénients doivent être soulignés.

Premièrement, les propriétés théoriques des estimateurs sur modèles de panel dynamiques sont asymptotiques et leurs propriétés en échantillon fini ne sont pas bien connues. La seule façon de connaître leurs propriétés en échantillon fini serait de conduire des simulations de Monte-Carlo (Islam, 1998a).

Deuxièmement, Barro (1997) souligne que les transformations en différences premières utilisées pour éliminer les effets fixes réduisent la précision des estimateurs (qui n'utilisent plus que la transformation *within*) et peuvent exacerber les effets des erreurs de mesure.

Troisièmement, Durlauf et Quah (1999) estiment que l'approche, consistant à éliminer les effets fixes en utilisant les différences premières avec effets fixes, a pour inconvénient de laisser inexplicables les variations de la croissance de long terme entre les économies, c'est-à-dire finalement ce qui a motivé le développement de cette littérature. Les estimations faites à l'aide de ces méthodes seraient donc plus pertinentes pour analyser les cycles d'affaires que pour étudier la croissance.

Quatrièmement, Shioji (1997a,b) et De la Fuente (1998), pensent que les estimations en panel ont tendance à surestimer systématiquement les vitesses de convergence. En effet, les échantillons utilisés sont le plus souvent de taille trop restreinte et les estimations captent plus vraisemblablement des ajustements de court terme autour de la tendance que la dynamique de long terme. Ces auteurs montrent alors que corriger ce biais ramène vers des taux de convergence proches des 2% obtenus habituellement.

121.3 Les clubs de convergence

L'hypothèse de β -convergence absolue est souvent rejetée pour de grands échantillons de pays mais généralement acceptée lorsqu'elle est testée sur un sous-ensemble de pays sélectionnés sur la base de la proximité de leurs conditions initiales ou terminales. Il convient dans ces conditions de s'interroger sur la possibilité de l'existence de *clubs de convergence*, c'est-à-dire des groupes de pays ou de régions pour lesquels on accepte l'hypothèse de convergence (Baumol, 1986 ; Durlauf et Johnson, 1995).

Un club de convergence est un groupe d'économies dont les caractéristiques structurelles sont identiques et dont les conditions initiales sont suffisamment proches pour converger vers un même équilibre de long terme. Au sein d'un même échantillon, il peut ainsi y avoir plusieurs clubs de convergence qui ne convergent pas entre eux. Il existe alors des seuils de convergence qui permettent à certaines économies de converger vers des équilibres caractérisés par de hauts niveaux de produit par tête, alors que d'autres économies convergent vers des équilibres caractérisés par de bas niveaux de produit par tête et tombent ainsi dans des pièges de sous-développement. L'hypothèse de convergence en clubs est ainsi liée à

l'existence d'équilibres multiples (paragraphe 121.31) et nécessite le développement de procédures empiriques adaptées (paragraphe 121.32).

121.31 Equilibres multiples dans les modèles de croissance

Dans la théorie néoclassique standard, le résultat de convergence conditionnelle repose sur l'unicité de la solution stationnaire de long terme. En revanche, l'existence de clubs de convergence peut être basée sur des modèles de croissance caractérisés par la possibilité d'équilibres multiples, comme par exemple, les modèles de croissance endogène. Ainsi, des seuils dans l'accumulation des facteurs de production ou dans le niveau de développement des secteurs clés de l'économie permettent de générer des équilibres multiples dans le long terme (Azariadis et Drazen, 1990). Des équilibres multiples apparaissent également dans le modèle de Becker *et al.* (1990), qui est un modèle en économie fermée où le progrès technique et la croissance de la population sont endogènes.

La mise en évidence de clubs de convergence ne constitue pas pour autant une preuve de la validité des modèles de croissance endogène. En effet, dans le cadre des modèles de croissance néoclassique, Galor (1996) montre que l'unicité de l'équilibre constitue seulement un cas particulier, obtenu notamment sous des hypothèses fortes sur les comportements des agents et la technologie de la production. L'existence d'équilibres multiples de long terme et donc de clubs de convergence reste ainsi parfaitement compatible avec le modèle de croissance néoclassique dès lors que l'on relâche certaines de ces hypothèses (comme par exemple l'introduction de l'hétérogénéité des comportements d'épargne), ou si l'on introduit des éléments additionnels tels que le capital humain, la distribution des revenus ou la fécondité.

121.32 Détection empirique des clubs de convergence

D'un point de vue économétrique, l'existence d'équilibres multiples se caractérise par une instabilité des paramètres dans les équations de convergence. Dans ces conditions, les équations de convergence traditionnelles, qui imposent une restriction forte de stabilité des paramètres, sont susceptibles d'être affectées par un problème de biais des estimateurs. Par exemple, supposons que l'économie mondiale soit polarisée en deux groupes distincts. On ne devrait pas estimer l'équation globale [1.32] mais plutôt les régressions suivantes :

$$\frac{1}{T} [\ln y_{iT} - \ln y_{i0}] = \alpha_1 + \beta_1 \ln y_{i0} + \varepsilon_t \quad \text{pour le groupe 1} \quad [1.39a]$$

$$\frac{1}{T} [\ln y_{iT} - \ln y_{i0}] = \alpha_2 + \beta_2 \ln y_{i0} + \varepsilon_t \quad \text{pour le groupe 2} \quad [1.39b]$$

Dans un tel cas, Bernard et Durlauf (1996) montrent que les régressions en coupe transversale, appliquées à des données générées par des économies qui convergent vers des

équilibres multiples, peuvent produire un coefficient associé au PIB par tête initial négatif. Le résultat de β -convergence standard apparaît alors comme un artéfact.

Un problème empirique important est la détermination de la composition des groupes et différentes méthodes ont été proposées.

Un premier groupe de travaux détermine cette composition à l'aide d'un ou de plusieurs critères fixés *a priori* pour vérifier ensuite que l'hypothèse de β -convergence est acceptée pour chacun des groupes. Ces critères sont le plus souvent l'appartenance à une zone géographique (Baumol, 1986 ; Grier et Tullock, 1989) ou à un système institutionnel (Baumol, 1986). D'autres auteurs préfèrent la fixation *a priori* des niveaux limites de PIB par habitant (Durlauf et Johnson, 1995 ; Beine *et al.*, 1999). Les résultats fournis par Mankiw *et al.* (1992) suggèrent également l'existence de clubs de convergence. En effet, le pouvoir explicatif de leur modèle est très faible pour leurs échantillons de 98 et de 75 pays. En revanche, la convergence absolue est acceptée lorsque l'échantillon est restreint aux pays de l'OCDE.

D'autres approches ont pour but d'éviter le recours à des voies arbitraires en cherchant à endogénéiser la sélection des clubs de convergence, c'est-à-dire en essayant de qualifier la forme de la distribution des PIB.

Certains auteurs utilisent, par exemple, des fonctions polynomiales du PIB par tête initial. Ainsi Baumol et Wolff (1988) régressent la différence des logarithmes de PIB par habitant entre 1950 et 1980 sur le niveau et le carré du PIB par habitant en 1950. La méthodologie proposée par Chatterji (1992) s'inspire de celle de Baumol et Wolff mais impose de déterminer un pays ou une région leader sur la période et d'étudier l'évolution des écarts par rapport à ce leader. Ainsi, il régresse pour 109 pays l'écart du PIB par habitant en logarithme de chacun d'entre eux, par rapport aux Etats-Unis, sur un polynôme d'ordre trois ayant pour variable l'écart précité. Cette méthode permet de dégager deux clubs de convergence parmi ces pays. Elle a également été appliquée sur les régions britanniques par Chatterji et Dewhurst (1996) et sur les régions européennes par Armstrong (1995b) et Dewhurst et Mutis-Gaitan (1995).

A titre d'illustration, nous présentons la méthodologie employée par Chatterji (1992) et Chatterji et Dewhurst (1996). La première étape consiste à identifier une économie leader sur toute la période étudiée. Ensuite, on étudie l'évolution de l'écart entre l'économie leader et les autres économies dans le temps. Ceci s'effectue à l'aide la régression suivante :

$$\ln \frac{Y_{LT}}{Y_{iT}} = \sum_{k=1}^n \alpha_k \left(\ln \frac{Y_{L0}}{Y_{i0}} \right)^k + \varepsilon_i \quad [1.40]$$

Y_{LT} est le PIB par tête de l'économie leader à la fin de la période, Y_{iT} est le PIB par tête de l'économie i à la fin de la période. Par conséquent, $\ln(Y_{LT} / Y_{iT})$ mesure l'écart entre l'économie i et l'économie leader à la fin de la période. De la même façon, $\ln(Y_{L0} / Y_{i0})$ mesure l'écart entre l'économie i et l'économie leader au début de la période.

Les auteurs choisissent enfin la valeur optimale de k sur la base des statistiques de Student et des critères d'information tout en reconnaissant que la multicollinéarité entre les diverses puissances de $\ln(Y_{L0} / Y_{i0})$ est susceptible de fausser cette démarche. Supposons qu'à la suite de l'application de cette méthodologie, on obtienne la courbe de régression estimée suivante :

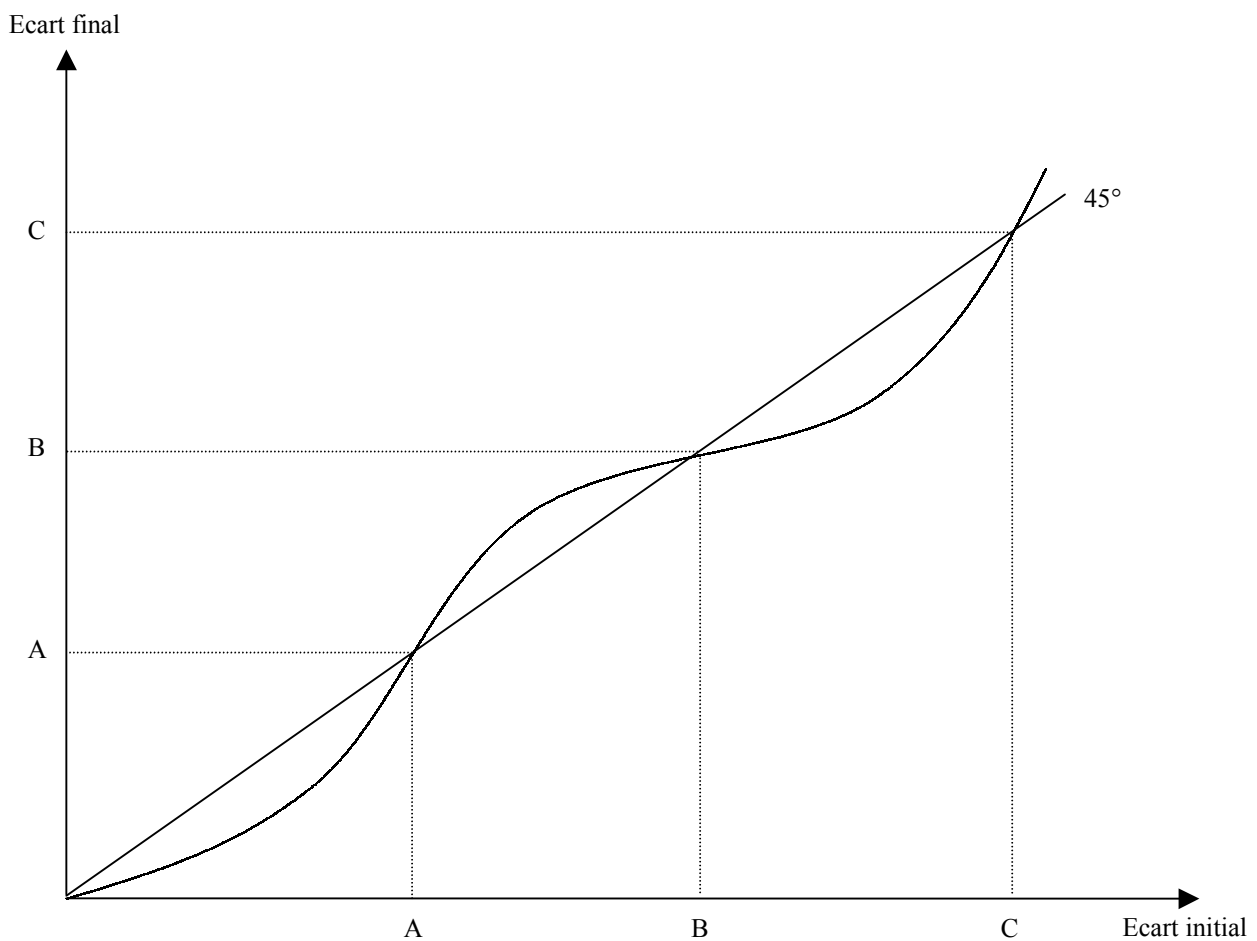


Figure 1.4 : Relation entre écart initial et écart final

Pour les économies dont l'écart initial se situe entre 0 et A, l'écart décroît dans le temps. Ces économies convergent alors vers le leader et constituent avec ce dernier un club de convergence. Pour les économies dont l'écart initial se situe entre A et B, l'écart augmente dans le temps et elles convergent vers un écart d'équilibre avec le leader qui se situe en B. Les

économies dont l'écart initial avec le leader se situe entre B et C tendent également vers l'écart d'équilibre B mais dans leur cas, l'écart avec le leader a tendance à diminuer. Ainsi, ces économies dont l'écart initial est compris entre A et C forment un deuxième club de convergence. Enfin, les économies dont l'écart initial est supérieur à C divergent à la fois du leader et du second club de convergence.

D'autres méthodes consistent à employer divers algorithmes tels que la méthode des arbres de régression avec données en coupe (Breinan *et al.*, 1984) pour mettre en évidence les clubs de convergence. Différents critères, ou variables de contrôle, doivent être mobilisés pour révéler les clubs de convergence. Il peut s'agir du PIB par tête et du taux d'alphabétisation (Durlauf et Johnson, 1995), du niveau de développement financier et du stock de capital humain (Berthélemy et Varoudakis, 1995, 1996), ou encore le PIB par tête initial, le montant de R&D et l'aide à l'investissement de l'Union Européenne (Fagerberg et Verspagen, 1996). Durlauf et Johnson (1995) utilisent la base de données de Summers et Heston sur la période 1960-1995. Ils obtiennent 4 clubs de convergence et montrent qu'il n'y a pas de convergence entre les régions les plus riches. Ces résultats sont qualitativement similaires à ceux obtenus par De Long (1988). Cependant, les auteurs reconnaissent que cette méthode est relativement compliquée à mettre en œuvre. En outre, elle ne permet pas de tester la significativité du nombre de régimes.

Jean-Pierre (1997) propose la méthode des régressions sur données réordonnées. Cette méthodologie permet de déterminer les seuils, sur la variable de classement, choisis à partir d'un graphique représentant les t de Student récursifs. Sur la base d'un échantillon mondial de 91 pays sur la période 1960-1989, l'auteur utilise le niveau de PIB par tête en 1960 pour ordonner les observations. L'application de la méthodologie fait apparaître deux régimes de convergence. Le premier comprend les pays avec un PIB par tête en 1960 inférieur à 2 400 \$ et le second comprend les pays dont le PIB par tête en 1960 est supérieur à cette valeur.

L'ensemble des développements présentés jusque-là ont en commun l'étude de la convergence « rattrapage », celle-ci étant définie comme la faculté qu'ont les économies à converger vers un état d'équilibre de long terme. Cette faculté dépend de deux éléments : les conditions initiales et la valeur de cet état régulier. Les concepts de β -convergence absolue et de β -convergence conditionnelle sont associés au fait que, indépendamment des conditions initiales, les économies convergent vers le même état régulier (dans le cas absolu) ou vers leur propre état régulier (dans le cas conditionnel). La convergence en clubs signifie que, suivant les conditions initiales, différents états réguliers peuvent être ou ne pas être atteints. Dans ces conditions, la vision simpliste, consistant à interpréter la notion de convergence-rattrapage comme un rattrapage des économies riches par les économies pauvres, doit être rejetée. Les critiques du concept de β -convergence, basées sur l'observation que les économies pauvres ne rattrapent pas les économies riches, sont tout autant peu fondées. En revanche, nous avons souligné que différentes limites pertinentes peuvent être formulées quant à la recherche de méthodes empiriques permettant d'estimer cette forme de convergence.

Dans un autre ordre d'idée, un ensemble de critiques s'est développé pour essayer d'évaluer d'autres facettes de l'idée de convergence.

122 Les autres concepts de convergence

Dans ce paragraphe, nous présentons des approches alternatives et d'autres concepts de convergence utilisés dans la littérature. Il s'agit, d'une part, des analyses basées sur une étude de la distribution des PIB par tête conduisant au concept de σ -convergence (paragraphe 122.1) et à l'analyse de la dynamique de transition des états de richesse des économies reposant sur l'étude des propriétés de chaînes de Markov (paragraphe 122.2). Il s'agit, d'autre part, des études basées sur les propriétés des séries temporelles et qui aboutissent au concept de convergence stochastique (paragraphe 122.3).

122.1 La σ -convergence

Dans cette approche, la convergence est associée à l'étude de l'évolution dynamique d'un indicateur de dispersion des PIB par tête entre les économies. On conclut à la convergence de l'ensemble des économies de l'échantillon lorsque cette dispersion diminue dans le temps. Le concept de σ -convergence pourrait donc être associé à une forme de réduction des inégalités.

Le test de la σ -convergence consiste à comparer un indicateur de dispersion, calculé pour la fin de la période, à cet indicateur calculé pour le début de la période. Deux indicateurs de dispersion sont habituellement utilisés : l'écart-type des PIB par tête en log [1.41] ou le coefficient de variation [1.42] :

$$v_t = \sqrt{\frac{1}{N} \sum_{i=1}^N \left[\log \left(\frac{y_{it}}{y_t^*} \right) \right]^2} \quad \log y_t^* = \frac{1}{N} \sum_{i=1}^N \log y_{it} \quad [1.41]$$

$$c_t = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{y_{it} - \bar{y}_t}{\bar{y}_t} \right)^2} \quad \bar{y}_t \equiv \frac{1}{N} \sum_{i=1}^N y_{it} \quad [1.42]$$

où y_{it} est le produit par tête de l'économie i à la période t , N est le nombre d'économies dans l'échantillon.

Comme le notent Dalgaard et Vastrup (2001), ces deux indicateurs attribuent des poids différents aux performances de croissance des économies de l'échantillon et peuvent parfois fournir des résultats différents. Nous étudions plus particulièrement le premier indicateur qui

permet d'établir un lien entre σ - et β -convergence (paragraphe 122.11) avant d'en souligner les principales limites (paragraphe 122.12).

122.11 La relation entre σ - et β -convergence

Dans la littérature « classique », selon la terminologie de Sala-I-Martin (1996b), deux concepts de convergence sont habituellement utilisés : la β -convergence (étudiée dans les paragraphes précédents) et la σ -convergence. Les liens entre ces deux concepts de convergence ont souvent été analysés.

La figure 1.5 représente l'évolution dans le temps de deux économies pour lesquelles on accepterait à la fois l'hypothèse de β -convergence absolue et celle de σ -convergence. En effet, l'économie B, la plus pauvre au début de la période, connaît un plus fort taux de croissance en début de période que l'économie riche, l'économie A, qui doit faire face à un taux de croissance négatif. Il y a donc un mouvement de β -convergence. Ce mouvement explique le rapprochement des deux économies au cours de la période et donc une diminution de la dispersion. Dans ce cas de figure, le processus de σ -convergence est induit par la β -convergence.

Cependant, il ne faut pas en conclure que ces deux notions fournissent la même information et conduisent systématiquement au même résultat. Ainsi, l'idée selon laquelle la croissance plus rapide des pays pauvres par rapport à celle des pays riches, entraîne une réduction de l'inégalité, est fautive. Procéder ainsi revient à commettre une erreur de type *Galton's fallacy*. Ce terme a une signification historique particulière et se réfère à l'origine à une étude effectuée par Francis Galton, le cousin de Charles Darwin. Il a examiné un échantillon de pères de très grande taille et noté que leurs fils étaient en moyenne d'une taille plus proche de la moyenne de la population (β -convergence). « L'erreur » de Galton a été d'inférer de ces observations que l'existence de cette propriété de retour à la moyenne entraînait une réduction dans la dispersion des tailles au sein de la population (σ -convergence).

En revanche, on peut montrer que la β -convergence est une condition nécessaire mais non suffisante à la σ -convergence (Friedman, 1992 ; Quah, 1993b ; Hart, 1995 ; Bliss, 1999) en s'appuyant sur la relation formelle définie par Quah (1993b) et Sala-I-Martin (1996a)¹³. Le point de départ est l'équation de convergence absolue où la variable expliquée est le taux de croissance cumulé :

$$\log(y_{iT}/y_{i0}) = a + \beta \log(y_{i0}) + u_i \quad [1.43]$$

¹³ Nous reprenons ici l'analyse proposée par Hénin et Le Pen (1995) pour illustrer ce point.

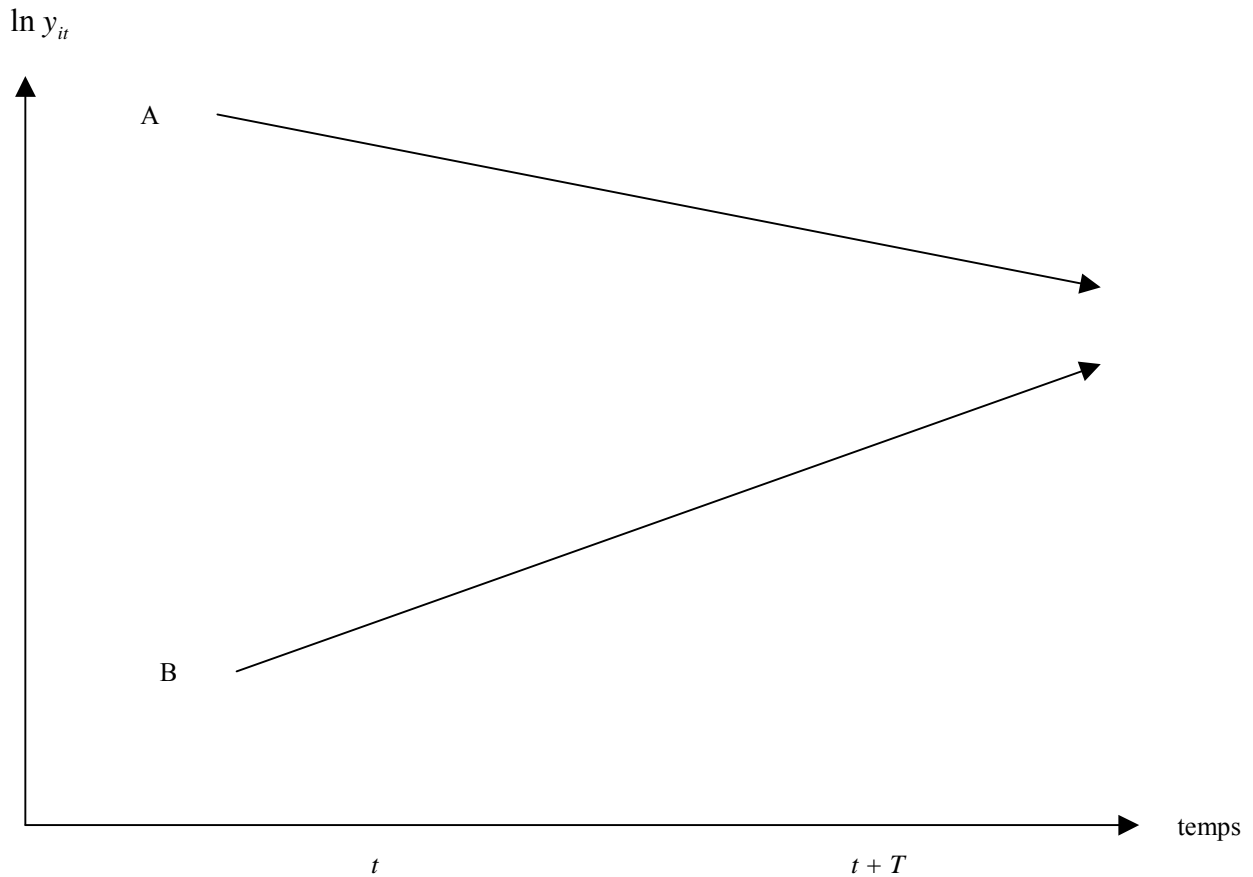


Figure 1.5 : β - et σ -convergence

Source : Adaptation de Sala-I-Martin (1996b), p. 1021

Cette équation se réécrit de la façon suivante :

$$\log(y_{iT}) = a + (1 + \beta) \log(y_{i0}) + u_i \quad [1.44]$$

En prenant la variance de chacun des membres de l'équation, on trouve :

$$V[\log(y_{iT})] = (1 + \beta)^2 V[\log(y_{i0})] + V(u_i) \quad [1.45]$$

Différentes réécritures de [1.45] permettent d'aboutir à l'équation suivante (Fuss, 1999) :

$$VR = \frac{V[\log(y_{iT})]}{V[\log(y_{i0})]} = \frac{(1 + \beta)^2}{R^2} \quad [1.46]$$

où R^2 est le coefficient de corrélation multiple de la régression [1.44].

On vérifie alors immédiatement que la β -convergence ($\beta < 0$) est une condition nécessaire mais non suffisante pour la σ -convergence ($VR < 1$). Le résultat final dépend en fait de la confrontation entre deux mécanismes qui ont des effets opposés. Le premier est l'existence de la β -convergence traduisant un comportement de retour à la moyenne qui a tendance à ramener toutes les séries vers le même niveau. Le second est relié à l'existence de chocs spécifiques auxquels sont soumises les économies et qui renouvellent en permanence la dispersion des produits par tête. La σ -convergence est la résultante de ces deux mécanismes et elle n'existe que lorsque les effets bénéfiques de la β -convergence dominent les effets négatifs des perturbations que subissent les économies.

En outre, de l'équation [1.46], on peut dériver une condition formelle pour l'existence de la σ -convergence. Cette dernière n'existe que si $VR < 1$, c'est-à-dire si :

$$(1 + \beta)^2 < R^2 \quad [1.47]$$

Il est alors possible de représenter cette condition par la figure 1.6. dans le plan $[(1 + \beta)^2, 1 - R^2]$. La condition de σ -convergence est appelée frontière de convergence dans le cas où $VR = 1$, c'est-à-dire lorsque $(1 + \beta)^2 + 1 - R^2 = 1$.

12.12 Les limites du concept de σ -convergence

La première critique porte sur la dépendance de ce concept de convergence à la date initiale. En effet, Sala-I-Martin (1996a) montre que la dispersion des PIB par tête est décroissante dans le temps lorsque la dispersion à la date initiale est supérieure à la valeur d'équilibre¹⁴. Cette dispersion est croissante si la valeur initiale est inférieure à la valeur d'équilibre. En conséquence, si les économies subissent un choc au début de la période ayant temporairement réduit la dispersion en coupe transversale inférieure à sa valeur d'équilibre, on observe un accroissement de cette dispersion alors que le choix d'une autre période initiale aurait pu aboutir au résultat inverse.

L'utilisation du concept de σ -convergence a également été critiquée par Quah (1993a, 1993b). Il s'agit d'un concept relativement réducteur de l'évolution de la distribution de la richesse des économies en coupe transversale. En effet, il ne permet pas de fournir une information pertinente sur la dynamique de la répartition des PIB par tête et sur l'évolution des économies au sein de cette répartition.

¹⁴ Cette valeur est définie par : $(\sigma^2)^* = \sigma_u^2 / [1 - (1 + \beta)^2]$, où σ_u^2 est la variance de l'erreur de la régression [1.44], supposée constante pour toutes les économies et supposée indépendante dans le temps et entre les économies.

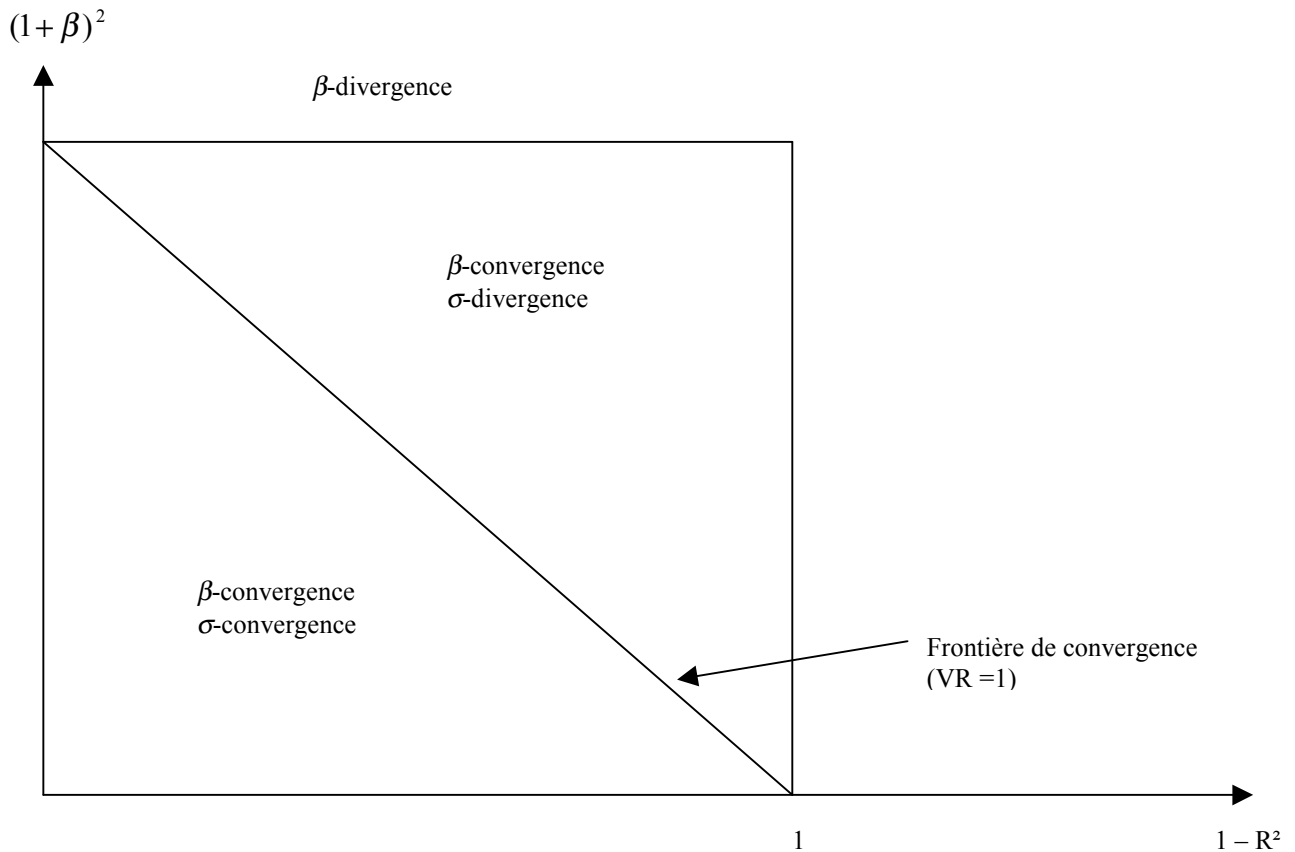


Figure 1.6 : Représentation des conditions de β - et de σ -convergence

Source : Hénin et Le Pen (1995), p. 670

Ainsi, la figure 1.7a représente deux pays ayant échangé leurs positions relatives entre la date initiale et la date finale sans que l'écart qui les sépare soit modifié. Dans ce cas, l'écart-type de cette répartition est constant sur la période alors que la situation entre les deux économies est très différente. Dans le même ordre d'idée, la figure 1.7b représente deux pays ayant inversé leurs positions relatives entre les deux dates de telle sorte que l'écart final est plus petit que l'écart initial. Dans ce cas, on ne peut pas rejeter les hypothèses de β - ou de σ -convergence. Pourtant, dans ces deux cas, il serait sans doute plus judicieux de conclure à un processus de divergence.

D'autres méthodes, mobilisant l'ensemble des informations contenues dans les distributions en coupe transversale des PIB par tête sur la période étudiée, ont ainsi été proposées dans le but d'analyser les trajectoires individuelles de chaque économie au sein de ces distributions.

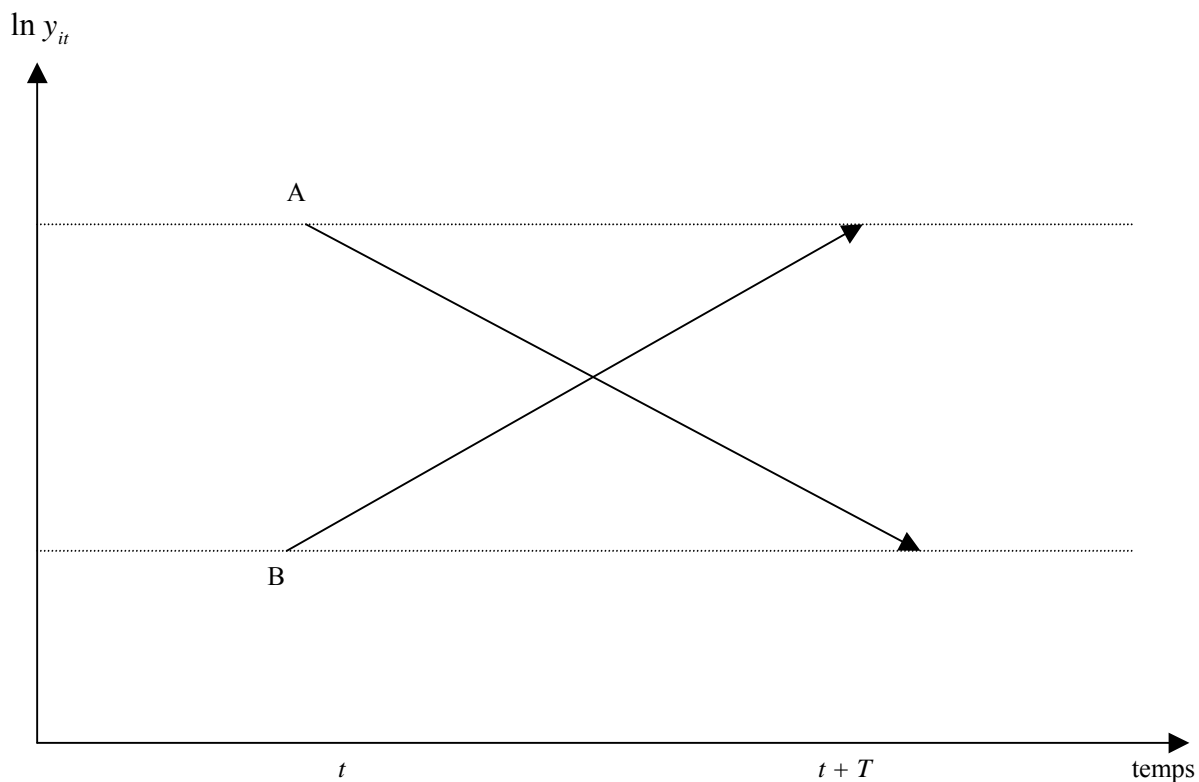


Figure 1.7a : σ -convergence, limites (1)

Source : Le Pen (1997), p. 733

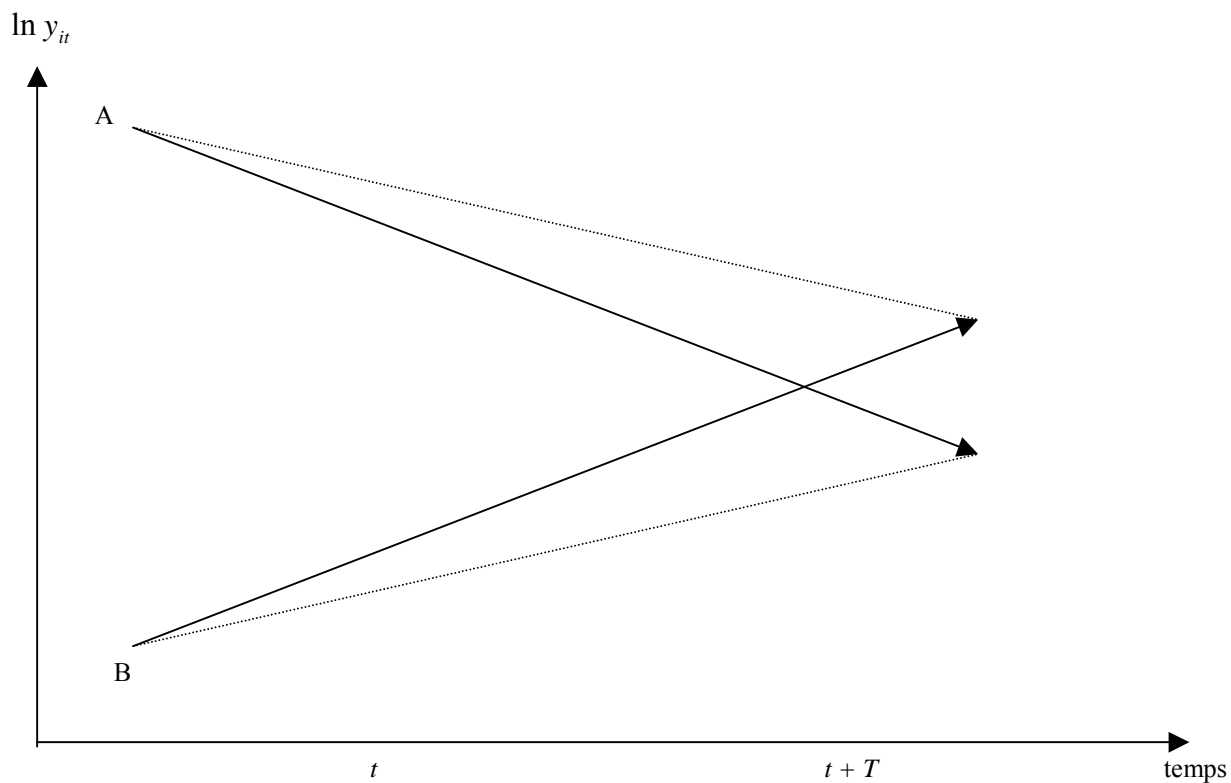


Figure 1.7b : σ -convergence, limites (2)

Source : Le Pen (1997), p. 733

122.2 L'étude de l'évolution des distributions en coupe transversale

Pour pallier les limites attachées au concept de σ -convergence, Quah (1993a, 1993b, 1996a) propose d'étudier la distribution en coupe transversale de ces séries dans son intégralité (c'est-à-dire sans se contenter d'analyser l'évolution d'un indicateur synthétique comme la dispersion) et la dynamique à l'intérieur de cette distribution. En d'autres termes, cet auteur cherche, d'une part, à étudier l'évolution de la forme de la distribution des niveaux de richesse au cours du temps et d'autre part, à identifier la position des économies ou de groupes d'économies à l'intérieur de cette distribution des PIB par tête. Ceci lui permet d'évaluer l'ampleur des changements de positions relatives de ses économies au sein de cette distribution.

Dans le premier cas, il s'agit d'estimer, à l'aide de techniques non-paramétriques, la fonction de densité des PIB par tête pour l'échantillon et d'examiner l'évolution de la forme de cette densité. La figure 1.8 représente deux modes d'évolution de la distribution au cours du temps, chacun représentant deux types de convergence. Si, à partir de la distribution initiale, les économies évoluent vers une distribution « resserrée », alors il y a convergence globale de toutes les économies vers un même niveau de PIB par tête. En revanche, si la distribution devient bimodale (ou multimodale), alors les économies convergent vers des niveaux différents, ce qui indique la constitution de deux (ou plusieurs clubs) de convergence.

Dans le second cas, on cherche à étudier si, au cours de la période d'étude, les économies ont changé leur position dans la distribution des richesses. On procède par l'estimation de matrices de transition, ou chaînes de Markov, qui sont construites à partir d'une discrétisation de la distribution des PIB par tête en plusieurs classes (en utilisant, par exemple, les quartiles ou les quintiles de la distribution). Les matrices de transition permettent ainsi d'estimer les probabilités de passage d'une classe à l'autre entre deux dates. L'étude de ces probabilités permet alors d'évaluer la force et les tendances de mobilité des régions au sein de la distribution des PIB par tête : si les probabilités de passage d'une classe à l'autre sont élevées alors la mobilité est élevée, si les probabilités de rester dans la même classe sont élevées alors la mobilité est faible, si les probabilités de passage d'une classe à une classe immédiatement supérieure sont plus élevées que celles de sauter plusieurs classes, alors la mobilité est plutôt progressive... Par extension, elle permet d'observer si le niveau de PIB par tête des économies a tendance à s'homogénéiser ou si, au contraire, des regroupements distincts s'opèrent. Cette méthode a été appliquée sur divers types d'échantillons : les pays du monde (Quah, 1993a, 1993b), les régions japonaises (Kawagoe, 1999) ou finlandaises (Kangasharju, 1999).

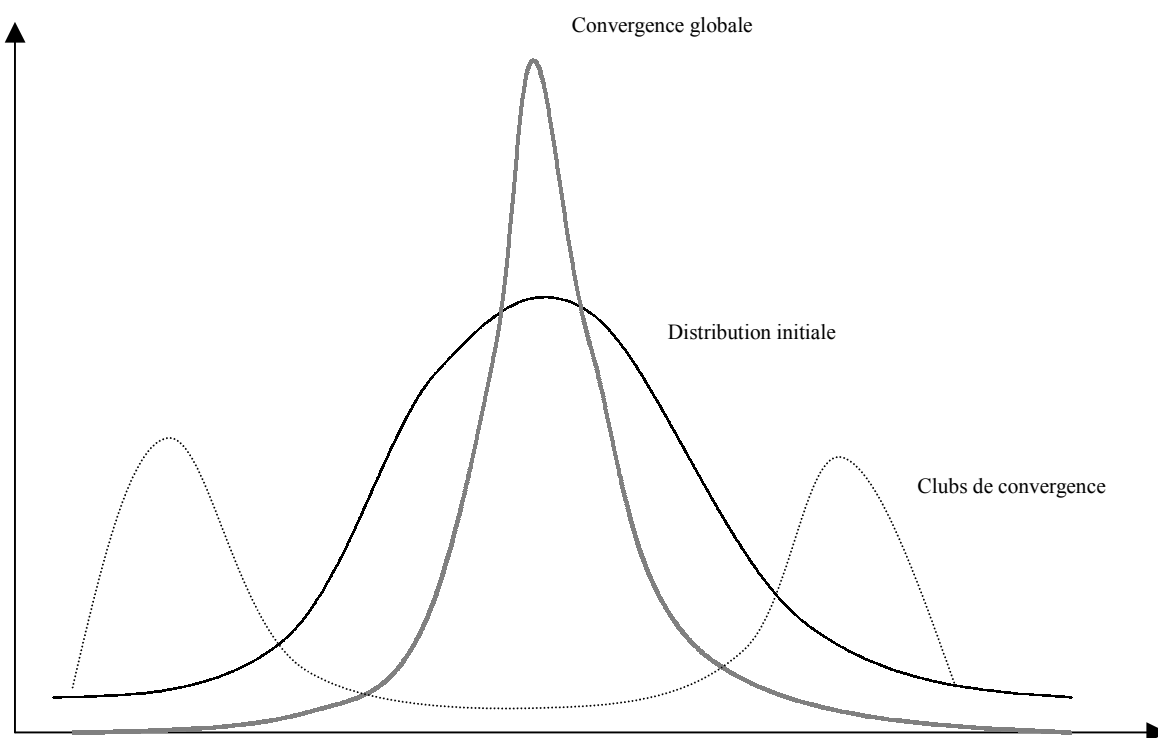


Figure 1.8 : Convergence et évolution de la distribution

Source : Adaptation de Fuss (1999), p. 236

122.3 La convergence comme persistance des écarts de PIB par tête

Le dernier concept de convergence que nous présentons s'appuie sur l'étude des propriétés temporelles des échantillons et se distingue du concept de β -convergence appliqué aux études en coupe transversale. En particulier, le concept de convergence défini par Bernard (1991) et Bernard et Durlauf (1995, 1996) repose sur la propriété de stationnarité des séries temporelles. Pour cette raison, on parle dans ce cas de convergence stochastique. On dit qu'il y a convergence stochastique si les prévisions à long terme des écarts de PIB par habitant entre deux ou plusieurs économies tendent vers zéro. Comme le notent Bernard et Durlauf (1996), cette définition n'est pas respectée si les chocs spécifiques, que subit chacune des économies, exercent des effets permanents sur leur trajectoire de long terme.

Pour tester cette hypothèse de stationnarité des écarts de PIB par tête avec une moyenne nulle, deux cas doivent être distingués. Dans le cas bivarié, le test de cette hypothèse de convergence revient à tester la présence d'une racine unitaire dans la série des écarts de

PIB par habitant. De nombreuses procédures de test de la racine unitaire sont présentées dans la littérature et les plus fréquemment utilisées sont la procédure de test de Dickey et Fuller (1981) ou de Phillips et Perron (1988). On peut noter toutefois que, dans ce contexte, l'hypothèse nulle testée est celle de non-stationnarité et par conséquent de non-convergence. Dans le cas multivarié, plus intéressant, on teste si les PIB par habitant des N économies de l'échantillon présentent une tendance commune en utilisant, par exemple, la méthodologie de Johansen et Juselius (1990). Le test de convergence revient donc, dans ce contexte, à tester la présence de $N - 1$ relations de cointégration entre les PIB par tête en logarithmes.

Les résultats empiriques obtenus dans les études en coupe transversale (β -convergence) et ceux obtenus en série temporelle (convergence stochastique) semblent être contradictoires. En effet, les tests effectués en coupe transversale attestent généralement la présence de la convergence tandis que les tests effectués en séries temporelles ne parviennent pas, généralement, à rejeter l'hypothèse de non-convergence. Cette apparente contradiction peut s'expliquer par la différence des concepts de convergence testés : convergence-rattrapage (β -convergence) ou convergence-stationnarité (convergence stochastique). Comme le notent Bernard et Durlauf (1996), les hypothèses implicites sous-jacentes à ces deux approches sont contradictoires. Dans la première, on suppose que les économies ont déjà atteint leurs trajectoires de long terme et que tout choc sur l'une des économies n'aura aucun effet persistant. En revanche, dans la seconde, les économies sont encore en phase de transition vers leurs états réguliers. Le choix entre ces deux méthodes repose donc sur les hypothèses adoptées sur la position des économies par rapport à leurs trajectoires de long terme. En outre, le fait qu'on ne puisse pas rejeter l'hypothèse de non-convergence stochastique peut aussi être dû à la faible puissance des tests de racine unitaire habituels et à la présence de chocs exogènes majeurs tels que les 2 guerres mondiales, la grande crise de 1929, les 2 chocs pétroliers, etc. En effet, dans ce cas, Perron (1989) indique que les tests de racine unitaire sont biaisés en faveur de l'hypothèse nulle de la racine unitaire, ce qui s'interprète, dans notre contexte, comme un biais en faveur de l'hypothèse de non-convergence.

Conclusion de section

Le concept de convergence a été compris et étudié de multiples façons : convergence-rattrapage, σ -convergence ou convergence stochastique. Pour tous ces tests, l'analyse empirique porte sur des régions ou des pays observés soit en coupe transversale, soit en série temporelle, soit en données de panel. Ces analyses nécessitent donc le recours à des données localisées. Cependant, il apparaît que la *dimension spatiale* des données est négligée dans ces approches car on y suppose implicitement que chacune des économies est une entité géographiquement indépendante des autres. On ignore ainsi la possibilité d'interactions spatiales. Les théories relevant de l'économie géographique et de la synthèse géographie-croissance offrent pourtant un fondement théorique pour l'intégration explicite de l'espace dans l'analyse des processus de croissance et de convergence.

SECTION 13 : ECONOMIE GEOGRAPHIQUE DES DISPARITES REGIONALES

Alors que les interactions économiques s'opèrent à l'échelle internationale, que les politiques d'intégration économique touchent de plus en plus de nations, que les politiques de développement à l'échelle européenne ou à l'intérieur de beaucoup de pays cherchent à réduire les inégalités régionales, l'analyse économique considère encore trop souvent que les économies sont a-territorialisées, oubliant la dimension géographique des échanges, définie comme les schémas de répartition des activités économiques sur un territoire. Les théories de l'économie géographique montrent que cette dimension n'est ni négligeable, ni neutre. En effet, les interactions économiques sont liées, d'une part, à toutes les formes d'accessibilité aux marchés et d'autre part, aux différents avantages de la proximité procurés par la concentration des activités économiques. Les choix de localisation, fondés sur ces principes, débouchent sur une répartition plus ou moins inégalitaire des activités économiques, des richesses et de la croissance entre les territoires.

Dans ces conditions, la dimension géographique des phénomènes économiques, si elle est observée, révèle une répartition particulière des données dans l'espace étudié. Les travaux empiriques sur la convergence doivent alors se préoccuper des conséquences économétriques impliquées par cette répartition (paragraphe 131). Les mécanismes économiques conduisant à ces schémas géographiques particuliers doivent également être analysés. Nous présentons tout d'abord les principes fondamentaux de l'économie géographique (paragraphe 132). Nous insistons ensuite sur la façon dont les effets de l'intégration régionale sur les disparités spatiales peuvent être étudiés dans le cadre du modèle de base de l'économie géographique proposé par Krugman (1991a, 1991b) et de ses raffinements (paragraphe 133). Nous soulignons enfin comment les modèles de la « synthèse géographie-croissance » permettent d'étudier les liens entre les processus de croissance et les mécanismes de répartition des activités économiques entre les régions (paragraphe 134).

131 Concentrations spatiales et tests de convergence

L'inégale répartition géographique de l'activité économique et de la croissance est une des caractéristiques les plus marquantes des économies contemporaines : les facteurs de production (incluant le capital humain et le capital physique) ont en effet tendance à se regrouper en certains points particuliers de l'espace, ce qui conduit à une concentration géographique des activités économiques. Ce « fait stylisé », pour reprendre les termes d'Easterly et Levine (2001), se retrouve à tous les niveaux de l'échelle spatiale : au niveau mondial, national ou régional.

Au niveau mondial, l'agglomération des activités correspond au dualisme Nord-Sud. Les pays à hauts revenus sont regroupés dans un petit nombre de zones tempérées dans l'hémisphère nord alors que le revenu décroît constamment avec la distance à ces zones. Il apparaît ainsi que les 20 pays les plus riches produisent 50% du PIB mondial avec seulement 15% de la population mondiale et que 54% du PIB mondial est produit par les pays occupant seulement 10% de la superficie totale (Easterly et Levine, 2001 ; Henderson *et al.*, 2001). Ce phénomène de concentration géographique se retrouve également au niveau des taux de croissance, comme l'illustre la croissance rapide des pays de l'Asie de l'Est et la stagnation des pays de l'Afrique sub-saharienne durant les dernières décennies.

A l'échelle infra-nationale, d'autres types d'agglomérations apparaissent sous la forme de fortes disparités régionales dans un même pays. Ainsi, aux Etats-Unis, les activités économiques sont essentiellement regroupées en deux grands arcs, l'un longeant la côte Ouest, et l'autre longeant la côte Est et les Grands Lacs. En 1990, 41,2% de l'emploi manufacturé se trouvaient dans 100 comtés occupant seulement 1,5% de l'aire totale (Hanson, 1998a). En Europe, la « banane bleue », qui comprend les régions se trouvant sur l'axe Londres-Francfort-Milan, concentre population et activités productives. Plus de la moitié de la population de l'Union Européenne est actuellement regroupée sur seulement 4% de son territoire, là où sont produits plus des deux tiers de la richesse (Commission Européenne, 1996, p. 24).

Cette concentration géographique dans les pays va de pair avec une urbanisation croissante, mettant alors en évidence une troisième forme de disparité spatiale : celle opposant les espaces urbains aux espaces ruraux. En Europe par exemple, la proportion de la population vivant dans les villes a connu, au cours du 20^{ème} siècle, une croissance explosive : 38% de la population vivait dans les villes en 1900 contre 75 % aujourd'hui (Bairoch, 1985). Ces agglomérations urbaines jouent par ailleurs un rôle fondamental dans le dynamisme économique. Ainsi, en 1997, l'Ile-de-France comptait 18,8% de la population française et représentait 2,2% du territoire national mais produisait 29% du PIB national (IAURIF, 2001). Au Japon, les préfectures urbaines de Tokyo, de Kanagawa, de Aichi, d'Osaka et de Hyogo sont regroupées sur 5,2% du territoire japonais mais elles comptabilisent 33% de la population totale et produisent 40% du PIB (Fujita et Thisse, 2002a).

Alors que ces faits stylisés sont bien connus, leurs implications sur les propriétés des séries en coupe transversale sont peu considérées, même si les phénomènes de concentration géographique sont susceptibles de remettre en cause les régressions de convergence absolue ou conditionnelle habituellement utilisées. En effet, ces régressions relient uniquement la croissance d'une économie à sa propre histoire, et non au système interrégional ou inter-pays auquel elle appartient (Martin et Sunley, 1998 ; Martin, 1999). Cette approche a deux conséquences.

D'une part, l'observation d'une concentration géographique de richesse ou de croissance dans certaines zones pourrait laisser penser à un processus de convergence en clubs, dans lequel les clubs sont définis sur une base géographique : les économies voisines

convergeraient alors vers un niveau de produit par tête propre à leur groupe. Au contraire, lorsque les observations sont regroupées, cela revient à supposer que le processus de convergence est identique pour tous les pays ou toutes les régions.

D'autre part, cette approche ne permet pas de prendre en compte la façon dont les différentes économies sont reliées les unes aux autres. Pourtant, si les économies à croissance rapide sont regroupées dans l'espace, de même que les économies à croissance lente, alors le taux de croissance d'une économie peut dépendre d'une façon importante de la croissance des économies voisines.

L'importance de l'impact de la localisation sur la croissance, dans les études internationales, a parfois été prise en compte en incluant, dans les équations de convergence conditionnelle, des variables muettes nationales ou continentales. Ainsi, Barro (1991) montre que les deux variables muettes représentant les pays d'Afrique subsaharienne et les pays d'Amérique Latine ont un effet significatif et négatif. D'autres études se sont penchées sur la possibilité de clubs de convergence construits sur la base de critères géographiques en distinguant, parmi les pays n'appartenant pas à l'OCDE, les pays du continent asiatique, du continent africain et du continent américain (Grier et Tullock, 1989).

Ces tentatives mises à part, l'interdépendance spatiale des économies reste le plus souvent ignorée et les approches traditionnelles traitent les observations comme si elles étaient indépendantes. Sur le plan économétrique, l'interdépendance des économies est pourtant susceptible d'avoir des effets importants sur la fiabilité des régressions de convergence, lorsque ces dernières sont estimées en supposant que les erreurs ne sont pas corrélées, ni dans le temps, ni dans *l'espace*. Ce problème est relevé par De Long et Summers (1991, p. 487) qui notent en annexe de leur article que les trajectoires de croissance dans les pays voisins sont le plus souvent très proches :

« Many comparative cross-country regressions have assumed that there is no dependence across residuals, and that each country provides as informative and independent an observation as any other. Yet it is difficult to believe that Belgian and Dutch economic growth would ever significantly diverge, or that substantial productivity gaps would occur within Scandinavia. »

En conséquence, si des variables sont omises dans les équations de convergence, et que ces variables prennent des valeurs similaires dans les pays voisins, alors les résidus de la régression seront vraisemblablement corrélés. L'inférence statistique, effectuée sur la base d'une estimation par les MCO n'est donc pas fiable (Temple, 1999a). Ce fait est repris et souligné par Mankiw (1995, p. 304-305) :

« For the reported standard errors to be correct, the residual for Canada must be uncorrelated with the residual for United States. If country residuals are in fact correlated, as is plausible, then the data most likely contain less information than the reported standard errors indicate. That is, statistical inference is overstated. »

Pour vérifier s'il existe effectivement une forme de corrélation spatiale entre les résidus des régressions de convergence, De Long et Summers (1991) forment, pour chaque paire de pays i et j , le produit $u_i u_j / \sigma^2$ des résidus estimés de leur régression, normalisé par l'estimation de la variance résiduelle. Ils régressent ensuite cette quantité sur la distance δ_{ij} (en miles) entre les capitales des pays i et j , en utilisant deux formes fonctionnelles (équations [1.48] et [1.49]) :

$$\frac{E(u_i u_j)}{\sigma^2} = \alpha + \frac{1}{1 + \lambda \delta_{ij}} + \varepsilon \quad [1.48]$$

$$\frac{E(u_i u_j)}{\sigma^2} = \alpha + \exp(-\lambda \delta_{ij}) + \varepsilon \quad [1.49]$$

Dans chaque cas, le coefficient de détermination ajusté est négatif et la valeur estimée de λ est très faible. Il apparaît que la corrélation estimée entre les résidus varie entre 0,03 pour les pays dont les capitales sont adjacentes, et 0,05 pour les pays dont les capitales sont distantes de plus de 10 000 miles (soit environ 16 090 km). Les auteurs construisent également un diagramme représentant, en ordonnée, les produits croisés des résidus estimés, normalisés par la variance résiduelle, et, en abscisse, la distance. Ce diagramme ne permet pas de confirmer l'hypothèse selon laquelle les pays voisins ont des résidus similaires.

Les auteurs sont cependant surpris de l'absence apparente d'un degré significatif de corrélation spatiale dans leur échantillon. Ceci pourrait être dû à la méthode utilisée qui ne permet pas de prendre en compte, d'une façon appropriée, le problème de la corrélation entre les erreurs d'une régression. Moreno et Trehan (1997), en utilisant les méthodes de l'économétrie spatiale, que nous présentons dans le chapitre 2 (p. 95) et que nous utilisons dans nos propres estimations, parviennent effectivement à montrer que le taux de croissance d'un pays est fortement relié à celui de ses régions voisines.

A l'issue de cette discussion, il apparaît donc surprenant que ces différents effets géographiques, même s'ils sont reconnus, ne sont pas davantage étudiés dans la littérature macroéconomique. Deux arguments, se renforçant mutuellement, justifient pourtant la prise en compte de la localisation géographique des observations et de leurs interdépendances dans le contexte des modèles de convergence. D'une part, différentes études théoriques, basées sur les modèles de l'économie géographique, expliquent pourquoi les activités économiques s'agglomèrent en un petit nombre de lieux et comment les schémas géographiques interagissent avec les processus de croissance (Krugman, 1991a ; Kubo, 1995 ; Martin et Ottaviano, 1999). D'autre part, les techniques statistiques et économétriques appropriées pour analyser ces effets sont développées dans le cadre de la littérature en économétrie spatiale (Cliff et Ord, 1981 ; Anselin, 1988a ; Anselin et Bera, 1998).

Alors que nous détaillons ces techniques dans le deuxième chapitre de la thèse, nous nous concentrons ici sur les modèles théoriques permettant de rendre compte de ces différents faits stylisés.

132 La formation des agglomérations : fondements et principes

L'économie géographique a pour objectif d'expliquer l'inégale répartition des activités et des hommes dans l'espace, c'est-à-dire la formation d'agglomérations en un petit nombre de localisations. Le paradigme compétitif, issu du modèle d'équilibre général, n'est pas capable de générer de telles agglomérations sans supposer *a priori* l'existence de fortes hétérogénéités dans l'espace ayant un pouvoir attractif (paragraphe 132.1). Dans ces conditions, l'explication de l'existence et de la persistance d'agglomérations doit passer, soit par l'abandon de l'hypothèse de marchés compétitifs et de l'hypothèse des rendements décroissants, soit par le recours aux externalités (paragraphe 132.2). La formation des agglomérations apparaît alors comme le résultat d'un équilibre entre des forces d'agglomération et de dispersion dont le contenu et l'intensité varient selon l'échelle spatiale retenue et l'époque (paragraphe 132.3).

132.1 L'agglomération en présence d'hétérogénéités

Dans le modèle d'équilibre général d'Arrow-Debreu (1954), l'économie est formée par des agents (firmes et ménages) et par des biens. Chaque firme est caractérisée par un plan de production alors que les consommateurs sont identifiés par une relation de préférence et un ensemble de ressources. Pour intégrer l'espace dans ce modèle d'équilibre général, les auteurs supposent qu'un bien n'est pas seulement défini par ses caractéristiques physiques, mais également par son lieu de disponibilité. Deux biens physiquement identiques sont donc considérés comme différents s'ils sont disponibles en deux lieux distincts. Dans ce cadre d'analyse, choisir une localisation revient à choisir un bien selon les mêmes critères (Thisse, 1992, 2000).

Pour montrer l'existence et l'unicité d'un système de prix qui équilibre simultanément tous les marchés, Arrow et Debreu (1954) doivent supposer la convexité des préférences des consommateurs et des plans de production des firmes. Ces hypothèses, déjà restrictives, deviennent intenable dans le contexte d'une économie spatiale.

En effet, du côté du consommateur, la convexité des préférences signifie qu'un consommateur préfère toujours une petite quantité d'un grand nombre de biens. Si l'on raisonne sur le choix d'une résidence, cela signifie qu'un consommateur préfère résider en un grand nombre de lieux, ce qui n'est pas le comportement habituel. Du côté du producteur, l'hypothèse de convexité des ensembles de production est encore plus problématique car elle

implique des rendements non croissants sur la production. Ainsi, il est plus efficace de produire dans des unités de production de plus en plus petites. A la limite, si la distribution des ressources naturelles est uniforme dans l'espace, l'économie est telle que chacun produit pour sa propre consommation : on se trouve dans un monde à la Robison Crusoë et une économie de type *backyard capitalism*. Même si le nombre de firmes est donné, cette hypothèse suppose que chaque firme préfère toujours ouvrir des établissements dans un grand nombre de lieux.

En conséquence, on ne peut expliquer la concentration en un même lieu de plusieurs unités de production et l'hypothèse de rendements d'échelle non croissants n'est compatible qu'avec un paysage recouvert en chaque point d'une économie autarcique. Les rendements d'échelle croissants deviennent alors essentiels pour expliquer la distribution géographique des activités dans l'espace. Ce résultat est appelé *Folk Theorem* de l'économie géographique (Scotchmer et Thisse, 1992).

L'hypothèse de convexité a une autre conséquence : celle de supposer que tous les biens et les agents sont divisibles. Cette hypothèse étant irréaliste, il est naturel de vouloir introduire l'indivisibilité des activités économiques. Cependant, la prise en compte simultanée d'indivisibilité et de coût de transport n'est pas compatible avec un équilibre concurrentiel. Ce point a été démontré par Starrett (1978). Supposons un espace homogène avec un nombre fini de localisations. L'espace est homogène lorsque l'ensemble de production est le même dans toutes les localisations, les préférences des consommateurs sont les mêmes dans toutes les localisations et les ressources naturelles sont réparties également dans l'espace. Alors, sans faire d'hypothèses sur la convexité des préférences et des technologies, Starrett (1978) énonce le théorème d'impossibilité spatiale suivant : si l'espace est homogène et le transport coûteux, alors il n'existe pas d'équilibre compétitif impliquant du transport.

Ce théorème a deux conséquences. D'une part, si les activités économiques sont parfaitement divisibles, alors un tel équilibre existe et il est tel que chaque localisation est une économie autarcique. On retrouve le monde de Robison Crusoë. D'autre part, si les activités économiques ne sont pas parfaitement divisibles, alors le transport des biens entre deux localisations devient inévitable. Dans ce cas, le théorème implique la non-existence d'un équilibre concurrentiel (Fujita et Thisse, 2002a, chap. 2).

Finalement, pour expliquer la formation d'agglomérations dans un cadre concurrentiel, il faut supposer que l'espace est hétérogène. On peut ainsi introduire une distribution inégale des ressources naturelles, comme dans la théorie néoclassique du commerce international. On peut aussi postuler *a priori* l'existence d'un ou de plusieurs centres d'attraction. C'est la voie choisie par la Nouvelle Economie Urbaine (Alonso, 1964 ; Muth, 1969 ; Fujita, 1989) : les choix de résidence des consommateurs au sein des espaces urbains sont expliqués par l'accessibilité à un ou plusieurs centres d'emplois fixés de manière exogène. Cependant, cette solution ne permet pas *d'expliquer* pourquoi les activités s'agglomèrent en un ou plusieurs lieux (Baumont et Huriot, 2000). Pour cela, il faut, soit abandonner le cadre de la concurrence parfaite et des rendements constants, soit introduire des externalités dans la production.

132.2 Concurrence monopolistique et externalités

Les mécanismes de formation des agglomérations s'appuient sur deux concepts fondamentaux : la concurrence monopolistique et les externalités.

1) *La concurrence monopolistique*

La concurrence monopolistique définit une situation de concurrence imparfaite ayant les caractéristiques suivantes (Matsuyama, 1995) :

- Les produits sont différenciés. Par conséquent, chaque firme est consciente de son pouvoir de monopole et peut fixer le prix de son bien.
- Le nombre de firmes (et de produits) est si grand que chaque firme ignore les interactions stratégiques avec les autres firmes. De ce fait, les actions d'une firme particulière ont un impact négligeable sur l'économie agrégée.
- L'entrée est libre et se réalise jusqu'à ce que les profits des firmes tendent vers 0.

La structure de concurrence monopolistique, basée sur le modèle de Dixit-Stiglitz (1977), est très largement utilisée dans les théories de l'agglomération bien que ce modèle ne soit pas explicitement « spatial ». Ce modèle permet d'abord d'intégrer la concurrence imparfaite et les rendements croissants dans une grande souplesse formelle. Ensuite, au contraire des modèles oligopolistiques, les problèmes d'interactions stratégiques entre les firmes, en termes de concurrence en prix, sont évacués puisque chaque produit n'a pas de concurrent immédiat. Enfin, l'intérêt de ce modèle repose sur la structure de préférence des agents qui ont une fonction d'utilité de type CES (Thisse, 1997). Les agents ont alors une préférence pour la diversité des produits. Comme nous le verrons dans le paragraphe 133.1 (p. 76), ces différents éléments poussent à l'agglomération des producteurs et des consommateurs par un effet « boule de neige » : la présence d'un grand nombre de producteurs en une localisation entraîne une plus grande disponibilité de produits locaux. Ceci incite les consommateurs, qui ont une préférence pour la variété, à venir se localiser en ce site, ce qui, en retour, attire d'autres firmes fabriquant des biens différenciés.

2) *Les externalités*

Depuis Scitvosky (1954), il est courant de classer les effets externes en deux catégories suivant la manière dont ils se manifestent : les externalités pécuniaires et les externalités technologiques.

- Les premières passent par les mécanismes traditionnels de marché et « renvoient aux bénéfices des interactions économiques qui se concrétisent au travers des mécanismes habituels de marché (*i.e.* au travers des prix). » (Fujita et Thisse, 1997, p. 43).

- Les secondes recouvrent les effets d'interactions qui se réalisent en dehors des mécanismes de marché et qui affectent directement l'utilité des consommateurs ou la fonction de production des entreprises, en raison de la présence au même lieu d'autres agents économiques. Les externalités informationnelles sont, par exemple, souvent classées dans cette catégorie.

Il faut noter que cette terminologie est relativement critiquable car les externalités pécuniaires ne sont pas des externalités au sens d'Arrow (1970) ou de Baumol et Oates (1975) pour qui l'externalité, positive ou négative, se place nécessairement en dehors du marché concurrentiel¹⁵. D'une manière générale, les deux formes sont sollicitées dans les théories de la formation des agglomérations.

Les externalités correspondent souvent à des non-convexités dans la production ou la consommation. Elles peuvent se réaliser à trois niveaux différents et expliquent l'intérêt d'une forme particulière de concentration (Duranton, 1997). Suivant la terminologie de Hoover (1936), les économies d'échelle sont internes à la firme et justifient le développement de grandes firmes. Les économies de localisation sont externes à la firme mais affectent toutes les firmes appartenant au même secteur : elles pourraient ainsi justifier la concentration géographique de firmes travaillant dans la même branche et sont souvent évoquées pour expliquer la spécialisation productive de certains espaces. Enfin, les économies d'urbanisation sont externes à la branche mais internes au système économique : elles apparaissent lors de l'agglomération d'industries différentes et sont ainsi représentatives des espaces urbains.

Les externalités semblent alors déterminantes dans l'explication des processus de concentration spatiale des individus et des firmes. Elles peuvent agir en tant que forces d'attraction, si elles sont positives, ou de forces de répulsion, si elles sont négatives. Elles restent par ailleurs souvent associées à d'autres forces d'agglomération et de dispersion.

132.3 L'agglomération : un équilibre des forces centripètes et centrifuges

La présence d'avantages comparatifs dans une localisation particulière, comme la présence de ressources naturelles ou la proximité à un nœud de transport, sont des « causes de première nature » (Krugman, 1993a). Ces éléments permettent d'expliquer pourquoi une activité a préféré se localiser en un endroit particulier plutôt qu'en un autre lieu. Si plusieurs agents envisagent la proximité à ces facteurs de première nature comme importante, alors une agglomération se formera en ce lieu. Cependant, les causes de première nature restent des éléments exogènes et ne permettent pas d'expliquer les forces qui poussent les agents à se regrouper les uns près des autres. Par ailleurs, les causes de première nature ne peuvent pas toujours rendre compte de la persistance d'une agglomération. Pour cela, il faut déterminer les

¹⁵ Guillain (2000) fournit une analyse détaillée de ce point.

« causes de seconde nature », c'est-à-dire les raisons pour lesquelles les agents ont un intérêt à se localiser à proximité les uns des autres. Par exemple, si un agent choisit de manière aléatoire sa localisation dans un ensemble de lieux identiques, cette première décision, qui crée une hétérogénéité entre les lieux, est à même d'enclencher un processus d'agglomération des autres agents en cet endroit. Il s'agit ainsi d'identifier les forces d'agglomération, ou forces centripètes (paragraphe 132.31). Cependant, la croissance d'une agglomération n'est pas infinie et il existe des forces de dispersion (ou forces centrifuges) qui poussent les agents à s'éloigner les uns des autres (paragraphe 132.32). La formation des agglomérations est alors décrite comme un équilibre entre les forces d'agglomération et les forces de dispersion (paragraphe 132.33).

Nous présentons rapidement les forces d'agglomération et de dispersion les plus utilisées. Pour un panorama plus complet, on peut consulter Duranton (1997) ou Fujita et Thisse (1997, 2000a, 2000b).

132.31 Les forces d'agglomération

Du côté des forces d'agglomération, on trouve d'abord les interactions de marché. Ces dernières recouvrent deux aspects. D'une part, la structure de concurrence monopolistique joue en faveur de la concentration spatiale des activités puisque la différenciation des produits permet d'affaiblir la concurrence en prix et que les consommateurs, ayant une préférence pour la variété, sont attirés par ces concentrations. D'autre part, l'existence de relations verticales entre les firmes donne lieu à des mouvements de localisation en cascades. En effet, une localisation comprenant de nombreuses firmes attire des firmes produisant des biens intermédiaires. En retour, ceci a pour conséquence d'en faire une localisation attractive pour les firmes qui utilisent ces biens intermédiaires. Cette interdépendance entre les décisions de localisation des firmes amont et aval avait déjà été mise en avant dans les années 50 par Myrdal (1957) ou Hirschman (1958).

Les interactions hors marché, ou externalités, jouent également un rôle important d'agglomération lorsqu'elles sont positives. Elles se réfèrent surtout aux externalités de communication et aux externalités de capital humain. Dans le premier cas, les agents ont intérêt à être géographiquement concentrés pour pouvoir échanger directement des informations et éviter le phénomène de détérioration de la transmission d'information lorsque le nombre d'intermédiaires est important (Banerjee, 1993 ; Baumont *et al.*, 1998 ; Thisse et Van Ypersele, 1999). Dans le second cas, les analyses proposées par Lucas (1988) ou Rauch (1993) suggèrent que le regroupement d'individus a pour conséquence d'augmenter le capital humain d'une ville, ce qui permet d'augmenter la productivité du travail et la qualité du capital humain.

Enfin, les grandes villes, de par la diversité des firmes et des travailleurs, permettent d'améliorer l'appariement sur les marchés de l'emploi et des produits (Thisse et Van Ypersele, 1999 ; Montfort et Ottaviano, 2000 ; Combes et Duranton, 2001). En effet, les entreprises trouvent plus facilement une main d'œuvre qualifiée qui correspond à leurs

besoins si le marché du travail est diversifié. De même, les travailleurs ont plus de chances de trouver un emploi qui correspond à leur qualification lorsque les firmes sont diversifiées (Thisse et Zénou, 1995). Les ménages profitent également de cette diversité dans la mesure où ils ont à leur disposition une plus grande variété de biens, de services et de loisirs. On peut encore étendre ces avantages à l'appariement sur les marchés de biens ou de services à la production : plus les firmes avales sont nombreuses et plus les fournisseurs sont incités à se développer. De même, plus les fournisseurs sont nombreux et plus les firmes avales ont une garantie d'être servies. Les espaces urbains sont alors sur ces nombreux points considérés comme des « assurances tous risques » dans les choix de localisation (Baumont et Huriot, 1995).

132.32 Les forces de dispersion

Du côté des forces de dispersion, on peut déjà citer le rôle des facteurs immobiles. En effet, lorsque les individus sont immobiles, le facteur travail est dispersé et ceci encourage une localisation dispersée des firmes pour des raisons d'offre : les firmes utilisant le facteur travail doivent se localiser près de la main d'œuvre, et pour des raisons de demande : il faut livrer à ces individus immobiles les biens produits. Dans le modèle de formation des agglomérations de Krugman (1991a), le secteur agricole est immobile et c'est cet élément qui joue comme une force de rappel et qui limite l'extension de l'agglomération. D'une façon générale, on peut montrer que tout obstacle à la mobilité géographique des travailleurs constitue un frein à la concentration géographique des firmes (Puga, 1999 ; Monfort et Nicolini, 2000). Cette force de rappel peut encore être constituée par la présence de biens intransportables, comme les infrastructures publiques (Walz, 1996) ou les logements (Helpman, 1996).

Au niveau urbain, on invoque souvent la hausse des coûts fonciers, résultant de l'attrait que les grandes villes exercent sur les agents économiques, consommateurs ou producteurs. L'agglomération est ainsi freinée par un accroissement de la concurrence pour l'occupation de sol qui, toutes choses égales par ailleurs, engendre une hausse des prix du sol. A partir d'un certain niveau, ces coûts ont un effet répulsif et découragent l'installation de nouveaux agents tout en favorisant l'étalement urbain et le développement de la périurbanisation. Les externalités négatives sont également souvent évoquées : l'agglomération des activités et des hommes ayant pour effet d'augmenter les problèmes de congestion des voies de communication, les problèmes de pollution ou de criminalité (Glaeser, 1994 ; Brakman *et al.*, 1996 ; Tabuchi et Thisse, 2002).

Le rôle des coûts de transport¹⁶ mérite une attention particulière. Des coûts de transport élevés, signes d'une forme de tyrannie de la distance, pour reprendre l'expression de Duranton (1997), constituent une force importante de dispersion. Ainsi, selon Bairoch (1985), le transport des biens agricoles vers les aires urbaines a limité le développement des villes

¹⁶ ou plus généralement aujourd'hui l'ensemble des coûts de réalisation des échanges pouvant être désignés sous le terme de coûts d'interaction.

dans les pays industrialisés et ce jusqu'à la révolution industrielle. On peut également considérer que le transport des biens industriels freine le développement de l'agglomération. De même, la réalisation d'une relation de service nécessite souvent la rencontre entre le prestataire et le client. S'il existe une demande immobile à l'extérieur des villes, les entreprises, ou les sociétés, ont tendance à se rapprocher de cette demande pour pouvoir la servir à un coût plus bas. Enfin, les coûts de transport intra-urbain des personnes pour travailler ou effectuer des achats représentent le fondement des modèles d'économie urbaine et constituent aujourd'hui un frein à l'urbanisation.

Le rôle des coûts d'interaction a évolué dans le temps et, si on assiste aujourd'hui à une diminution des seuls coûts monétaires de transport, les coûts de transaction et les coûts temporels restent importants. Les coûts de transport des marchandises ou de déplacement des personnes restent fortement conditionnés par l'accessibilité et donc par l'état de développement des réseaux de transport et leur facilité d'accès. En revanche, les coûts de transaction sont plutôt dépendants de la proximité géographique et l'agglomération constitue un moyen de les diminuer (Cappellin, 1988).

Dans les modèles d'économie géographique, la distribution des activités économiques est largement gouvernée par l'arbitrage entre les coûts de transport et les rendements d'échelle croissants (Thisse et Van Ypersele, 1999). La tendance à la diminution des coûts de transport des marchandises n'a pas conduit à une dispersion des activités. En effet, si les coûts de transport deviennent négligeables, alors les firmes ne sont pas obligées de se rapprocher de la demande immobile ou des ressources naturelles. Cependant, cette baisse a également pour effet d'intensifier la concurrence en réduisant le pouvoir de monopole des entreprises fondé sur l'éloignement géographique : la distance n'est plus une protection contre la concurrence. Pourtant, si les entreprises choisissent de différencier leurs produits, elles ne craignent plus la concurrence en prix et préfèrent s'agglomérer pour profiter des avantages liés à cette concentration (bassins de consommation plus importants, proximité aux autres firmes...) qui réduisent les coûts de transaction. Ainsi, contrairement aux résultats attendus intuitivement, la baisse des coûts de transport permet à d'autres forces d'agglomération de s'exprimer.

132.33 Les équilibres géographiques

Cette présentation rapide montre que les forces d'agglomération ou de dispersion engagent de multiples facteurs dans des combinaisons qui peuvent être complexes. Il faut ainsi souligner que c'est la valeur de ces facteurs, support de ces forces, qui détermine le sens des forces. Par exemple, des coûts de transport élevés agissent comme force de dispersion. En revanche, s'ils sont faibles, ils favorisent l'agglomération. Il faut, par ailleurs, retenir que ces facteurs interagissent au sein d'arbitrages : arbitrage entre coûts de transport et coûts fonciers, arbitrage entre concurrence en prix et taille de marché ... Une fois définis les facteurs support des forces d'agglomération et de dispersion, la formation des agglomérations apparaît alors comme le résultat d'un équilibre entre ces différentes forces. Pour limiter les problèmes mathématiques liés à la résolution des modèles et à l'interprétation de leurs résultats, il n'est généralement pas possible d'intégrer toutes ces forces simultanément. Les modèles

d'économie géographique sont alors amenés à combiner au moins l'une des forces d'agglomération et l'une des forces de dispersion (Krugman, 1999). Par exemple, le modèle de base défini par Krugman (1991a, 1991b) considère comme facteurs déterminants : la valeur des coûts de transport des biens industriels, la part de la dépense industrielle des ménages et le degré de concurrence monopolistique.

Les équilibres géographiques qui émergent de ces modélisations sont généralement de deux types : il s'agit soit d'équilibres d'équi-répartition où les activités économiques sont réparties également entre les différentes régions, soit d'équilibres centre-périphérie où les activités économiques sont concentrées dans une seule région. Les équilibres géographiques obtenus sont très dépendants des valeurs des paramètres choisis comme support des forces de dispersion et de concentration. On montre qu'une variation de ces valeurs conduit à des changements discontinus dans la configuration géographique d'équilibre. En d'autres termes, il existe des points de bifurcation à partir desquels on passe d'un type d'équilibre géographique à un autre. On montre en particulier que l'équilibre de dispersion est souvent de nature instable et se « transforme » en équilibre d'agglomération (Papageorgiou et Smith, 1983). Par ailleurs, ce sont des petites différences initiales qui déterminent quelle localisation va héberger une agglomération plutôt qu'une autre. Cependant, dès qu'une localisation est devenue un centre d'activité, un phénomène de « verrouillage » opère le plus souvent. Même si l'attractivité d'un site diminue, suite à une modification des circonstances exogènes, les agents économiques restent localisés au même endroit pour continuer de bénéficier des avantages de l'agglomération. En conséquence, l'histoire joue un rôle (Arthur, 1989 ; Krugman, 1991c).

Ces différents développements ont mis en lumière le rôle joué par les coûts de transport dans la formation et la persistance des agglomérations. Dans le contexte actuel d'intégration économique au sein de l'Union Européenne, il est intéressant de savoir quelles vont être les conséquences, en termes de localisation des activités économiques, de la baisse des coûts de transport et d'interaction. Plus généralement, les différents impacts de l'intégration européenne sur les disparités régionales peuvent être étudiés dans le cadre des modèles d'équilibre interrégionaux initiés par Krugman (1991a, 1991b).

133 Schéma centre-périphérie et intégration économique

Les modèles d'économie géographique, issus du modèle d'équilibre interrégional de Krugman (1991a, 1991b), offrent un cadre analytique propice à la description d'une économie régionale en situation d'intégration. En effet, dans ces modèles, plusieurs éléments permettant de décrire les effets de l'intégration économique sur les interactions entre les économies sont pris en compte. Ces éléments sont les coûts de transport des marchandises, ou plus généralement les coûts liés aux échanges de marchandises, la taille du marché, la structure

productive, le degré de mobilité des travailleurs, du capital ou des innovations ou encore la structure de la demande... Ces modèles peuvent en particulier être utilisés pour étudier l'impact de l'intégration économique sur le développement économique régional (Ottaviano et Puga, 1997, 1998 ; Baumont, 1998a, 1998b ; Walz, 1998 ; Puga, 2001) ou pour comparer le fonctionnement du système économique « régional » aux Etats-Unis et en Europe (Puga, 1999). Bien sûr, ces modèles ne sont pas les seuls capables de décrire des systèmes économiques intégrés. Ils le font dans le cadre d'hypothèses qui est le leur en s'intéressant aux équilibres de répartition des activités économiques entre les régions.

Ainsi, les modèles d'économie géographique ont pour objectif de montrer comment deux régions peuvent se différencier, de manière endogène, entre un centre industrialisé et une périphérie désindustrialisée. Dans le modèle de Krugman (1991a, 1991b), les mécanismes permettant d'aboutir à ce résultat sont liés à la présence d'externalités pécuniaires et de causalités circulaires et cumulatives trouvant leur source dans l'interconnexion entre production et consommation finale (paragraphe 133.1). Le schéma centre-périphérie, qui émerge notamment à la suite d'une baisse des coûts de transport, reste par ailleurs un résultat fort des modèles d'économie géographique développés à la suite du modèle de base de Krugman, même lorsque les forces d'agglomération sont atténuées ou modifiées (paragraphe 133.2).

133.1 L'équilibre centre-périphérie

Le modèle de Krugman (1991a, 1991b) est constitué de deux régions, de deux secteurs : le secteur des biens industriels et celui du bien agricole, et de deux types de travailleurs.

Du côté de l'offre, deux types de biens sont produits.

D'une part, le secteur agricole produit un bien homogène au moyen d'une technologie à rendements constants. Ce secteur n'emploie qu'un seul type de travailleurs, la main d'œuvre agricole, supposée immobile. Les biens agricoles s'échangent sans coûts de transport.

D'autre part, le secteur industriel produit des biens différenciés au moyen d'une technologie à rendements croissants et dans une situation de concurrence monopolistique. Ce secteur emploie l'ensemble des ouvriers de l'économie et ces derniers sont parfaitement mobiles. Les biens industriels s'exportent d'une région vers l'autre avec coût de transport qui prend une forme dite de l'*iceberg*, introduite par Samuelson (1954). On suppose dans ce cas qu'une partie du produit « fond » au cours du transport. En conséquence, chaque fois qu'un bien est exporté d'une région à une autre, une partie seulement de ce bien arrive à destination.

Du côté de la demande, on suppose que tous les ménages ont la même fonction d'utilité. Elle est de type Cobb-Douglas entre le bien agricole et les biens industriels. La préférence entre ces derniers est modélisée par une fonction de type CES à la Dixit et Stiglitz

(1977) et permet d'incorporer une préférence pour la variété : l'utilité du consommateur augmente avec le nombre de biens industriels.

Dans ce modèle, l'immobilité des travailleurs agricoles est la principale force de dispersion car ils consomment les deux types de biens. Certaines entreprises ont donc intérêt à rester proche de ces travailleurs afin de satisfaire leur demande sans supporter de coûts de transport sur les biens industriels. La concurrence en prix constitue également une force de dispersion. En effet, elle est plus importante lorsque plusieurs entreprises sont concentrées dans le même lieu. Cependant, si la différenciation des produits est suffisamment importante, alors la concurrence en prix et la force de dispersion associée à cet effet sont affaiblies.

Les forces d'agglomération sont plus difficiles à mettre en évidence car elles sont le résultat d'un jeu complexe de mécanismes économiques impliquant différents éléments (Fujita et Thisse, 1997) :

- L'effet d'entraînement amont ou *backward linkage* : puisque les rendements sont croissants, les producteurs sont incités à se localiser près des consommateurs pour exploiter ces rendements et satisfaire une large demande. En outre, les consommateurs, ayant une préférence pour la variété, sont également incités à se localiser près des producteurs afin d'avoir accès à un plus grand nombre de biens sans supporter les coûts de transport de ces biens. Ces deux effets donnent lieu à des externalités pécuniaires et à un effet « taille de marché ».
- L'effet d'entraînement aval ou *forward linkage* : les ménages et les entreprises sont également liés par le marché du travail. Lorsque de nombreux ouvriers sont localisés dans une région, le nombre de produits régionaux est plus grand et les prix d'équilibre sont plus bas, du fait de la plus grande concurrence entre les firmes agglomérées. Ceci entraîne en effet de revenu réel pour les consommateurs, qui sont alors attirés dans cette région. Il se produit donc un effet « d'indice des prix ».

Ces deux effets sont à l'origine d'une causalité circulaire et cumulative reliant les processus d'agglomération des producteurs et des travailleurs, comme l'illustre la figure 1.9.

L'intensité relative des forces d'agglomération et de dispersion conduit à l'émergence de différents types de configurations spatiales d'équilibre caractérisées par l'agglomération ou la dispersion des activités économiques. L'équilibre interrégional de l'économie est obtenu par la détermination des quantités et des prix d'équilibre sur les marchés des biens régionaux, de la répartition de la main d'œuvre ouvrière et des firmes entre les deux régions et des salaires d'équilibre sur le marché du travail. Les interactions décrites précédemment sont modélisées à l'aide de différents facteurs dont l'importance et l'ampleur sont traduites par des paramètres tels que le degré de préférence pour la variété, le coût de transport *iceberg*, la part de la main d'œuvre industrielle ou encore la part de la dépense en bien industriel.

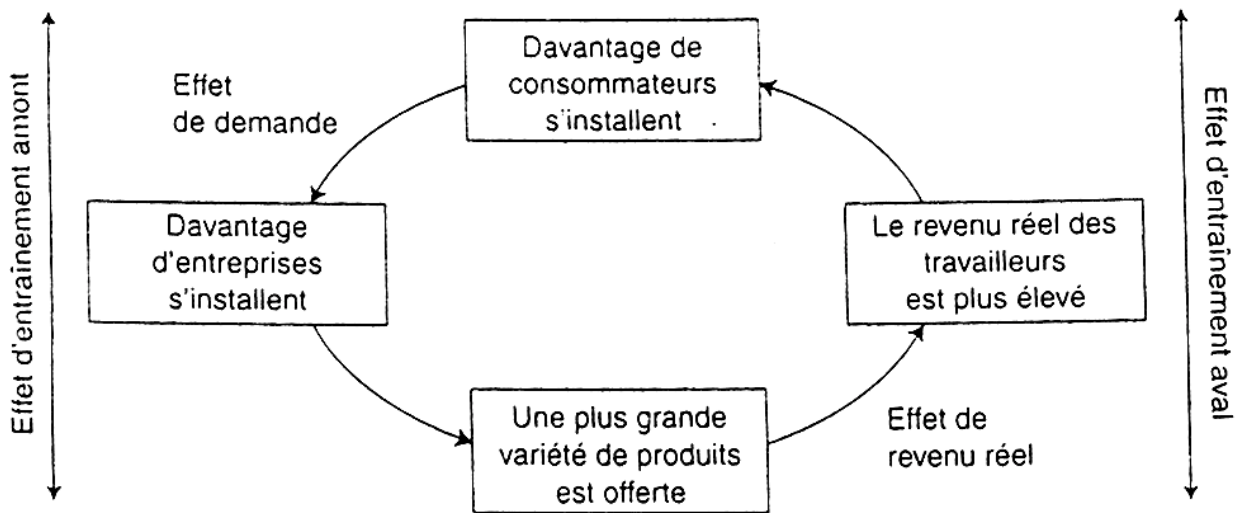


Figure 1.9 : Causalité circulaire dans la formation d'une agglomération d'entreprises et de travailleurs

Source : Fujita et Thisse (1997), p. 59

Compte tenu de la complexité de ces interactions et du nombre élevé de paramètres considérés, l'équilibre interrégional de l'économie ne peut pas être dérivé analytiquement. Krugman est donc amené à procéder par simulations en fixant la valeur de certains paramètres. Il montre que les différents mécanismes décrits précédemment sont susceptibles d'aboutir à long terme sur une configuration de type centre-périphérie. En particulier, cette dernière est favorisée lorsque les biens industriels sont suffisamment différenciés afin de diminuer la concurrence en prix, lorsque la part du secteur industriel est important et/ou lorsque le coût de transport des biens industriels est relativement faible.

La stabilité des équilibres géographiques ainsi obtenus est également étudiée et elle dépend de l'importance respective des forces centrifuges et des forces centripètes. Le coût de transport sur les biens industriels joue à cet égard un rôle particulièrement important. Supposons, toutes choses égales par ailleurs, que l'industrie se répartisse de manière égale entre les deux pays et que les coûts de transport soient très élevés. Dans ce cas, une diminution modérée de ces coûts ne modifie pas les décisions de localisation des firmes qui préfèrent conserver une répartition régionale égalitaire car cela permet de répondre à la demande finale de manière locale. Lorsque les coûts de transport tombent en dessous d'une valeur critique, une relocalisation d'une firme dans l'une des régions augmente les profits des firmes dans cette région et diminue celle des firmes localisées dans l'autre région. On observe alors un mouvement auto-renforçant conduisant à une différenciation endogène de deux

régions *a priori* identiques en une région centre qui possède toute l'industrie et une région périphérique désindustrialisée.

Plus généralement, il apparaît que le comportement qualitatif de ce modèle change brusquement lorsque les valeurs des paramètres clés passent un niveau critique : les équilibres de localisation et leur stabilité sont ainsi très sensibles aux conditions initiales (Krugman, 1999) : la région qui possède un léger avantage en profite. Toute différence, même petite, dans les conditions initiales est susceptible de conduire vers des schémas de développement inégaux qui sont stables. En outre, ce sont ces conditions initiales, géographiques ou économiques, qui déterminent les localisations où auront lieu l'agglomération des activités : « le temps réapparaît : la géographie économique a une histoire en ce sens que les conditions initiales sont essentielles pour la sélection d'un équilibre » (Fujita et Thisse, 1997, p. 76).

L'une des critiques portées à ce modèle de base concerne son incapacité à produire des résultats sous forme analytique. Cependant, Ottaviano *et al.* (2002) confirment les résultats précédents en utilisant une spécification alternative pour la demande qui permet l'obtention de résultats analytiques. Ils trouvent ainsi une configuration centre-périphérie dans des conditions similaires à celles identifiées par Krugman (1991a, 1991b). Il apparaît alors que les principales tendances vers l'agglomération sont valides pour une classe de modèles plus large et ne dépendent plus d'une spécification pour les préférences.

133.2 D'un modèle d'agglomération à l'autre

Le modèle de Krugman (1991a, 1991b) a donné lieu à de nombreuses extensions dans lesquelles le schéma centre-périphérie apparaît être une caractéristique principale en présence d'intégration régionale.

1) Développement des infrastructures de transport et échanges

Martin et Rogers (1995) introduisent le rôle des infrastructures qui permettent de faire diminuer les coûts de transport. Ceux-ci sont considérés sous deux angles puisqu'il s'agit à la fois des coûts de transport intra- et interrégionaux. Dans ces conditions, la localisation des entreprises dépend, en partie, des dotations en infrastructure des régions. Il apparaît ainsi qu'au cours de l'intégration économique, les firmes du secteur à rendements croissants se localisent dans la région la mieux dotée en infrastructures pour pouvoir profiter d'une demande relative plus élevée pour les biens domestiques dont le prix est plus faible.

En procédant par simulation, les auteurs montrent alors que le développement d'infrastructures dans toutes les régions d'une façon symétrique a pour effet de renforcer les écarts de développement et le processus d'agglomération. En effet, l'amélioration des infrastructures interrégionales, en augmentant le commerce interrégional, permet aux entreprises de la région initialement favorisée (celle dont les dotations en capital sont les plus élevées ou celle dont la taille est la plus importante) de rester localisées dans cette région en

servant à moindre coût la région plus défavorisée. Dans une perspective d'aménagement du territoire, la possibilité d'une convergence, en termes de structures industrielles, de la région périphérique avec la région centrale doit donc plutôt passer par le développement d'infrastructures permettant de faciliter le commerce intra-régional. Dans ce cas, en effet, la demande domestique pour les biens produits dans la région périphérique est accrue, incitant les firmes à se relocaliser dans cette région.

Dans ces conditions, la politique de la CEE, visant à accélérer le développement des régions périphériques en finançant les infrastructures serait non seulement inefficace, mais elle aurait en plus les effets contraires à ceux escomptés en favorisant encore plus les forces d'agglomération dans les régions initialement les plus riches. Ainsi, d'après Faini (1983), la réduction des coûts de transport entre le nord et le sud de l'Italie dans les années 50 a privé les firmes du Mezzogiorno d'une protection contre la concurrence, accélérant ainsi la désindustrialisation de cette région.

2) Développement des infrastructures publiques et productivité des firmes

Charlot (1999) s'intéresse au rôle des infrastructures publiques en exploitant les effets de complémentarité entre le secteur public et le secteur privé. Plus précisément, au lieu d'influencer les conditions d'échanges intra- et interrégional, l'auteur considère que le développement des infrastructures publiques est bénéfique à la productivité des firmes. Celles-ci sont financées par un impôt national et réparties par les instances publiques entre les différentes régions de façon exogène. Elles participent de deux manières différentes à l'amélioration de la productivité des firmes : soit en diminuant le coût variable des entreprises locales (exprimé en termes de quantité de travail nécessaire à la production), soit en réduisant le coût fixe en favorisant l'environnement général de production. Les hypothèses sur les secteurs agricole et industriel sont, pour le reste, identiques à celles du modèle de Krugman (1991a).

Au contraire du modèle de Martin et Rogers (1995), dans lequel une politique de développement des infrastructures peut avoir des effets ambigus, l'accroissement des investissements publics dans une région augmente toujours, dans ce modèle, l'attractivité de cette région. Les principaux résultats obtenus montrent que les investissements publics jouant sur le coût variable de production ont un impact plus important sur les équilibres de localisation que les investissements publics affectant le coût fixe de production. Par ailleurs, les simulations réalisées indiquent que lorsque les coûts de transport sont élevés ou la préférence pour la variété est faible, les investissements publics ont la possibilité de réduire les déséquilibres entre les régions. En revanche, lorsque les coûts de transport diminuent, l'intervention publique, *via* le financement d'infrastructures, ne peut plus jouer son rôle de rééquilibrage en contrecarrant les processus cumulatifs conduisant à l'agglomération des activités. Les seuls équilibres géographiques stables obtenus dans ce cas sont les équilibres centre-périphérie.

On retrouve ici le caractère dominant de l'équilibre centre-périphérie. L'amélioration de la productivité des firmes constitue une force d'agglomération supplémentaire alors que le frein potentiel lié à la politique fiscale n'est pas très important dans ce modèle.

3) *Mobilité des firmes et relations intersectorielles*

La causalité cumulative dans le modèle de Krugman (1991a) repose sur l'hypothèse que, lorsqu'une région attire des firmes, elle est également capable d'attirer des travailleurs par des salaires plus élevés et un plus grand choix de produits. Cependant, en Europe, les migrations interrégionales apparaissent être relativement faibles en comparaison de celles observées aux Etats-Unis (Bentivogli et Pagano, 1999). Dans ces conditions, le mécanisme précédemment décrit ne tient plus. D'autres modèles d'économie géographique introduisent alors les liaisons amont-aval entre industries comme nouvelle force d'agglomération (Krugman et Venables, 1995, 1996 ; Venables, 1996).

Dans le modèle de Krugman et Venables (1995) par exemple, le secteur industriel produit à la fois le bien final, utilisé pour la consommation finale, et des biens intermédiaires, utilisés comme inputs dans la production d'autres biens industriels. Au contraire du modèle de Krugman (1991a), les auteurs supposent que la main d'œuvre industrielle est immobile entre les deux régions, mais qu'elle est mobile entre les secteurs. Par conséquent, le processus de causalité cumulative a lieu car les entreprises ayant besoin de biens intermédiaires diversifiés se localisent près d'autres entreprises, pour minimiser les coûts de production du bien final (*forward linkage*), ce qui a pour conséquence d'augmenter la demande de biens intermédiaires et en retour, d'attirer encore d'autres firmes dans cette région (*backward linkage*). Suivant la valeur des coûts de transport, trois résultats émergent. Pour des coûts de transport importants, les firmes sont réparties entre les deux pays. Lorsque les coûts de transport tombent en dessous d'une valeur critique, un schéma de type centre-périphérie se forme spontanément. Cependant, une baisse encore plus accentuée de ces coûts a pour conséquence de favoriser la relocalisation d'un certain nombre de firmes dans la région périphérique et conduit à un processus de convergence des structures industrielles régionales. L'intégration économique a ainsi pour effet de mettre en évidence une courbe en U pour le développement. Cette conclusion se retrouve dans le modèle de Venables (1996) avec un continuum de secteurs en concurrence imparfaite.

En Europe, une agglomération des activités à rendements croissants a bien eu lieu selon un schéma de type centre-périphérie durant les premières étapes de l'intégration économique (Brülhart et Torstensson, 1996) : les industries caractérisées par de fortes économies d'échelles étaient très concentrées au début de l'intégration européenne mais sont aujourd'hui mieux réparties sur le territoire. Ceci confirme la relation en U entre intégration et polarisation.

4) Mobilité géographique et mobilité intersectorielle des travailleurs

Nombre d'industries régionales ne sont pas concentrées dans un pays particulier, ce qui laisserait supposer que les migrations interrégionales et internationales suivent des logiques différentes. L'intégration régionale européenne ne peut donc s'examiner de la même façon que l'intégration économique au sein d'un système homogène tel que le système des états des Etats-Unis. Pour rendre compte de ces spécificités, plusieurs auteurs ont cherché à modéliser plus finement la mobilité de la main d'œuvre.

Ainsi, Monfort et Nicolini (2000) construisent un modèle à deux pays et quatre régions dans lequel chaque pays est constitué de deux régions. Le travail entre les régions d'un même pays est mobile alors qu'il est immobile entre régions de pays différents. Les coûts de transport sont également plus ou moins élevés selon que les échanges s'effectuent entre régions d'un même pays et régions de pays différents. Comme dans le modèle de Krugman (1991a), il n'y a pas de mobilité intersectorielle du travail. Les résultats de cette analyse suggèrent que l'agglomération des économies nationales est intensifiée par le processus d'intégration internationale. Par conséquent, l'intégration économique dans l'Union Européenne est susceptible d'augmenter les disparités régionales à l'intérieur des pays européens.

Une autre analyse, proposée par Puga (1999) consiste à intégrer dans un cadre unifié le modèle de Krugman (1991a) et celui de Krugman et Venables (1995) : il considère un modèle dans lequel il existe à la fois des migrations interrégionales de la main d'œuvre et des relations intersectorielles entre les firmes qui poussent à l'agglomération. L'auteur montre alors que, dans ce cadre unifié, les migrations interrégionales ont pour conséquence d'accélérer le processus d'agglomération. En revanche, si les travailleurs sont immobiles, les écarts de salaires persistent et agissent comme une force de dispersion en augmentant les coûts de production pour les firmes produisant dans la localisation où les salaires sont les plus élevés. Ceci modère l'agglomération et il apparaît des équilibres dans lesquels toutes les régions possèdent une industrie, même dans des proportions différentes. Ainsi, l'absence de migrations interrégionales retarde le processus d'agglomération pendant le processus d'intégration régionale et l'affaiblit lorsqu'il a lieu. L'auteur montre également que l'absence de mobilité interrégionale rend la relation entre coûts de transport et agglomération non-monotone. Pour des coûts de transport élevés, les firmes sont réparties également dans les deux régions pour servir la demande finale. Pour des coûts de transport intermédiaires, les disparités régionales augmentent puisque l'une des régions attire plus d'industrie que l'autre mais sans que cela aboutisse à une situation dans laquelle toute l'industrie est concentrée dans l'une des régions. Enfin, pour des coûts de transport faibles, l'agglomération diminue à nouveau pendant que la part de l'industrie dans la région avec des salaires plus faibles augmente graduellement jusqu'à un équilibre symétrique stable. Ce modèle suggère ainsi qu'en présence de freins à la mobilité interrégionale du travail, l'intégration économique peut conduire à un équilibre des structures productives entre les régions, même si, durant les

périodes intermédiaires du processus d'intégration, il peut exister d'importantes disparités interrégionales des salaires réels.

Ces résultats sont à relier aux réglementations visant à assouplir les conditions de la mobilité géographique du travail, notamment par l'harmonisation des niveaux de qualification. En outre, la relation de développement en U trouvée entre agglomération et intégration peut être interprétée de la façon suivante (Puga, 1999, 2001) : en raison d'une faible mobilité du travail, l'intégration européenne, à terme, peut conduire en elle-même à une convergence régionale en termes de salaires réels et de structure de production. Pourtant, le rattrapage des régions riches par les régions pauvres doit, dans ce cas, passer par une intégration suffisante et par des salaires flexibles aux changements d'emplois industriels. Concernant ce dernier point, Puga (1999) suggère que si les différences de salaires reflètent les conditions locales entre les pays et non les conditions locales entre les régions, l'intégration européenne pourrait avoir pour effet d'accentuer encore l'agglomération des activités et l'augmentation des disparités de revenu entre les régions d'un même pays¹⁷.

Ces différents éléments suggèrent ainsi que l'intégration économique, sous la forme de la diminution des coûts d'interaction (coûts de transport, coûts de transaction...) a pour conséquence de favoriser l'agglomération des activités et ainsi de renforcer les disparités régionales en termes de revenus. Ce constat, défavorable pour les politiques visant à renforcer l'intégration régionale, a été en partie imputé au fait que les forces de concentration sont surtout prises en compte et qu'en contrepartie des forces de dispersion sont négligées.

5) Des forces de dispersion à prendre en compte

Les modèles d'économie géographique présentés jusqu'à présent font abstraction d'un certain nombre de forces de dispersion susceptibles de remettre en cause, ou d'atténuer, l'émergence d'un schéma centre-périphérie. Nous présentons ici rapidement trois extensions du modèle de Krugman (1991a, 1991b).

Calmette et Le Pottier (1995) introduisent un coût de transport, de type *iceberg*, supporté par le consommateur, sur le bien homogène produit dans une région et consommé dans l'autre. Dans ce cas, l'équilibre géographique d'agglomération se produit moins facilement que dans le cas où ces coûts sont supposés nuls.

Helpman (1996) introduit la consommation de logement dans le but de prendre en compte le besoin d'espace nécessaire au bien-être des agents. Ce bien est intransportable et agit comme un bien dont le coût de transport est infiniment grand. Dès lors, c'est l'équilibre d'équi-répartition qui apparaît le plus souvent comme un schéma d'équilibre.

On peut enfin introduire dans l'analyse le fait que les agglomérations doivent supporter des coûts de fonctionnement, comme les coûts fonciers, les coûts de congestion, les coûts de pollution ou les coûts d'insécurité. Ces coûts, qui augmentent avec le nombre de

¹⁷ D'autres modèles, développés en particulier par Puga et Venables (1997, 1998, 1999) tendent pourtant à montrer que la mise en place d'accords multilatéraux est le meilleur moyen d'aboutir à une convergence des structures industrielles régionales (Ghio, 1999a).

firmes ou d'individus présents, réduisent alors l'attractivité des agglomérations et peuvent conduire à la relocalisation des entreprises vers les régions pauvres. Le modèle de Brakman *et al.* (1996) prend en compte ces phénomènes de congestion en supposant que les coûts de production dans le secteur à rendements croissants ne sont pas constants mais qu'ils sont une fonction croissante du nombre de firmes localisées dans la région. La prise en compte de ces nouveaux mécanismes apporte un élément de complexité supplémentaire. En fixant une congestion affectant à la fois le coût variable et le coût fixe de production, les résultats obtenus par simulation montrent que l'équilibre d'équi-répartition est le seul équilibre stable pour des coûts de transport très élevés ou très faibles alors que pour des coûts de transport intermédiaires, les équilibres sont multiples.

Ce tour d'horizon de quelques modèles d'économie géographique nous a permis de présenter les mécanismes fondamentaux de formation des agglomérations qui sont considérés dans le cadre du modèle de Krugman (1991a, 1991b). Il apparaît ainsi qu'il existe des forces de dispersion « pures » liées à la présence de facteurs de production ou de biens immobiliers ou à des phénomènes de congestion. Par ailleurs, il existe des configurations de facteurs qui agissent, soit comme forces de dispersion, soit comme forces d'agglomération, suivant les valeurs et les combinaisons de ces différents éléments. Ainsi en est-il de la combinaison entre les trois paramètres clés du modèle de Krugman (1991a, 1991b) : la préférence pour la variété, la part de la dépense en biens industriels et le coût de transport des biens industriels. Des effets de seuils, liés aux valeurs de ces paramètres, existent pour lesquels l'arbitrage entre les coûts d'accessibilité aux marchés régionaux et les avantages de la proximité conduit, soit à l'expression de forces nettes d'agglomération, soit à l'expression de forces nettes de dispersion. On montre néanmoins, compte tenu des valeurs des paramètres caractérisant les systèmes économiques contemporains (faible coût de transport, forte part de la dépense en biens industriels et forte préférence pour la variété), que les effets « de taille de marché » et « d'indice de prix » jouent en faveur de l'agglomération tandis que les coûts d'accessibilité ne sont pas suffisamment forts.

Tous ces travaux supposent néanmoins que le « stock » de firmes et/ou d'agents est donné. Il est alors intéressant de considérer les implications que peuvent avoir, sur la répartition des activités entre les régions, un phénomène de croissance. Ceci est réalisé à l'aide des modèles qui effectuent une « synthèse géographie-croissance » (Baumont, 1998a ; Baumont et Huriot, 1999).

134 La synthèse géographie-croissance

L'agglomération spatiale des activités dans les villes et la croissance économique sont des phénomènes concomitants (Bairoch, 1985 ; Hohenberg et Lees, 1985). Ainsi, durant la seconde phase de la révolution industrielle en Europe, le taux de croissance a fortement

augmenté en même temps que le taux d'urbanisation s'est fortement accru et que des centres industriels se formaient au cœur de l'Europe. En outre, la ville est souvent le lieu d'agglomération spatiale des activités porteuses de croissance, comme les services aux producteurs, le secteur de la R&D, les activités tertiaires supérieures, etc. Il semble alors que la croissance économique et l'agglomération spatiale des activités sont des processus se renforçant mutuellement.

Les théories de l'économie géographique et celles de la croissance endogène ont récemment été intégrées afin de mieux comprendre les relations qui lient croissance et agglomération. Ce courant, dénommé « synthèse géographie-croissance » (Baumont et Huriot, 1999) découle de la prise en compte d'un certain nombre de points communs entre les processus d'agglomération spatiale des activités économiques et les processus d'accumulation temporelle des facteurs de croissance (paragraphe 134.1). Dans ces conditions, il est possible de mettre en lumière la façon dont la croissance modifie les équilibres spatiaux atteints dans le cadre d'une économie statique. En particulier, la nature locale ou globale des externalités de connaissance joue un rôle fondamental (paragraphe 134.2). On peut enfin montrer que les processus d'agglomération et les mécanismes de croissance endogène s'influencent et se renforcent mutuellement (paragraphe 134.3).

134.1 Les fondements de la synthèse

Le rapprochement entre les théories de la croissance endogène et les théories de l'économie géographique est important pour deux raisons.

D'une part, ce rapprochement permet d'intégrer une dimension temporelle aux modèles d'économie géographique, qui sont essentiellement des modèles de statique comparative. Par exemple, les effets de l'intégration économique sur les schémas de localisation sont étudiés à partir d'une variation paramétrique des coûts de transport. Comme le soulignent Baumont et Huriot (1999, p. 144) : « *Les modèles d'agglomération de type "centre-périphérie" expliquent seulement comment, en fonction des perturbations initiales, on est passé d'un équilibre "dispersé" à un équilibre "aggloméré". Les phases intermédiaires ne sont pas décrites et n'ont pas d'intérêt en soi. La "dynamique cumulative" n'est qu'une technique d'ajustement et elle n'a pas de signification historique. Seul l'équilibre peut être caractérisé et relié aux valeurs des paramètres clés* ». Dans les modèles d'économie géographique étudiés précédemment, le nombre de firmes, c'est-à-dire celui de la quantité de main d'œuvre disponible, est fixé. Introduire une dynamique de croissance permet au contraire d'étudier les modalités de localisation ou de délocalisation des activités lorsque le nombre de firme augmente

D'autre part, les théories de la croissance endogène sont le plus souvent basées sur l'existence d'externalités positives sur la production et sur la connaissance. Dans ces modèles,

on suppose que ces externalités sont parfaitement mobiles entre les industries, entre les secteurs et entre les différentes économies. En conséquence, ces externalités sont dénuées de toute dimension géographique. Dans le cas des externalités de connaissance, cela implique que, quelles que soient les localisations des firmes innovantes, ces dernières bénéficient d'un stock de connaissances d'une manière symétrique. En conséquence, il n'existe pas de bénéfices à être localisé près d'une source d'externalités. Les externalités sont alors qualifiées de globales, c'est-à-dire que les agents en bénéficient quelles que soient leurs localisations. Pourtant, Lucas (1988) relève que les effets externes à l'origine de la croissance endogène sont principalement spatialisés, les externalités ayant un caractère local et non global : plus un agent est à proximité de la source des externalités, plus il en bénéficie. Lucas (1988) reconnaît alors dans la ville le niveau spatial auquel se produisent les différentes interactions, source de croissance endogène. Par ailleurs, une littérature empirique importante suggère que les externalités informationnelles ont une diffusion limitée dans l'espace, les entreprises fortement consommatrices d'informations ayant intérêt à se localiser à proximité les unes des autres afin d'obtenir des informations (Guillain, 2000 ; Guillain et Huriot, 2001). Ces études tendent à renforcer l'idée que les externalités sont locales et non globales.

Dans ces conditions, le rapprochement des théories de l'économie géographique et des théories de la croissance endogène permet, en modélisant simultanément la création de nouvelles firmes et leurs processus de localisation, d'évaluer les interactions entre localisation des industries, taux de croissance de long terme et intégration économique.

D'un point de vue technique, la réunion entre ces domaines apparaît réalisable car les deux littératures partagent souvent des hypothèses communes, comme la présence d'effets externes et le recours à la concurrence monopolistique, qui sont formalisées dans les modèles d'économie géographique de la croissance (Martin et Ottaviano, 1999 ; Baldwin et Forslid, 2000). Il est donc important d'étudier comment les résultats précédemment obtenus sur les schémas de localisation des firmes sont influencés par l'introduction d'une dynamique de croissance. Pour cela, le caractère local ou global des externalités de connaissance joue un rôle prépondérant.

134.2 Effets de débordement et équilibres spatiaux

L'intégration d'une dynamique de croissance dans les modèles d'économie géographique passe par l'introduction d'un secteur de R&D, qui produit des innovations sous la forme de nouvelles variétés de biens intermédiaires, dans Englmann et Walz (1995) et Walz (1996), et sous la forme « d'idées » pour de nouvelles variétés de biens différenciés dans Martin et Ottaviano (1999). Compte tenu de la structure de concurrence monopolistique adoptée par les auteurs, la création de nouvelles variétés de biens se traduit en fait par la création de nouvelles firmes.

Ce secteur de la R&D bénéficie d'externalités de connaissance. Dans Englmann et Walz (1995) et Walz (1996), le secteur de R&D bénéficie d'externalités de type « Marshall-

Arrow-Romer » (ou MAR) : la connaissance produite par ce secteur accroît sa propre productivité, de telle sorte que la création de nouvelles variétés de biens intermédiaires est une fonction du nombre total de variétés déjà existantes. Dans Martin et Ottaviano (1999), l'activité de R&D de chaque firme bénéficie d'externalités de type « Jacobs », résultant de la diversité de l'environnement productif au sein duquel les firmes évoluent¹⁸. Dans ce cadre, la productivité de la R&D est d'autant plus élevée que l'éventail de processus de production de biens technologiques différenciés est large. Dans tous les cas, différents équilibres géographiques émergent selon que les effets de débordement ainsi mis en évidence sont globaux ou locaux.

Dans le modèle d'Englmann et Walz (1995), il y a deux régions, deux facteurs de production (le travail qualifié et non qualifié) et trois types de biens. Le travail qualifié est parfaitement mobile entre les deux régions alors que le travail non qualifié est immobile. Il existe deux types de biens qui sont échangés sans coût entre les régions : le bien traditionnel, produit avec du travail qualifié et du travail non qualifié, et le bien différencié, produit avec du travail qualifié, du travail non qualifié et des biens intermédiaires. Ces derniers sont des biens et services produits localement et non-échangeables entre les régions. Avec le travail non-qualifié, ils jouent ainsi le rôle de force de dispersion. Les forces d'agglomération sont dues à l'existence de liens verticaux entre les firmes produisant des biens industriels et celles produisant les biens intermédiaires localisés. Les auteurs utilisent un mécanisme de croissance endogène à la Romer (1990) et Grossman et Helpman (1991) dans lequel la croissance est le résultat d'un changement technique endogène provenant du développement permanent et coûteux de nouveaux biens intermédiaires. Dans ce contexte, deux cas extrêmes, concernant la portée spatiale des externalités de connaissance, sont considérés.

- D'une part, si les externalités sont strictement locales (elles ne profitent qu'au secteur de R&D de la région elle-même), alors le schéma centre-périphérie émerge. La région ayant un avantage initial dans le nombre de producteurs de biens ou de services intermédiaires devient le seul centre industriel. Par ailleurs, cette région est la seule à connaître un taux de croissance positif à l'état régulier. Au contraire, la région périphérique se spécialise dans la production du bien traditionnel. Dans ce cas, le caractère local des externalités devient donc une force supplémentaire d'agglomération.

- D'autre part, si les externalités sont globales, elles ont une portée interrégionale et se diffusent indépendamment de la distance. Le nombre total de variétés dans les deux régions abaisse alors le coût de production d'une nouvelle variété quel que soit le lieu de production. Dans ce cas, il y a une plus grande variété d'équilibres géographiques possibles pour différentes valeurs des paramètres. On peut alors obtenir des configurations dans lesquelles l'industrie se répartit également dans les régions qui ont toutes deux un taux de croissance positif ou des configurations aboutissant à une spécialisation relative ou totale d'une région.

¹⁸ Les externalités de type MAR et de type Jacobs peuvent être considérées respectivement comme les versions dynamiques des économies de localisation et d'urbanisation.

Walz (1996) reprend le cadre précédent mais suppose que les biens intermédiaires sont échangeables avec coûts de transport de type iceberg et ne considère que le cas d'effets de débordement interrégionaux. L'immobilité de la main d'œuvre joue toujours le rôle de force de dispersion alors que la force d'agglomération repose sur les liens verticaux entre biens industriels et biens intermédiaires. L'auteur montre alors que, si les dotations en facteurs sont identiques entre les régions, le modèle conduit, selon la valeur des paramètres, soit à un schéma centre-périphérie dans lequel les activités d'innovation sont concentrées dans l'une des deux régions, soit à un schéma d'équi-répartition pour lequel l'innovation se développe à un taux identique dans chacune des régions. En revanche, si les dotations diffèrent entre les régions, il existe une trajectoire unique qui aboutit à un équilibre stable de long terme dans lequel le secteur de R&D est concentré dans une seule des deux régions.

Dans le modèle de Martin et Ottaviano (1999), les coûts de R&D dépendent du nombre de producteurs différenciés, et non de la connaissance produite par ce secteur comme précédemment. Les auteurs considèrent deux régions, le « Nord » et le « Sud ». Ces régions sont identiques à tous points de vue, excepté dans leurs stocks initiaux de capital humain, le « Nord » étant mieux doté. Il existe trois secteurs. Le premier fabrique, en situation de concurrence pure, un bien homogène, considéré comme le numéraire. Le second fabrique, en concurrence monopolistique, un ensemble de biens différenciés qui sont échangés entre le Nord et le Sud avec un coût de transport de type *iceberg*. Le troisième est le secteur de la R&D fournissant des idées pour de nouvelles variétés de biens différenciés. Le facteur travail est supposé mobile entre les secteurs mais pas entre les régions.

Les équilibres obtenus varient à nouveau selon que les externalités dans le secteur de la R&D sont locales ou globales.

- Si les externalités sont globales, alors la création de nouvelles firmes de biens différenciés diminue les coûts futurs de R&D de manière identique entre les deux régions. La localisation des firmes n'est donc pas liée à la recherche de proximité. Les choix de localisation des firmes dépendent alors de deux forces opposées. D'une part, la région à faible capital attire les firmes car la concurrence y est moins vive. D'autre part, les firmes sont également attirées par la région la mieux dotée en capital car les revenus régionaux et la demande locale y sont plus importants. Le sens des relocalisations des firmes, du Nord vers le Sud, ou du Sud vers le Nord, dépend alors de l'intensité relative de ces deux forces et de la valeur des coûts de transport. Lorsque ces derniers sont élevés, les firmes préfèrent desservir les marchés locaux. En revanche, si les coûts de transport passent en dessous d'une valeur critique, on peut observer un mouvement de relocalisation relative dans la région la moins riche initialement.

- Si les externalités sont locales, alors le besoin de proximité entre les firmes devient une force d'agglomération car les coûts de R&D sont les plus faibles dans la région possédant le plus grand nombre de firmes différenciées. Dans ce cas, toutes les activités de R&D s'agglomèrent dans le Nord où les firmes sont plus nombreuses et le taux de croissance plus élevé. Les auteurs montrent alors qu'une décroissance des coûts de transport conduit les firmes à s'agglomérer dans la région comprenant les activités de R&D. Cependant, cette agglomération n'est pas totale.

Kubo (1995) propose enfin un modèle, s'appuyant sur le modèle de développement inégal de Krugman (1981), dans lequel les externalités sont à la fois locales et globales. Dans ce cas, une source de croissance endogène localisée dans une région est susceptible d'abaisser les coûts dans l'une ou dans les deux régions. Du point de vue formel, ce modèle s'éloigne des précédents car la croissance endogène n'est plus le résultat d'une accumulation des connaissances dans le secteur de R&D. Au contraire, elle est entretenue par un processus d'accumulation du capital physique alimenté par les profits. L'accumulation du capital dans une région peut réduire les coûts de production dans cette région (économies d'échelle) et également dans l'autre région (économies externes régionales). Dans ces conditions, différentes configurations spatiales d'équilibre apparaissent, selon les intensités relatives des économies d'échelle internes à une région et les économies externes régionales. Dans le premier cas, les effets internes à la région dominent les économies externes régionales et les configurations de type centre-périphérie de Krugman (1991a) apparaissent. Dans le second cas, si les effets externes l'emportent dans les deux régions, il y a croissance dans les deux régions dont la taille industrielle est stable. Le dernier cas correspond à des effets externes dominant dans une région et des effets internes dominant dans l'autre région. Le développement prend alors une forme intermédiaire et dépend de la taille initiale de la région.

La présentation rapide de ces modèles a fait apparaître des configurations géographiques d'équilibre différentes selon que les externalités de connaissance sont locales ou globales. Par rapport aux modèles sans dynamique de croissance, on observe par ailleurs une plus grande variété dans les schémas d'équilibre. En particulier, des schémas centre-périphérie « impurs » (Baumont et Huriot, 1999) peuvent émerger, dans lesquels la totalité du secteur de R&D, mais seulement une partie du secteur des biens différenciés, se localisent dans une région. Ces modèles ont permis d'étudier comment la croissance du nombre des firmes produisant des biens différenciés influence les schémas d'équilibre spatiaux. Il convient maintenant de se demander quels sont, en retour, les effets des configurations géographiques sur le taux de croissance.

134.3 Croissance et agglomération

Peu de modèles se sont encore penchés sur la double causalité existant entre agglomération spatiale et croissance économique (Martin et Ottaviano, 1999, 2001 ; Baldwin et Forslid, 2000 ; Baldwin *et al.*, 2001 ; Fujita et Thisse, 2002b). Ces modèles cherchent à comprendre comment la croissance endogène engendre certaines configurations spatiales qui, en retour, agissent sur le taux de croissance. Cette approche permet alors d'aboutir à plusieurs résultats nouveaux sur les effets de l'intégration économique. Nous détaillons ici les deux modèles proposés par Martin et Ottaviano (1999, 2000).

Dans le modèle de Martin et Ottaviano (1999), nous avons vu que plusieurs configurations spatiales différentes émergent selon le caractère global ou local des externalités. En retour, ces configurations ont des effets différenciés sur le taux de croissance. Ainsi, en présence d'externalités globales, les auteurs montrent que la configuration spatiale n'a pas d'influence sur le taux de croissance des régions, puisque les externalités de connaissance produites dans une région profite à toutes les firmes quelle que soit leur localisation. En revanche, en présence d'externalités locales, il apparaît que la concentration spatiale des activités bénéficie à la croissance. Les coûts de transport jouent, là encore, un rôle important. Lorsqu'ils sont peu élevés, les firmes sont incitées à se localiser au Nord, l'effet externe sur les coûts de R&D y devient donc plus important (compte tenu du caractère local des externalités) ainsi que le taux de croissance de cette région. La chute des coûts de R&D a cependant pour conséquence de stimuler la création de nouvelles firmes dont une partie ira se délocaliser au Sud car la concurrence en prix y est moins importante. Par conséquent, si dans un premier temps, la baisse des coûts de transport a pour conséquence d'augmenter les disparités spatiales entre les régions tout en ayant un effet positif sur la croissance, la relocalisation de nouvelles firmes au Sud permet, dans un deuxième temps, la diffusion de la croissance à l'ensemble de l'économie. Ce deuxième effet, absent des modèles statiques, permet alors d'expliquer les relocalisations au Sud suite à une intégration commerciale. Le schéma centre-périphérie et les tendances à l'extrême concentration doivent donc être nuancés.

Le second modèle de Martin et Ottaviano (2001) permet de montrer que la croissance et l'agglomération sont des processus se renforçant mutuellement. Les auteurs se concentrent exclusivement sur les externalités globales dans le but d'éliminer l'effet d'agglomération lié à la présence d'externalités localisées. Deux résultats importants sont à noter.

D'une part, ils montrent que la croissance économique est une cause suffisante d'agglomération. En effet, les travailleurs sont immobiles, à la fois entre les régions et entre les secteurs. Dans ces conditions, il ne peut pas se produire de causalité circulaire et cumulative dans les modèles d'économie géographique sans croissance. Cependant, les auteurs montrent que, si une dynamique de croissance est introduite, il se produit un processus cumulatif aboutissant à l'agglomération.

D'autre part, l'effet de l'agglomération sur la croissance provient de l'introduction d'une nouvelle liaison verticale : les auteurs supposent que le secteur de R&D consomme des biens différenciés comme biens intermédiaires. Par conséquent, ce secteur est affecté par deux types d'externalités : des externalités de connaissance comme précédemment et des externalités pécuniaires liés à la consommation de biens intermédiaires. Dans ce cas, le coût de l'innovation est plus bas dans la région où sont localisées les firmes du secteur intermédiaire car le bien intermédiaire est moins cher. En conséquence, la concentration favorise le secteur de R&D et donc la croissance de l'économie. Grâce à cette nouvelle hypothèse, les auteurs peuvent donc montrer que l'agglomération des activités compte pour la croissance, même en l'absence d'externalités locales, seules les interactions de marché entre le secteur de la R&D et le secteur des biens différenciés sont nécessaires.

Dans ces conditions, si l'on démarre d'une situation initiale où les deux régions sont identiques, les auteurs montrent que lorsque l'économie commence à croître, le seul équilibre est un équilibre spatial dans lequel l'une des régions capte tout le secteur de R&D et la majeure partie de la production. On retrouve donc un schéma centre-périphérie « impur ».

Les modèles d'économie géographique intégrant un processus de croissance amènent à s'interroger de façon plus nuancée sur l'évolution des disparités entre les régions. Ces nuances sont de trois natures. La première tient à la structure économique des régions qui est plus complexe puisqu'un secteur productif porteur de croissance est maintenant considéré : les économies régionales en croissance sont celles qui innovent en fabriquant une plus grande variété de biens. La deuxième tient à l'émergence d'équilibres géographiques intermédiaires entre le schéma centre-périphérie et le schéma d'équi-répartition. Dans ces schémas intermédiaires, les activités industrielles peuvent être réparties entre les économies régionales tandis que les capacités d'innovation, portées par exemple par le secteur de R&D, restent concentrées dans une des régions. Les conditions d'émergence de tels équilibres centre-périphérie « impurs » restent néanmoins largement subordonnées à l'existence d'effets de débordement géographiques interrégionaux sur les externalités de connaissance. La troisième nuance concerne la réciprocity des liens entre schémas géographiques et schémas de croissance puisque l'on peut établir que, si la croissance économique influence la répartition des activités économiques entre les territoires, en retour, cette même répartition n'est pas neutre sur les perspectives de croissance des économies régionales.

Conclusion de section

L'inégale répartition géographique des activités économiques, de la population, de la croissance économique sont des phénomènes observés qui relèvent largement du domaine des faits stylisés. Pourtant, ces phénomènes sont rarement pris en compte dans la modélisation et dans l'estimation des régressions de convergence.

Pour l'économétrie, cet « oubli » de l'espace dans l'analyse de la convergence occasionne deux problèmes. D'une part, si les valeurs observées ne sont pas réparties au hasard dans l'espace, alors les erreurs sont susceptibles d'être spatialement autocorrélées. D'autre part, s'il existe un phénomène de polarisation spatiale des activités, alors les paramètres des coefficients de régressions peuvent ne pas être constants dans l'espace, aboutissant alors à un raisonnement en termes de clubs de convergence.

Pour l'économie, ces différents phénomènes peuvent être reliés aux théories de l'économie géographique qui analysent les mécanismes de formation des agglomérations, c'est-à-dire la répartition inégale dans l'espace des activités économiques et des individus. Le modèle de base d'équilibre interrégional que nous considérons est celui proposé par Krugman (1991a, 1991b) et il s'appuie sur une structure productive à rendements croissants, en concurrence monopolistique et intégrant des externalités. Les firmes industrielles fondent leur choix de localisation sur un arbitrage entre les besoins de proximité, qui sont facilités par l'agglomération, et la recherche d'accessibilité pour réaliser les échanges, qui est conditionnée

par les coûts d'interaction. Dans ce cadre, les théories de l'économie géographique permettent en particulier de rendre compte pourquoi les politiques d'intégration économique, qui tendent à favoriser les échanges en abaissant les coûts de transport et en facilitant la mobilité des individus et des biens, stimulent l'émergence d'un schéma polarisé de type centre-périphérie. Les modèles d'économie géographique de la croissance se sont développés récemment en intégrant les mécanismes d'accumulation spatiale et temporelle des activités économiques. La croissance économique est portée par les innovations et les externalités de connaissance qui peuvent avoir une diffusion interrégionale. Les mécanismes d'interaction entre croissance et formation des agglomérations sont formalisés et montrent que les schémas géographiques ne sont pas neutres sur les potentialités de croissance économique.

Enfin, tous ces éléments plaident en faveur d'une prise en compte explicite de l'espace dans l'analyse de la convergence des économies.

Conclusion du chapitre 1

Dans ce chapitre, nous avons dressé un bilan des recherches menées sur le thème de la convergence des économies. L'hypothèse de convergence est d'abord basée sur le modèle néoclassique de Solow. Deux concepts de convergence sont déduits de ce cadre d'analyse, selon que les états réguliers des économies sont supposés identiques ou non. Dans le premier cas, les économies convergent vers un même état régulier, c'est le concept de β -convergence absolue. Dans le second cas, les économies convergent chacune vers leur état régulier respectif, c'est le concept de β -convergence conditionnelle. L'hypothèse de β -convergence conditionnelle a donné lieu à un très grand nombre de travaux, dont les résultats sont souvent entachés par de nombreux problèmes. L'hypothèse de clubs de convergence se rattache, quant à elle, aux modèles de croissance endogène à équilibres multiples ou au modèle de croissance néoclassique avec hétérogénéité des agents. Enfin, l'hypothèse de σ -convergence n'est pas spécifiquement liée à un modèle de croissance mais entretient avec le concept de β -convergence des liens étroits.

D'autres approches suggèrent que ni la β -convergence, ni la σ -convergence ne permettent de « résoudre » le problème de la convergence. En effet, ces concepts n'utilisent pas toute l'information associée aux séries de PIB par tête sur différentes années. Ainsi, Quah (1993a, 1996a, 1996b) propose de caractériser entièrement la dynamique de la répartition des revenus par tête par l'étude des propriétés des chaînes de Markov associées au processus de convergence. Son approche permet alors de révéler les éventuels clubs de convergence et d'évaluer le degré de mobilité des pays d'un état de richesse à l'autre. De leurs côtés, Evans et Karras (1993, 1996), ou encore Bernard (1991) et Bernard et Durlauf (1995) ont pour objectif d'exploiter pleinement la dimension temporelle des séries en vérifiant l'hypothèse de convergence des séries par la caractérisation des écarts de revenu par tête.

Toutes ces approches négligent cependant la dimension spatiale associée aux données utilisées. Nous avons alors montré que dans ce cas, d'autres problèmes économétriques apparaissent si une éventuelle corrélation des erreurs, due à l'interdépendance spatiale des économies, n'est pas prise en compte. Cet aspect est rarement étudié dans la littérature alors même que les théories de l'économie géographique permettent d'expliquer pourquoi les activités et les hommes sont concentrés en un petit nombre de lieux et comment ces schémas de localisation interagissent avec les schémas de croissance.

Concernant les modèles « statiques » de l'économie géographique, nous avons montré que suite à la présence de nombreuses forces d'agglomération, les activités industrielles, tertiaires et de R&D tendent à se concentrer dans un petit nombre de régions dans les pays développés. Par ailleurs, la distribution géographique des aires caractérisées par de hautes ou de faibles densités d'activités économiques est rarement aléatoire, les endroits où les agglomérations se localisent étant identifiés par les conditions de première ou de seconde

nature. Ces modèles se caractérisent enfin le plus souvent par l'émergence d'un schéma centre-périphérie¹⁹ en présence d'intégration régionale (baisse des coûts de transport).

Concernant la synthèse géographie-croissance, plusieurs résultats pour l'analyse des processus de croissance ont été obtenus. Tout d'abord, on montre que la concentration spatiale des activités favorise la croissance. Ensuite, la croissance économique peut en retour renforcer les processus de polarisation. Enfin, les effets de débordement géographiques ont plusieurs conséquences sur les schémas de localisation selon leur caractère local ou global. Dans le premier cas, la polarisation des territoires est renforcée. Dans le second cas, ces effets peuvent conduire à une certaine répartition plus égalitaire entre les régions. Des résultats intermédiaires peuvent encore être obtenus si la concentration de firmes dans une région produit à la fois des effets de débordement globaux et locaux (Kubo, 1995) : des schémas d'équilibre ou de développement inégal apparaissent selon les intensités relatives des effets de débordement entre une région et ses voisines.

Tous ces résultats théoriques montrent que les schémas géographiques peuvent être structurés par les processus de croissance et qu'ils peuvent, en retour, orienter les schémas de croissance. En appliquant ces résultats à l'analyse d'un espace régional intégré, on aboutit alors aux observations suivantes :

- 1/ Puisque les activités économiques sont distribuées inégalement dans l'espace, des processus cumulatifs d'agglomération ont lieu et la plupart des activités économiques tendent à se concentrer dans un petit nombre de localisations.
- 2/ Puisque la croissance économique est stimulée par la concentration géographique des activités économiques, des schémas de développement inégal sont observés.
- 3/ Puisque l'histoire compte, à travers les conditions initiales et la nature cumulative des processus de croissance et d'agglomération, la distribution géographique observée de régions riches et de régions pauvres est relativement stable dans le temps.

Tous ces effets doivent être pris en compte dans l'analyse des processus de convergence. D'une part, il faut considérer la possibilité de l'existence d'un schéma de type centre-périphérie et de clubs de convergence qui en découle. L'homogénéité des schémas de convergence sur l'ensemble des régions est ici mise en cause et les formes « d'hétérogénéité spatiale » doivent au contraire être considérées. D'autre part, il faut intégrer le fait que les observations géographiques ne sont pas réparties aléatoirement dans l'espace. Le problème de l'indépendance spatiale est ici directement questionné et amène à s'interroger sur les phénomènes « d'autocorrélation spatiale » des productions ou des revenus par tête. Les méthodes de l'économétrie spatiale permettent de prendre en compte ces deux aspects et nous les mobilisons dans les études empiriques de la convergence des régions européennes réalisées dans les trois derniers chapitres de la thèse. Nous consacrons le chapitre suivant à la présentation détaillée de ces méthodes.

¹⁹ Dans des modèles multi-régionaux, on peut en outre montrer que deux agglomérations sont séparées par une distance minimale car il y a un effet d'ombre empêchant la formation d'une agglomération distincte dans le voisinage d'une autre agglomération (Krugman, 1993).

CHAPITRE 2

Econométrie spatiale

La majeure partie des travaux présentés dans ce chapitre est issue des textes suivants :

Le Gallo J. (2001a) Econométrie spatiale, 2. Hétérogénéité spatiale, *Document de travail du LATEC*, Dijon, n°2001-01 (Janvier 2001).

Le Gallo J. (2002) Econométrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire, *Economie et Prévision*, à paraître.

Chapitre 2

Econométrie spatiale

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

Historiquement, c'est à Cliff et Ord (1973) que l'on doit, après une série d'articles à la fin des années 60 et au début des années 70, un ouvrage présentant de manière synthétique l'état des savoirs en statistique et en économétrie spatiales. La fin des années 70 et les années 80 sont marquées par le raffinement du cadre original d'analyse et plus particulièrement par le développement de la théorie de l'estimation et des tests (Ord, 1975 ; Paelinck et Klassen, 1979 ; Anselin, 1980). Un ensemble d'ouvrages rendent compte de ces développements, comme par exemple ceux de Cliff et Ord (1981), Anselin (1988a), Griffith (1988a), Cressie (1993), Jayet (1993) ou Anselin et Florax (1995a).

Jusqu'au début des années 90, les techniques permettant de spécifier, d'estimer et de tester la présence de l'autocorrélation spatiale dans les modèles économétriques étaient principalement publiées dans les revues spécialisées et appliquées à des problèmes d'économie régionale, spatiale ou urbaine. Ainsi, Can (1992), Can et Megboluge (1997), Pace et Gilley (1997) ont abordé le traitement de l'autocorrélation spatiale dans les modèles hédoniques de prix immobiliers, alors qu'Anselin *et al.* (1997) et Varga (1998, 2000) ont étudié les externalités spatiales d'information dues à la recherche universitaire et à la R&D. Toutefois, les méthodes de l'économétrie spatiale sont maintenant appliquées à d'autres sujets tels que l'analyse de la demande (Case, 1991), l'économie internationale (Aten, 1997), l'économie publique (Case *et al.*, 1993 ; Brueckner, 1998), l'économie rurale (Benirschka et Binkley, 1996), les lois de Kaldor-Verdoorn (Fingleton, 2000, 2001, 2002) ou encore les phénomènes de croissance et de convergence (Moreno et Trehan, 1997 ; Fingleton, 1999 ; Rey et Montouri, 1999). Ces méthodes sont, en effet, potentiellement applicables à toutes les études empiriques nécessitant l'utilisation de données spatiales.

Selon Anselin et Bera (1998) et Anselin (2001b), deux raisons principales peuvent être attribuées au regain d'attention porté à la prise en compte des effets spatiaux dans les études empiriques. La première est le développement de nouveaux courants théoriques, tels que l'économie géographique et ses ramifications en économie industrielle, économie internationale, théories de la croissance ou économie du travail..., qui ont généralisé la prise en compte des interactions et des externalités spatiales dans l'analyse des décisions économiques des agents. La modélisation de l'autocorrélation spatiale permet en effet de capter l'existence, l'ampleur et l'influence des effets de débordement géographiques. La deuxième raison, d'ordre empirique, est la disponibilité croissante des données spatialisées, le fort développement actuel des logiciels de Systèmes d'Information Géographiques (SIG) et l'apparition de logiciels, tels que *SpaceStat* (Anselin, 1999), le module *S+SpatialStat* pour *SPlus* (Mathsoft, 1996) ou la librairie pour Matlab (LeSage, 1999a), proposant l'estimation des principaux modèles spatiaux.

L'objectif de ce chapitre est de présenter les outils nécessaires à une démarche économétrique visant à prendre en compte les effets d'autocorrélation et d'hétérogénéité spatiales. Notre présentation est organisée de la façon suivante.

Dans la première section, nous précisons les particularités des données spatiales ainsi que les méthodes de traitement de ces données. Ces dernières sont souvent caractérisées par deux effets : l'autocorrélation et l'hétérogénéité spatiales que nous définissons précisément. Nous montrons comment l'autocorrélation spatiale peut être formalisée à partir de la notion de matrice de poids, puis nous présentons les méthodes d'analyse exploratoire des données spatiales qui permettent de déterminer la présence et l'ampleur de ces deux effets.

La deuxième section est consacrée aux méthodes de traitement de l'autocorrélation spatiale dans les modèles de régression linéaire. Nous présentons les trois spécifications les plus utilisées ainsi que leurs principales propriétés, puis nous abordons les méthodes d'estimation adaptées à ces spécifications. Ces modèles se caractérisent, en effet, par une corrélation entre les erreurs et les variables explicatives. En conséquence, l'inférence statistique basée sur les Moindres Carrés Ordinaires (MCO) n'est pas fiable en présence d'autocorrélation spatiale et d'autres méthodes d'estimation doivent être utilisées : méthode du maximum de vraisemblance, des variables instrumentales et des moments généralisés. Nous exposons ensuite les méthodes d'inférence permettant de tester la présence d'autocorrélation spatiale dans les modèles. En se basant sur les propriétés des différents tests, nous proposons alors une démarche générale visant à déterminer la forme prise par l'autocorrélation spatiale dans les modèles de régression.

Dans la troisième section, nous étudions le phénomène d'hétérogénéité spatiale. Nous présentons d'abord les principales spécifications en coupe transversale permettant de capter les deux facettes de cet effet : instabilité des paramètres et hétéroscédasticité. Ensuite, nous abordons les liens qui unissent autocorrélation et hétérogénéité spatiales et la façon dont les différents tests présentés au long de ce chapitre doivent être amendés pour tenir compte des deux effets simultanément.

SECTION 21 : LES EFFETS SPATIAUX : DEFINITION ET DETECTION

Les données spatiales sont des observations d'une variable mesurée en des localisations différentes réparties dans l'espace. Par conséquent, les données spatiales sont « spéciales » : elles combinent une information sur les attributs ainsi qu'une information sur la localisation de ces attributs, mesurée par exemple par une adresse postale ou des coordonnées géographiques (latitude et longitude). Les données spatiales possèdent souvent des propriétés particulières et doivent être analysées différemment des données a-spatiales. Le but de cette section est de préciser la nature particulière de ces données.

Il existe différents types de données spatiales et chacun appelle des traitements statistiques et économétriques différents. En économie, on est généralement confronté à trois types d'observations. Il peut s'agir, tout d'abord, de points représentant, par exemple, des localisations d'unités de production ou de distribution... Ces points sont souvent mesurés par leur latitude et leur longitude. Ensuite, ces localisations peuvent être des lignes, connectées entre elles ou non, comme un réseau routier ou fluvial. Enfin, les données sont parfois fournies pour des aires géographiques comme des régions ou des pays. Par la suite, nous nous restreignons au 3^{ème} type d'observations.

Dans tous les cas, le nombre de ces points, de ces lignes ou de ces zones est *fini*. Cette caractéristique permet de distinguer les techniques de l'économétrie spatiale de celles de la géostatistique (Cressie, 1993 ; Griffith et Layne, 1999). L'économétrie spatiale est principalement utilisée lorsqu'on est en présence d'un *ensemble fini* (régulier ou non) de points ou de zones reliés entre eux par des relations de voisinage. La géostatistique concerne essentiellement les *données issues d'un processus spatial sous-jacent continu* sur l'espace étudié. L'utilisation de l'une ou l'autre méthodologie dépend de la disponibilité des données et du problème étudié. Les techniques de la géostatistique sont d'un intérêt moindre pour l'étude de données socio-économiques. Une exception notable est l'étude des modèles hédoniques de prix immobiliers, pour lesquels les deux types de traitement sont appropriés (Dubin *et al.*, 1999)²⁰. Dans le cadre de cette thèse, nous nous limitons à la présentation des méthodes de l'économétrie spatiale.

Les données spatiales sont souvent caractérisées par deux effets : l'autocorrélation spatiale qui se réfère à l'absence d'indépendance entre des observations géographiques, et l'hétérogénéité spatiale qui est liée à la différenciation des variables et des comportements dans l'espace (paragraphe 211). Des outils spécifiques sont nécessaires à la modélisation des interdépendances entre les observations : les matrices de poids et les variables spatiales décalées (paragraphe 212). Les méthodes d'analyse exploratoire des données spatiales ont

²⁰ Ainsi, Can (1990, 1992) ou encore Pace et Gilley (1997) utilisent les méthodes de l'économétrie spatiale alors que Dubin (1992, 1998), Basu et Thibodeau (1998) ou Gillen *et al.* (2001) font appel à la géostatistique.

pour but de caractériser les distributions ainsi que la présence et l'ampleur des effets spatiaux (paragraphe 213).

211 Autocorrélation et hétérogénéité spatiales

Depuis Anselin (1988a), l'habitude a été prise de classer les effets spatiaux en deux grandes catégories : tout d'abord, la dépendance spatiale qui se réfère à l'absence d'indépendance entre observations géographiques (paragraphe 211.1), et ensuite l'hétérogénéité spatiale qui est liée à la différenciation spatiale des comportements (paragraphe 211.2).

211.1 L'autocorrélation spatiale

Dès 1914, Student suspecte la présence d'une relation entre différentes observations géographiques, une idée qui entraîne l'abandon de l'hypothèse statistique fondamentale d'observations indépendantes. Ce phénomène, appelé autocorrélation spatiale, est défini par Anselin et Bera (1998) de la façon suivante :

« Spatial autocorrelation can be loosely defined as the coincidence of value similarity with locational similarity. »

En d'autres termes, *l'autocorrélation spatiale positive* se traduit par une tendance à la concentration dans l'espace de valeurs faibles ou élevées d'une variable aléatoire. En revanche, *l'autocorrélation spatiale négative* signifie que chaque localisation tend à être entourée par des localisations voisines pour lesquelles la variable aléatoire prend des valeurs très différentes. Enfin, *l'absence d'autocorrélation spatiale* indique que la répartition spatiale des valeurs de la variable est aléatoire. Détecter de l'autocorrélation spatiale dans une série spatiale fournit alors une information supplémentaire par rapport aux statistiques traditionnelles (telles que la moyenne ou l'écart-type) sur la façon dont les différentes valeurs sont disposées géographiquement (Griffith, 1992b).

La présence d'autocorrélation spatiale pour une variable signifie qu'il y a une relation fonctionnelle entre ce qui se passe en un point de l'espace et ce qui se passe ailleurs. Tobler (1979) l'avait déjà souligné en suggérant la première loi de la géographie suivante :

« Everything is related to everything else, but closer things more so. »

Les concepts de « proximité » et de « distance » sont pris en compte à travers l'utilisation de matrices de poids. L'autocorrélation spatiale est multidirectionnelle et « tout est relié à tout » ; elle diffère donc de l'autocorrélation temporelle qui est unidirectionnelle.

Cette interdépendance généralisée a pour conséquence de rendre plus complexes les méthodes de traitement de l'autocorrélation spatiale. Par exemple, certaines méthodes d'estimation valables pour les séries temporelles ne sont pas directement transposables au cas spatial (cf. paragraphe 222, p. 129).

L'autocorrélation spatiale a deux sources principales :

1/ Elle peut provenir du fait que les données sont affectées par des processus qui relient des lieux différents et qui sont à l'origine d'une organisation particulière des activités dans l'espace (Odland, 1988 ; Haining, 1990).

En effet, les processus d'interaction sont source d'autocorrélation spatiale lorsque les événements ou les circonstances en un lieu donné affectent les conditions en d'autres lieux si ces derniers interagissent, d'une manière ou d'une autre, par des mouvements de biens, de personnes, de capitaux, par des externalités spatiales ou par toutes les formes de comportements où un acteur économique réagit aux actions d'autres acteurs. Ainsi, la diffusion d'un phénomène (comme la diffusion technologique) à partir d'un ou de plusieurs lieux d'origine implique que l'intensité de la mesure de ce phénomène dépend de la distance à l'origine. Aux localisations proches les unes des autres, et donc à des distances comparables de l'origine, sont donc associées des intensités similaires pour le phénomène étudié.

2/ Elle peut également provenir d'une mauvaise spécification du modèle, comme des variables omises spatialement autocorrélées, d'une forme fonctionnelle incorrecte ou d'erreurs de mesure. C'est, en particulier, le cas lorsque l'étendue spatiale du phénomène étudié ne coïncide pas avec les unités spatiales d'observation. Elle est alors considérée comme un outil de diagnostic et de détection d'une mauvaise spécification du modèle.

211.2 L'hétérogénéité spatiale et ses deux facettes

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

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

de l'Europe et dans le « Sud », etc. Plus précisément, deux aspects sont liés à cette notion : l'instabilité et l'hétéroscédasticité.

Le premier, *l'instabilité*, provient de l'absence de stabilité dans l'espace des comportements ou d'autres relations étudiées : les formes fonctionnelles et les paramètres varient selon leurs localisations et ne sont donc pas homogènes. Il est donc nécessaire de mobiliser des modélisations prenant en compte les caractéristiques particulières de chaque localisation de l'échantillon.

Le deuxième aspect de l'hétérogénéité spatiale est *l'hétéroscédasticité*. Dans les modèles économétriques, elle peut venir de variables manquantes ou de toute autre forme de mauvaise spécification. Par exemple, les unités spatiales elles-mêmes ne sont généralement ni de formes régulières, ni homogènes : des régions peuvent avoir des formes et des aires différentes, des niveaux de développement technologique variables, des populations plus ou moins importantes, etc.

En conclusion, pour décrire des phénomènes quantitatifs ou qualitatifs ayant une dimension spatiale, les interactions spatiales sont fréquentes. Ainsi, il est nécessaire de considérer non seulement les dimensions et les structures des observations mais aussi leurs positions relatives. Ceci s'effectue à l'aide des matrices de poids.

212 Les matrices de poids et les variables spatiales décalées

Pour modéliser les interactions spatiales, il est nécessaire d'imposer une structure sur l'étendue de ces interactions. En effet, on ne peut pas estimer $N(N-2)/2$ termes d'interaction avec N observations. En économétrie spatiale, on définit alors un ensemble de voisins pour chaque région de l'échantillon, ce qui aboutit aux matrices de poids²¹. D'une façon générale, les matrices de poids permettent de spécifier de manière exogène la topologie du système spatial.

Une matrice de poids W est une matrice carrée, ayant autant de lignes et de colonnes qu'il y a de zones géographiques (on note N le nombre de régions) et où chaque terme w_{ij} représente la façon dont la région i et la région j sont connectées spatialement. Elles ne contiennent donc pas d'éléments à estimer.

Ces matrices se classent dans deux grandes catégories : les matrices de contiguïté (paragraphe 212.1) et les matrices de poids généralisées (paragraphe 212.2). Les matrices de poids sont souvent standardisées (paragraphe 212.3) et elles permettent de définir la notion de variable spatiale décalée utilisée pour la modélisation économétrique (paragraphe 212.4).

²¹ En géostatistique, la solution consiste à ordonner les paires de localisations i et j en fonction de la distance qui les sépare. Ceci est effectué à l'aide du semi-variogramme (Cressie, 1993).

212.1 Les matrices de contiguïté

Les matrices les plus utilisées sont les *matrices de contiguïté*. La contiguïté à l'ordre 1 entre deux régions se définit par le fait qu'elles ont une frontière commune. Chaque terme w_{ij} est alors défini comme suit :

$$w_{ij} = \begin{cases} 1 & \text{si les régions } i \text{ et } j \text{ sont contiguës à l'ordre 1} \\ 0 & \text{sinon} \end{cases} \quad [2.1]$$

Par convention, une région n'est pas contiguë avec elle-même : $w_{ii} = 0, \forall i$.

Si l'on désire connaître le nombre de régions contiguës à une région i , il suffit de calculer la somme des éléments de la ligne i de la matrice de contiguïté soit $L_i = \sum_j w_{ij}$. Le nombre total de liens existant dans le système régional est alors égal à $A = (1/2) \sum_i L_i$.

Dans le cas d'une disposition spatiale régulière, la définition de la contiguïté n'est pas unique. Ainsi, si l'on considère la figure 2.1, trois notions de contiguïté sont possibles. Ces dernières font référence au jeu d'échecs et au déplacement de la tour, du fou, de la dame. La tour se déplace horizontalement ou verticalement alors que le fou se déplace sur les diagonales. La dame peut se déplacer quant à elle dans toutes les directions. Les définitions de la contiguïté se déduisent de ces différents types de déplacement.

c	b	c
b	a	b
c	b	c

Figure 2.1 : Les définitions de la contiguïté pour une configuration spatiale régulière

Dans le cas de la tour, la région a compte 4 voisines : les régions de type b qui partagent avec la région a des côtés communs. Dans le cas du fou, la région a compte 4 voisines : les régions de type c qui partagent avec la région a des coins communs. Enfin, dans le cas de la dame, la région a compte 8 voisines : les régions de type b et de type c qui partagent un côté ou un coin commun. Ces diverses définitions sont souvent utilisées dans des exercices de simulation visant à établir les propriétés des différents tests.

Cette notion de contiguïté peut être généralisée : on dit que deux régions i et j sont contiguës à l'ordre k si k est le nombre *minimal* de frontières à traverser pour aller de i à j . Cependant, la matrice de contiguïté d'ordre k n'est pas égale à la matrice de contiguïté d'ordre 1 élevée à la puissance k . Cette opération produit, en effet, des « routes circulaires » (des routes qui passent plusieurs fois par une même région) et des « chemins redondants » (des

régions qui sont déjà contiguës à l'ordre $k - 1$ sont encore comptabilisées à l'ordre k). Blommestein (1985), Blommestein et Koper (1992) et Anselin et Smirnov (1996) ont développé des algorithmes permettant de passer d'une matrice de contiguïté à l'ordre 1 à une matrice de contiguïté d'ordre quelconque (cf. annexe 2.1 pour plus de précisions).

Ces matrices de contiguïté sont souvent utilisées en raison de leur simplicité mais apparaissent restrictives pour ce qui est de leur définition de la connexion spatiale entre régions. En outre, une même matrice de contiguïté peut représenter des arrangements très différents d'unités spatiales : c'est le problème de l'invariance topologique (Cliff et Ord, 1981, p. 21). D'autres matrices de poids apparaissent alors utiles.

212.2 Autres matrices de poids

Initialement, le concept simple d'une contiguïté binaire a été étendu par Cliff et Ord (1981) qui utilisent une combinaison d'une certaine mesure de distance et de la longueur relative de la frontière commune entre ces deux lieux, pour tenir compte de l'irrégularité des zonages. Formellement, un élément de la matrice de poids s'écrit :

$$w_{ij} = (d_{ij})^{-a} (\beta_{ij})^b \quad [2.2]$$

où d_{ij} est la distance entre l'unité i et l'unité j , β_{ij} la proportion de la frontière partagée par les régions i et j , a et b sont des paramètres exogènes déterminés *a priori*.

D'autres indicateurs peuvent aussi apparaître pertinents. Dans certains cas, les matrices de poids ne sont pas nécessairement symétriques²². Les diverses possibilités sont les suivantes :

- w_{ij} est la proportion de la frontière que la région i partage avec la région j . Cette matrice n'est pas symétrique.
- $w_{ij} = 1$ si la distance entre les régions i et la région j est inférieure à une valeur-seuil, et 0 sinon. La distance peut être la distance entre les capitales de régions ou la distance entre les centres géographiques.
- w_{ij} est une fonction décroissante de la distance (souvent égale à 0 si la distance excède une valeur-seuil). Diverses spécifications sont disponibles, les plus utilisées étant la fonction exponentielle négative [2.3] ou une fonction de l'inverse de la distance [2.4] :

$$w_{ij} = e^{-\alpha d_{ij}} \quad [2.3]$$

$$w_{ij} = \begin{cases} d_{ij}^{-\beta} & \text{si } d_{ij} < \bar{d} \\ 0 & \text{sinon} \end{cases} \quad [2.4]$$

²² Les propriétés de certaines matrices de poids ont été systématiquement étudiées par Bavaud (1998).

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

- $w_{ij} = 1$ si le centre de la région i est l'un des plus proches voisins de la région j , sinon $w_{ij} = 0$. Cette matrice de poids n'est pas symétrique.
- w_{ij} reflète l'accessibilité de la région i à la région j .

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

L'utilisation de telles matrices de poids peut poser problème. En effet, les procédures d'estimation et de tests dans les modèles spatiaux nécessitent des matrices de poids exogènes. En conséquence, les indicateurs socioéconomiques doivent être choisis de telle sorte que leur exogénéité par rapport aux variables explicatives du modèle soit assurée, sinon, l'endogénéité des poids doit être considérée explicitement dans la spécification du modèle.

212.3 Standardisation des matrices de poids

Les matrices de poids sont souvent standardisées et rendues telles que la somme des lignes est égale à 1. Chaque élément w_{ij} de la matrice est divisé par la somme totale de la ligne :

$$w_{ij}^s = \frac{w_{ij}}{\sum_j w_{ij}} \quad [2.5]$$

Les poids sont alors compris entre 0 et 1 et cette opération permet la comparabilité des paramètres spatiaux dans différents modèles économétriques (cf. paragraphe 221, p. 120). La procédure de standardisation donne une interprétation différente aux poids. Par exemple, dans le cas d'une matrice de distance, la connexion entre deux régions dépend, après standardisation, de la distance *relative* entre elles et non plus de la distance *absolue*. Par ailleurs, les matrices standardisées ne sont plus symétriques.

212.4 Variable spatiale décalée

Soit x une variable aléatoire et W une matrice de poids. On définit la *variable spatiale décalée* (*spatial lag*), pour N régions par le vecteur de dimension $(N,1)$: Wx .

Lorsque W est une matrice standardisée, le $i^{\text{ème}}$ élément de la variable spatiale décalée contient la moyenne pondérée des observations des régions voisines à la région i . Les décalages spatiaux d'ordre supérieur sont définis par W^2x , W^3x , etc.

Cette notion permet la comparaison entre la valeur de y associée à une localisation i et ses voisines. Ainsi, si y_i et $(Wy)_i$ sont similaires, il y a autocorrélation spatiale positive entre y_i et ses voisines alors que si y_i et $(Wy)_i$ sont dissemblables, il y a autocorrélation spatiale négative entre y_i et ses voisines. Cette constatation est à la base du diagramme de Moran servant à évaluer l'autocorrélation spatiale locale (cf. paragraphe 213.22, p. 114). La variable spatiale décalée joue en outre un rôle primordial dans la spécification des modèles économétriques spatiaux (cf. paragraphe 221, p. 120).

Dans ce paragraphe, nous avons défini les deux grands effets caractérisant les données spatiales. Nous présentons maintenant les méthodes d'analyse exploratoire permettant de détecter, dans les séries, la présence de ces deux effets.

213 Analyse exploratoire des données spatiales

L'analyse exploratoire des données spatiales (« ESDA »²³) est un ensemble de techniques destinées à décrire et à visualiser les distributions spatiales, à identifier les localisations atypiques, les observations extrêmes et les regroupements spatiaux, à détecter les schémas d'association spatiale et enfin à suggérer les régimes spatiaux et les autres formes d'hétérogénéité spatiale (Haining, 1990 ; Bailey et Gatrell, 1995 ; Anselin, 1998a, 1998b). Ces méthodes fournissent en particulier des mesures d'autocorrélation spatiale globale (paragraphe 213.1) et locale (paragraphe 213.2).

213.1 Autocorrélation spatiale globale

Lorsque les valeurs prises par une variable aléatoire, discrète ou continue, ne sont pas disposées au hasard sur une carte, mais sont proches pour deux régions voisines, on dit qu'il y a autocorrélation spatiale. Pour mesurer cette autocorrélation spatiale *globale*, on dispose de plusieurs statistiques permettant d'apporter une information non seulement sur les différentes valeurs prises par la variable aléatoire mais également sur la façon dont ces valeurs sont réparties dans l'espace. On distingue le cas des variables qualitatives (paragraphe 213.11) et celui des variables quantitatives pour lesquelles deux indicateurs existent : le coefficient de

²³ *Exploratory Spatial Data Analysis*.

Moran (paragraphe 213.12) et le coefficient de Geary (paragraphe 213.13). Une statistique générale englobant les cas précédents a également été suggérée (paragraphe 213.14).

213.11 Variables qualitatives dichotomiques

Dans la littérature, le cas où la variable étudiée prend deux valeurs correspond au « problème des couleurs ». En effet, on associe aux deux valeurs prises par la variable deux couleurs distinctes : blanc et noir. Dans ces conditions, pour savoir s'il y a autocorrélation spatiale, on considère les frontières qui relient deux observations voisines. Ces frontières peuvent être de trois types : noir-noir (*NN*) quand les deux régions qu'elles séparent sont noires, blanc-blanc (*BB*) quand les deux régions sont blanches, noir-blanc (*NB*) si les deux régions sont de couleurs opposées. Ainsi, intuitivement, pour une autocorrélation spatiale positive, les frontières *NN* et *BB* prédomineront de telle sorte que des concentrations de régions de même couleur apparaissent, alors que pour une autocorrélation négative, les frontières *NB* seront plus nombreuses et il y aura alternance quasi-régulière entre les régions noires et les régions blanches.

Formellement, considérons la variable aléatoire x_i prenant la valeur 1 si la $i^{\text{ème}}$ observation est noire et 0 sinon. Soit W une matrice de poids quelconque (symétrique ou non). Dans ces conditions, *NN*, *NB* et *BB* sont respectivement égaux à :

$$NN = (1/2) \sum_i \sum_j w_{ij} x_i x_j \quad [2.6]$$

$$NB = (1/2) \sum_i \sum_j w_{ij} (x_i - x_j)^2 \quad [2.7]$$

$$BB = (1/2) \sum_i \sum_j w_{ij} (1 - x_i)(1 - x_j) = (1/2) \sum_i \sum_j w_{ij} - NN - NB \quad [2.8]$$

Puisque la statistique *BB* est une fonction linéaire de *NN* et *NB*, elle ne fournit pas d'information supplémentaire et ne sera donc pas considérée par la suite.

Les lois de probabilité suivies par *NN* et *NB* sous l'hypothèse nulle d'absence d'autocorrélation spatiale sont usuellement données pour deux cas différents :

- *Hypothèse 1* : on suppose que chaque région est indépendamment noire ou blanche selon des probabilités p et q données (avec $p + q = 1$). La coloration de la carte peut alors être assimilée à un tirage avec remise. La variable aléatoire définie par le nombre de fois où une région noire (ou blanche) est tirée suit donc une loi binomiale.
- *Hypothèse 2* : on utilise un tirage aléatoire sans remise dans une urne contenant N_1 régions noires et N_2 régions blanches correspondant aux valeurs observées. La variable aléatoire définie par le nombre de fois où une région noire (ou blanche) est tirée suit donc une loi hypergéométrique.

Dans le tableau 2.1 sont donnés les moments d'ordre 1 et 2 des variables aléatoires *NN* et *NB* sous les deux hypothèses.

Hypothèse 1 :

$$E(NN) = (1/2)S_0p_1^2$$

$$V(NN) = (1/4)[S_1p_1^2 + (S_2 - 2S_1)p_1^3 + (S_1 - S_2)p_1^4]$$

$$E(NB) = S_0p_1p_2$$

$$V(NB) = (1/4)[2S_1p_1p_2 + (S_2 - 2S_1)p_1p_2(p_1 + p_2) + 4(S_1 - S_2)p_1^2p_2^2]$$

Hypothèse 2 :

$$E(NN) = (1/2)S_0[N_1^2 / N^2]$$

$$V(NN) = \frac{1}{4} \left[\frac{S_1N_1^2}{N^2} + \frac{(S_2 - 2S_1)N_1^3}{N^3} + \frac{(S_0^2 + S_1 - S_2)N_1^4}{N^4} \right] - E(NN)^2$$

$$E(NB) = S_0(N_1N_2) / N^2$$

$$V(NB) = \frac{1}{4} \left[\frac{2S_1N_1N_2}{N^2} + \frac{(S_2 - 2S_1)N_1N_2(N_1 + N_2 - 2)}{N^3} + \frac{4(S_0^2 + S_1 - S_2)N_1^2N_2^2}{N^4} \right] - E(NB)^2$$

Notations :

N = nombre total de régions

N_1 = nombre de régions noires et N_2 = nombre de régions blanches

p_1 = probabilité que la variable aléatoire prenne la valeur « noir »

p_2 = probabilité que la variable aléatoire prenne la valeur « blanc »

$$S_0 = \sum_i \sum_j w_{ij} \quad ; \quad S_1 = (1/2) \sum_i \sum_j (w_{ij} + w_{ji})^2 \quad ;$$

$$S_2 = \sum_i (w_i + w_i)^2 \quad ; \quad w_i = \sum_j w_{ij} \quad ; \quad w_i = \sum_j w_{ji}$$

Tableau 2.1 : Espérances et variances de NN et NB

Source : Cliff et Ord (1981, p. 19-21)

Considérons la statistique centrée-réduite suivante :

$$z = \frac{NN - E(NN)}{V^{1/2}(NN)} \quad [2.9]$$

Cliff et Ord (1981) montrent que si le nombre d'observations est suffisamment grand, la distribution de cette statistique z est correctement approchée par une loi normale d'espérance nulle et de variance unitaire. La même propriété vaut pour la statistique NB . Il est alors possible de tester les hypothèses 1 ou 2 avec le risque α . En outre, il est utile de faire les tests

pour plusieurs ordres de contiguïté ou pour des coefficients de contiguïté généralisés (différents de 0 ou 1 et dissymétriques) pour avoir une idée de la structure de l'espace²⁴.

Notons que l'inférence statistique peut également être basée sur l'approche de permutation. Dans cette approche, on suppose, que sous l'hypothèse nulle, chaque valeur observée aurait pu avoir lieu dans toutes les localisations avec une probabilité égale. Mais au lieu d'utiliser la moyenne et l'écart-type théoriques calculés sous l'hypothèse 2, une distribution de référence est empiriquement générée pour les coefficients, à partir de laquelle l'espérance et l'écart-type sont calculés. En pratique, ceci est effectué en permutant les valeurs observées sur toutes les localisations et en recalculant les statistiques pour chaque nouvel échantillon. L'espérance et l'écart-type des statistiques sont alors les moments calculés pour la distribution de référence pour toutes les permutations.

Pour étudier l'autocorrélation spatiale sur les variables quantitatives, deux coefficients sont habituellement utilisés : le coefficient de Moran (1948) et celui de Geary (1954).

213.12 La statistique de Moran

La statistique I de Moran (1948) est la statistique la plus connue d'autocorrélation spatiale (Cliff et Ord, 1981 ; Upton et Fingleton, 1985). Cette statistique, qui apparaît comme une extension de la statistique NN aux variables quantitatives, s'écrit de la façon suivante :

$$I = \frac{N}{S_0} \cdot \frac{\sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_t (x_t - \bar{x})^2} \quad [2.10]$$

où x_i est l'observation dans la région i , \bar{x} est la moyenne des observations des régions, N est le nombre de régions et w_{ij} est l'élément de la matrice de poids, S_0 est un facteur d'échelle égal à la somme de tous les éléments de W .

Le numérateur s'interprète comme la covariance entre unités contiguës, chaque contiguïté étant pondérée par w_{ij} / S_0 . Elle est normalisée par le dénominateur qui est la variance totale observée. L'espérance mathématique de I est égale à $-1/(N-1)$. Par conséquent, il y a autocorrélation spatiale positive lorsque I est supérieur à cette espérance et il y a autocorrélation spatiale négative lorsque I est inférieur.

²⁴ Il est également possible de généraliser ce test à plus de deux couleurs. Dans ce cas, il faut compter le nombre de liens entre régions de même couleur, le nombre de liens entre régions de deux couleurs différentes données, le nombre de liens entre régions de couleurs différentes sans spécifier les couleurs. Cliff et Ord (1981) ont déterminé les expressions de l'espérance mathématique et la variance de chacune de ces quantités pour les hypothèses 1 et 2.

En notant z le vecteur des N observations en déviation à la moyenne \bar{x} , [2.10] se réécrit de la façon suivante en forme matricielle :

$$I = \frac{N}{S_0} \cdot \frac{z'Wz}{z'z} \quad [2.11]$$

Si la matrice de poids W est standardisée, cette expression se simplifie puisque pour les matrices standardisées $S_0 = N$.

A l'origine proposée pour des matrices de contiguïté binaires, cette statistique a été étendue pour toutes les matrices de poids (Cliff et Ord, 1969, 1973) pour résoudre le problème de l'invariance topologique. Les valeurs extrémales du coefficient de Moran dépendent des valeurs propres d'une matrice contenant la matrice de poids W (De Jong *et al.*, 1984 ; Tiefelsdorf et Boots, 1995, 1996). Ce coefficient, à la différence des coefficients de corrélation habituels, n'est pas compris entre -1 et 1 et sa valeur dépend de la structure spatiale (Boots et Tiefelsdorf, 2000).

Les espérances et les variances de la statistique de Moran pour deux hypothèses différentes sont données dans le tableau 2.2 :

- *Hypothèse N* : on suppose que les valeurs x_i prises par la variable aléatoire X sur les différentes régions proviennent de N tirages indépendants d'une population normale.
- *Hypothèse R* : les valeurs x_i sont les réalisations d'une variable aléatoire X dont la distribution est inconnue. Il faut prendre en considération l'ensemble des $N!$ permutations possibles, chacune étant équiprobable.

L'hypothèse N est évidemment plus adaptée si l'on pense que les valeurs de la variable X suivent une loi proche de la loi normale. Si cette hypothèse n'est pas acceptable ou si l'on préfère ne pas se prononcer *a priori* sur ce point, l'hypothèse R est davantage conseillée.

213.13 La statistique de Geary

Un deuxième indicateur est le coefficient de Geary (1954) qui dérive plutôt de la statistique NB . Au lieu d'être construite à partir des produits croisés entre les déviations à la moyenne, la statistique de Geary utilise les carrés des différences entre les valeurs prises par la variable sur les régions voisines. Elle s'écrit de la façon suivante :

$$c = \frac{N-1}{2S_0} \cdot \frac{\sum_i \sum_j w_{ij} (x_i - x_j)^2}{\sum_i (x_i - \bar{x})^2} \quad [2.12]$$

Le terme du numérateur est, au facteur $1/2$ près, la variance pondérée des différences $x_i - x_j$ entre observations contiguës. Cette variance est faible pour une autocorrélation positive et très

élevée en présence d'une autocorrélation négative. En l'absence d'autocorrélation spatiale, le numérateur est approximativement égal à l'estimateur de la variance de l'échantillon, c'est-à-dire le dénominateur. Le coefficient de Geary est alors proche de 1.

Les espérances et les variances de la statistique de Geary sont fournies dans le tableau 2.2 pour les hypothèses N et R définies précédemment.

<p><u>Hypothèse N :</u></p> $E(I) = -1/(N - 1)$ $V(I) = \frac{N^2 S_1 - NS_2 + 3S_0^2}{S_0^2(N^2 - 1)} - E(I)^2$ $E(c) = 1$ $V(c) = \frac{(2S_1 + S_2)(N - 1) - 4S_0^2}{2(N + 1)S_0^2}$ <p><u>Hypothèse R :</u></p> $E(I) = -1/(N - 1)$ $V(I) = N \frac{(N^2 - 3N + 3)S_1 - NS_2 + 3S_0^2}{(N - 1)^3 S_0^2} - b_2 \frac{(N^2 - N)S_1 - 2NS_2 + 6S_0^2}{(N - 1)^3 S_0^2} - E(I)^2$ $E(c) = 1$ $V(c) = \frac{(N - 1)S_1(N^2 - 3N + 3 - (N - 1)b_2) - (1/4)(N - 1)S_2(N^2 + 3N - 6 - (N^2 - N + 2)b_2)}{N(N - 2)^2 S_0^2} + \frac{S_0^2(N^2 - 3 - (N - 1)^2 b_2)}{N(N - 2)^2 S_0^2}$ <p><u>Notations :</u></p> <p>N = nombre total de régions</p> $S_0 = \sum_i \sum_j w_{ij} \quad ; \quad S_1 = (1/2) \sum_i \sum_j (w_{ij} + w_{ji})^2 \quad ;$ $S_2 = \sum_i (w_i + w_i)^2 \quad ; \quad w_i = \sum_j w_{ij} \quad ; \quad w_i = \sum_j w_{ji} \quad ; \quad b_2 = \left(\sum_i (x_i - \bar{x})^4 / n \right) / \left(\sum_i (x_i - \bar{x})^2 / n \right)$

Tableau 2.2 : Espérances et variances des statistiques de Moran et de Geary
 Source : Cliff et Ord (1981, p. 42-46)

Comme pour les statistiques NN et NB , les statistiques de Moran et Geary centrées et réduites sont asymptotiquement distribuées selon des lois normales (0,1) (Cliff et Ord, 1973 ;

Sen, 1976)²⁵. Une troisième possibilité consiste à utiliser la procédure de permutation décrite précédemment pour les statistiques *NN* et *NB*.

Dans la pratique, on utilise plus souvent la statistique de Moran dont les résultats semblent un peu plus stables. Comme pour les statistiques *NN* et *NB*, il est souvent intéressant de calculer les statistiques de Moran et de Geary en prenant en compte diverses matrices de poids. Ainsi, prendre en compte les différents ordres de contiguïté supérieurs à 1 permet de voir comment la variable se comporte dans l'espace, en construisant un corrélogramme qui associe à chaque ordre ou niveau de contiguïté la valeur du coefficient d'autocorrélation (Oden, 1984 ; Upton et Fingleton, 1985). Des corrélogrammes peuvent également être construits à partir de matrices de distance à bande (dans ce cas, un élément w_{ij} est égal à 1 si la distance entre i et j est comprise entre deux seuils, tels que les quartiles de la distribution des distances). De cette façon, la structure de l'autocorrélation peut être inspectée à différentes échelles. Ce schéma permet habituellement de repérer, pour de faibles ordres de contiguïté, des phénomènes d'attraction (autocorrélation spatiale positive) ainsi que des phénomènes de répulsion pour des ordres supérieurs (LeBart, 1969 ; Sokal et Oden, 1978).

213.14 La statistique générale Γ

Toutes les statistiques précédemment décrites peuvent être considérées comme des cas particuliers d'une statistique générale Γ (Cliff et Ord, 1981 ; Hubert *et al.*, 1981, 1985 ; Hubert et Arabie, 1991). Cette statistique indique l'association entre deux matrices de similarité (ou de dissimilarité) pour un ensemble de régions. Elle s'écrit de la façon suivante :

$$\Gamma = \sum_i \sum_j w_{ij} Y_{ij} \quad [2.13]$$

où w_{ij} est un élément de la matrice de poids mesurant la proximité *spatiale* entre les localisations i et j , alors que Y_{ij} est une mesure de la proximité entre i et j sur une autre dimension. Cette statistique englobe les statistiques précédentes :

$$NB = \Gamma / 2 \quad \text{pour} \quad Y_{ij} = (x_i - x_j)^2 \quad [2.14a]$$

$$NN = \Gamma / 2 \quad \text{pour} \quad Y_{ij} = x_i x_j \quad [2.14b]$$

$$I = \Gamma / S_0 s^2 \quad \text{pour} \quad Y_{ij} = (x_i - \mu)(x_j - \mu) \quad \text{et} \quad s^2 = \sum_i (x_i - \mu)^2 / N \quad [2.14c]$$

$$c = \Gamma / 2S_0 \sigma^2 \quad \text{pour} \quad Y_{ij} = (x_i - x_j)^2 \quad \text{et} \quad \sigma^2 = \sum_i (x_i - \mu)^2 / N - 1 \quad [2.14d]$$

²⁵ Cette approximation n'est pas valable pour des configurations très particulières de la structure spatiale, telle qu'une configuration en étoile (Cliff et Ord, 1981, p. 50). En outre, l'approximation normale n'est pas valide pour de petits échantillons de 10 à 50 observations (Cliff et Ord, 1971 ; Terui et Kikuchi, 1994).

213.2 Autocorrélation spatiale locale

Les statistiques I de Moran et c de Geary sont des statistiques globales qui ne permettent pas d'apprécier la structure régionale de l'autocorrélation spatiale. Pourtant, il est possible de se demander, s'il y a des regroupements *locaux* de valeurs faibles ou élevées, quelles régions contribuent le plus à l'autocorrélation spatiale globale et finalement, jusqu'à quel point l'évaluation de l'autocorrélation spatiale globale masque des « localisations atypiques » ou « des poches de non-stationarité locale », c'est-à-dire respectivement des régions ou des groupes de régions qui dévient du schéma global d'autocorrélation spatiale.

Dans cette perspective, l'analyse de l'autocorrélation spatiale locale peut être effectuée en utilisant différentes techniques. Tout d'abord, les statistiques $G_i(d)$ (Getis et Ord, 1992 ; Ord et Getis, 1995) sont souvent utilisées pour étudier la significativité des regroupements spatiaux autour de localisations individuelles (paragraphe 213.21). Ensuite, le diagramme de Moran (Anselin, 1996) est utilisé pour visualiser l'instabilité spatiale locale et les observations extrêmes (paragraphe 213.22). Enfin, les indicateurs locaux d'association spatiale, ou « LISA »²⁶ (Anselin, 1995) sont également utilisés pour tester l'hypothèse d'une distribution aléatoire, en comparant les valeurs de chaque localisation spécifique avec les valeurs des localisations voisines (paragraphe 213.23).

213.21 Les statistiques de Getis-Ord

Getis et Ord (1992) suggèrent l'utilisation des statistiques $G_i(d)$ pour détecter les poches locales de dépendance spatiale qui ne peuvent pas être détectées par l'utilisation de statistiques globales. Ces statistiques sont souvent utilisées pour étudier la significativité des regroupements spatiaux autour de localisations individuelles. Pour chaque région i , cette statistique s'écrit de la façon suivante :

$$G_i(d) = \frac{\sum_{j \neq i} w_{ij}(d)x_j}{\sum_{j \neq i} x_j} \quad [2.15]$$

où $w_{ij}(d)$ sont les éléments d'une matrice de poids symétrique et binaire : ils sont égaux à 1 pour toutes les régions situées à une distance d d'une région donnée i et égaux à 0 pour toutes les autres régions. La variable x a une origine naturelle et est positive. La statistique mesure donc la proportion de la somme de tous les x_j qui sont situés à une distance inférieure à d de la région i .

²⁶ *Local Indicators of Spatial Association.*

Getis et Ord (1992) démontrent que les moments de cette statistique sous l'hypothèse R s'écrivent de la façon suivante :

$$E[G_i(d)] = W_i / N - 1 \quad [2.16a]$$

$$V[G_i(d)] = \frac{W_i(N-1-W_i)}{(N-1)^2(N-2)} \left(\frac{Y_{i2}}{Y_{i1}^2} \right) \quad [2.16b]$$

avec $W_i = \sum_{j \neq i} w_{ij}(d)$; $Y_{i1} = \sum_j x_j / (N-1)$ et $Y_{i2} = \sum_j x_j^2 / (N-1) - Y_{i1}^2$.

Une fois standardisée, une valeur positive de $G_i(d)$ indique un regroupement spatial de valeurs élevées autour de la région i alors qu'une valeur négative indique un regroupement spatial de valeurs faibles autour de la région i .

Cette statistique est étendue aux variables qui n'ont pas une origine naturelle (tels que les résidus d'une régression) et pour des matrices standardisées qui ne sont pas binaires ou symétriques (Ord et Getis, 1995). Cette statistique a alors l'expression suivante :

$$G_i(d) = \frac{\sum_{j \neq i} w_{ij}(d)x_j - W_i \bar{x}(i)}{\sigma(i) \left\{ [(N-1)S_{1i} - W_i^2] / (N-2) \right\}^{1/2}}, \quad j \neq i \quad [2.17]$$

où $W_i = \sum_{j \neq i} w_{ij}(d)$ et $S_{1i} = \sum_{j \neq i} w_{ij}^2$, $\bar{x}(i)$ et $\sigma(i)$ sont respectivement la moyenne et l'écart-type pour l'échantillon de taille $N-1$ en excluant la région i . Le signe de cette statistique s'interprète de la même façon que précédemment ²⁷.

Pour ces différentes statistiques, l'inférence est basée sur une approximation normale (Getis et Ord, 1992). Pourtant, ces auteurs admettent qu'il puisse exister un problème en présence d'autocorrélation spatiale globale (Ord et Getis, 1995, p. 298-299). En outre, l'inférence statistique est encore compliquée par le fait que les statistiques locales entre deux régions sont corrélées lorsque le voisinage de ces deux régions contient des éléments communs (Anselin, 1995 ; Ord et Getis, 1995). Il s'agit d'un problème de comparaisons multiples (Savin, 1984). Comme le souligne Anselin (1995, p. 96), cela signifie que lorsque la significativité totale associée avec les comparaisons multiples (tests corrélés) est fixée à α , et qu'il y a m comparaisons, alors la significativité individuelle α_i devrait être fixée à α/m (Bonferroni) ou à $1 - (1 - \alpha)^{1/m}$ (Sidák). La seconde procédure nécessite que les variables suivent une loi normale multivariée, ce qui est asymptotiquement le cas pour $G_i(d)$.

²⁷ La statistique, qui inclue la valeur prise par la variable dans la région i , a également été suggérée par Getis et Ord (1992) et Ord et Getis (1995).

213.22 *Le diagramme de Moran*

Le diagramme de Moran (Anselin, 1996) est utilisé pour visualiser les formes de l'autocorrélation spatiale locale, examiner l'instabilité locale spatiale et inspecter les observations extrêmes. Dans ce diagramme figure en abscisse la valeur standardisée de la variable, notée z , et en ordonnée son décalage spatial standardisé Wz . Les quatre différents quadrants du diagramme correspondent alors aux quatre différents types d'association spatiale locale existant entre une région et ses voisines (figure 2.2).

- HH : une région associée à une valeur élevée ²⁸ entourée de régions associées à des valeurs élevées.
- BH : une région associée à une valeur faible entourée de régions associées à des valeurs élevées.
- BB : une région associée à une valeur faible entourée de régions associées à des valeurs faibles.
- HB : une région associée à une valeur élevée entourée de régions associées à des valeurs faibles.

Les quadrants HH et BB sont associés à une autocorrélation spatiale positive car ils indiquent un regroupement spatial de valeurs similaires. En revanche, les quadrants BH et HB représentent une autocorrélation spatiale négative car ils indiquent un regroupement spatial de valeurs dissemblables. Par conséquent, le diagramme de Moran peut être utilisé pour visualiser les localisations atypiques, c'est-à-dire les régions qui se trouvent dans les quadrants BH et HB.

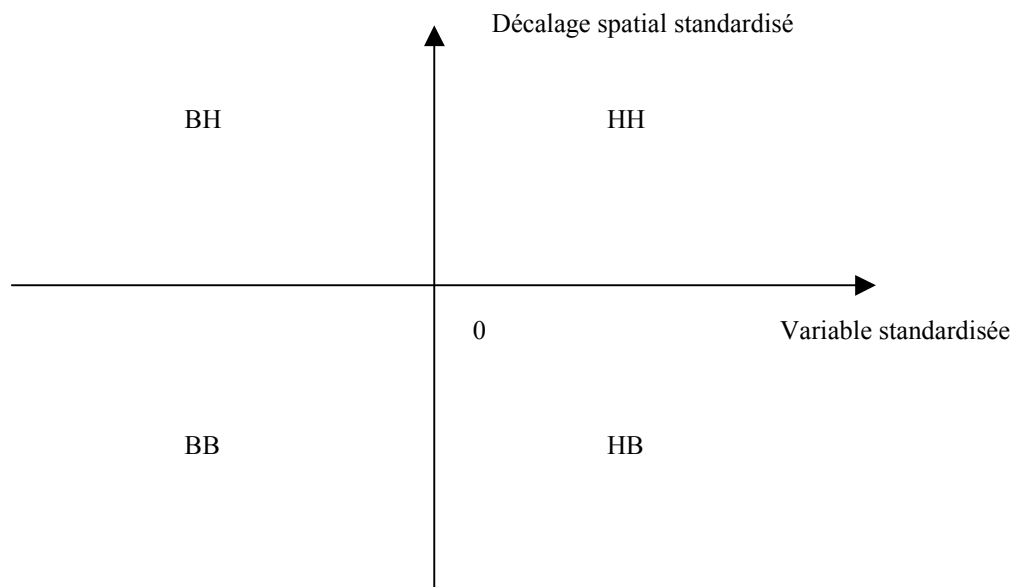


Figure 2.2 : Diagramme de Moran

²⁸ Elevé (resp. faible) signifie au-dessus (resp. en dessous) de la moyenne.

L'autocorrélation spatiale globale peut également être visualisée sur ce graphique, puisque, d'après [2.11], la statistique I de Moran est formellement équivalente à la pente de la régression linéaire de Wz sur z en utilisant une matrice de poids standardisée. La détection des observations extrêmes qui exercent une forte influence sur le I de Moran peut alors être basée sur les diagnostics standards de régression (Belsley *et al.*, 1980 ; Donald et Maddala, 1993).

Le diagramme de Moran ne fournit pas d'information sur la significativité des regroupements spatiaux. Celle-ci est obtenue à l'aide d'indicateurs locaux d'association spatiale.

213.23 Les indicateurs locaux d'association spatiale (LISA)

Anselin (1995) définit un indicateur local d'association spatiale ou « LISA » (*Local Indicator of Spatial Association*) comme toute statistique satisfaisant deux critères. Premièrement, pour chaque observation, le LISA donne une indication sur le regroupement spatial significatif de valeurs similaires autour de chaque observation. Deuxièmement, la somme des LISA associés à toutes les observations est proportionnelle à un indicateur global d'association spatiale.

Anselin fournit une version locale de la statistique Γ , du I de Moran et du c de Geary. Seule la deuxième est utilisée en pratique. Pour chaque région i elle s'écrit de la façon suivante :

$$I_i = \frac{(x_i - \bar{x})}{m_0} \sum_j w_{ij} (x_j - \bar{x}) \quad \text{avec} \quad m_0 = \sum_i (x_i - \bar{x})^2 / N \quad [2.18]$$

où x_i est l'observation pour la région i , \bar{x} est la moyenne des observations. La sommation sur j est telle que seules les valeurs voisines j de i sont incluses. Compte tenu de cette définition, la somme des statistiques locales peut s'écrire de la façon suivante :

$$\sum_i I_i = \frac{1}{m_0} \sum_i (x_i - \bar{x}) \sum_j w_{ij} (x_j - \bar{x}) = \frac{1}{m_0} \sum_i \sum_j w_{ij} (x_i - \bar{x})(x_j - \bar{x}) \quad [2.19]$$

De [2.10], il s'ensuit que la statistique globale de Moran est proportionnelle à la somme des statistiques locales de Moran :

$$I = \sum_i I_i / S_0 \quad [2.20]$$

Pour une matrice de poids standardisée en lignes, $S_0 = N$ de telle sorte que $I = (1/N) \sum_i I_i$: la statistique globale I de Moran est égale à la moyenne des statistiques locales de Moran.

Une valeur positive pour I_i indique le regroupement spatial de valeurs similaires (faibles ou élevées) alors qu'une valeur négative indique le regroupement spatial de valeurs dissemblables entre une région et ses voisines.

Cette statistique est basée sur des covariances spatiales plutôt que sur des accumulations spatiales. Elle mesure donc une forme d'association spatiale locale différente de celle donnée par Getis et Ord (1992, 1995).

Sous l'hypothèse R , les moments de I_i s'écrivent de la façon suivante :

$$E[I_i] = -W_i / (N - 1) \quad [2.21a]$$

$$V[I_i] = [W_{i(2)}(N - b_2) / (N - 1)] + [2W_{i(kh)}(2b_2 - N) / (N - 1)(N - 2)] - [W_i^2 / (N - 1)^2] \quad [2.21b]$$

où $W_i = \sum_j w_{ij}$, $b_2 = m_4 / m_2^2$, $m_2 = \sum_i (x_i - \bar{x})^2 / N$, $m_4 = \sum_i (x_i - \bar{x})^4 / N$,
 $W_{i(2)} = \sum_{j \neq i} w_{ij}^2$, $2W_{i(kh)} = \sum_{k \neq i} \sum_{h \neq i} w_{ik} w_{ih}$

Un test basé sur ces moments pourrait être dérivé mais la distribution de la statistique standardisée qui en résulterait est inconnue, au contraire des statistiques de Getis-Ord. En outre, comme le montrent les simulations effectuées par Boots et Tiefelsdorf (2000) sur configurations spatiales régulières, la distribution pour ces statistiques ne peut pas être approximée par une distribution normale. L'inférence statistique doit alors être basée sur l'approche de permutation (Anselin, 1995)²⁹. Dans ce cas, les probabilités critiques obtenues pour les statistiques locales de Moran sont, en fait, des pseudo-niveaux de significativité. De plus, comme la normalité est peu probable pour les LISA, il faut utiliser la correction de Bonferroni au lieu de la correction de Sidák.

Anselin (1995) donne deux interprétations pour les LISA. Premièrement, ils peuvent être utilisés comme indicateurs de regroupements spatiaux locaux, de la même façon que les statistiques $G_i(d)$. Deuxièmement, ils permettent de diagnostiquer l'instabilité locale (les localisations atypiques ou les poches de non-stationarité), les observations extrêmes et les régimes spatiaux. Cette seconde interprétation est similaire à l'utilisation du diagramme de Moran pour identifier les observations extrêmes pour le I de Moran : puisqu'il y a un lien entre les indicateurs locaux et la statistique globale, les observations extrêmes pour les LISA seront associées aux régions qui contribuent le plus à l'autocorrélation spatiale globale. Finalement, en combinant l'information donnée par le diagramme de Moran et la significativité des LISA, on obtient des cartes de significativité de Moran, qui montrent les

²⁹ Il s'agit alors d'une approche en termes de permutation conditionnelle dans le sens où la valeur y_i pour la localisation i est maintenue fixée pendant que les valeurs restantes sont permutées sur toutes les autres localisations de l'échantillon.

régions associées avec un LISA significatif et indiquent par un code de couleur les quadrants du diagramme de Moran auquel ces régions appartiennent (Anselin et Bao, 1997).

213.3 Quelques précautions méthodologiques

D'une façon générale, la mise en œuvre de ces différents tests d'autocorrélation spatiale, ainsi que les estimations et les autres tests présentés dans les sections suivantes, nécessitent quelques précautions méthodologiques liées à l'utilisation de données spatiales. Deux grandes difficultés se posent en particulier : la dépendance à la définition de la structure spatiale des interactions (paragraphe 213.31) et l'agrégation (paragraphe 213.32).

213.31 La dépendance à la définition de la matrice de poids

La modélisation des interdépendances spatiales se fait de diverses façons : ordres de contiguïté différents, absence ou présence de pondérations, etc. En conséquence, rejeter l'absence ou la présence d'autocorrélation spatiale pour une définition particulière du voisinage n'implique pas toujours qu'on aboutisse à la même conclusion avec d'autres définitions du voisinage.

Un exemple classique est donné par Cliff et Ord pour une représentation spatiale régulière (Cliff et Ord, 1973, p. 16-19). Ils considèrent un système de 16 régions auquel les trois définitions de la contiguïté précédemment définies sont associées : tour, fou et dame. Soit la figure suivante :

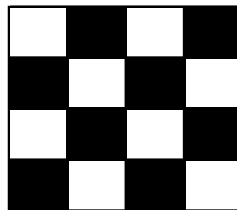


Figure 2.3 : Echiquier

Source : Cliff et Ord (1973, p. 16)

Dans le cas de la tour, où les déplacements se font sur une même ligne horizontale ou verticale, l'analyse de la figure montre que l'on trouve une autocorrélation spatiale négative puisqu'il y a alors alternance régulière entre carrés noirs et carrés blancs. La conclusion est inversée pour le cas du fou : les carrés de même couleur sont regroupés entre eux, il y a donc autocorrélation spatiale positive. Enfin, dans le cas de la dame, où les déplacements se font à la fois sur les lignes et sur les diagonales, on conclut à une absence d'autocorrélation spatiale. Bien sûr, cet exemple est un cas extrême - rares sont les cas où les tests divergeront autant pour un même ensemble d'observations - mais il montre comment les conclusions sont susceptibles de varier selon la notion de contiguïté choisie.

Dans ces conditions, puisque la mesure de l'autocorrélation spatiale est dépendante de la matrice de poids utilisée, il est donc utile d'évaluer, dans la mesure du possible, la robustesse des résultats obtenus au choix de la matrice de poids.

213.32 Le problème « MAUP »

On observe fréquemment que les différentes mesures d'autocorrélation spatiale sont sensibles à la façon dont les données sont agrégées (Openshaw et Taylor, 1979 ; Arbia, 1989 ; Fotheringham et Wong, 1991). Dans la mesure où souvent, seules des données observées sur des divisions administratives sont disponibles, de l'autocorrélation spatiale peut apparaître, non pas parce qu'elle est le résultat d'effets de proximité ou d'externalités, mais parce que les divisions considérées sont arbitraires. Plus précisément, le *Modifiable Areal Unit Problem*, ou MAUP recouvre deux aspects.

Premièrement, la mesure de l'autocorrélation spatiale est affectée par le niveau d'agrégation utilisé, c'est l'effet d'échelle. Ainsi, les résultats peuvent varier selon qu'on utilise des observations sur les régions ou sur les départements français. Par exemple, la figure 2.4 illustre comment deux niveaux d'agrégation différents affectent la valeur de la statistique de Moran : on aboutit à une autocorrélation spatiale négative dans le cas agrégé et à une autocorrélation spatiale faiblement positive dans le cas désagrégé. Dans ce cas, il est préférable de disposer d'informations sur « l'aire pertinente » associée au phénomène étudié. Dans l'illustration suivante, si on pense que l'aire de définition du phénomène est la région, plutôt que le département, alors la structure agrégée est plus pertinente que la structure désagrégée. Il s'en suit que le territoire national décrit par la figure 2.4. est caractérisé par une autocorrélation négative.



Figure 2.4 : Autocorrélation spatiale et effet d'échelle

Source : Chou (1991, p. 230)

Deuxièmement, l'autocorrélation spatiale est également sensible à la forme des unités spatiales. En effet, une région peut être découpée en plusieurs subdivisions de multiples

manières, ce qui donne lieu à de nombreuses configurations spatiales. Griffith (1992b) montre alors comment la statistique de Moran est affectée par les différentes façons dont les données sont agrégées (figure 2.5).

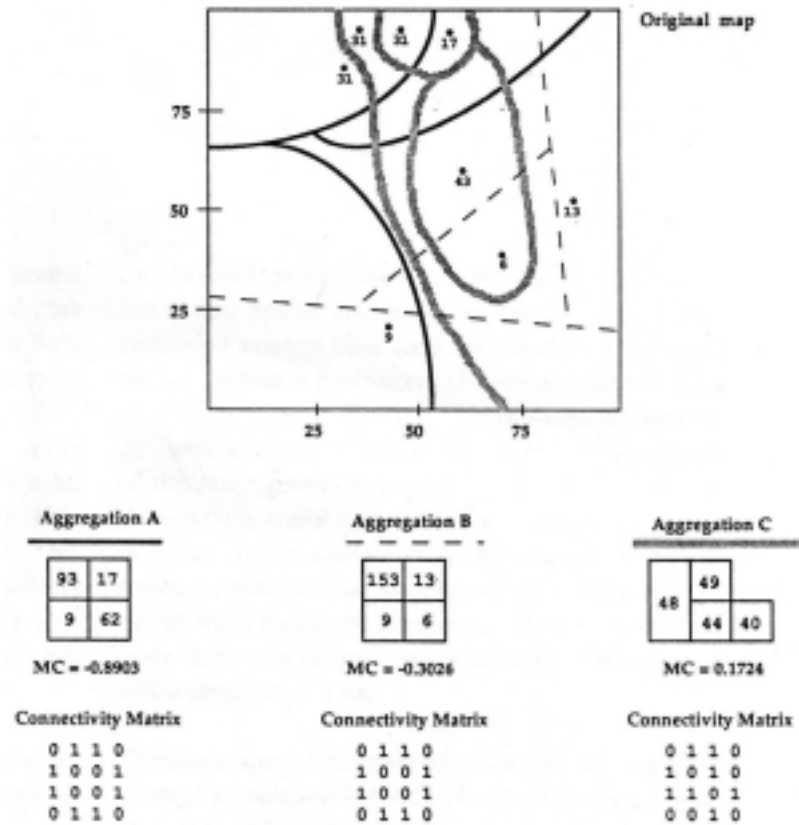


Figure 2.5 : Autocorrélation spatiale et effet de forme
Source : Griffith (1992b, p. 277)

Conclusion de section

Dans cette section, nous avons défini les deux grands effets caractérisant les données spatiales. D'une part, l'autocorrélation spatiale est définie par l'absence d'indépendance entre observations géographiques. Les matrices de poids permettent de modéliser la structure spatiale des données et d'aboutir à la notion de variable spatiale décalée. Cette dernière est d'une importance fondamentale pour introduire l'autocorrélation spatiale dans les modèles économétriques. D'autre part, l'hétérogénéité spatiale se réfère à la différenciation géographique des comportements.

Nous abordons maintenant la façon dont le premier effet est introduit dans les modèles économétriques et quelles sont les conséquences de cette introduction.

SECTION 22 :

AUTOCORRELATION SPATIALE

Sur le plan économétrique, la non-vérification de l'hypothèse d'indépendance des observations conduit à remettre en cause l'inférence statistique basée sur l'estimation d'une régression par les MCO. Le test de l'autocorrélation spatiale s'avère alors être une précaution indispensable à toute étude empirique mobilisant des données géographiques. La démarche à adopter s'articule autour de deux opérations : détecter et modéliser l'autocorrélation spatiale.

Les tests de l'autocorrélation spatiale permettent à la fois de détecter la présence ainsi que la forme de l'autocorrélation dans les régressions, orientant ainsi l'économiste vers la spécification économétrique appropriée. Cependant, dans les études empiriques, on peut aussi disposer *a priori* d'informations théoriques conduisant à préférer une forme particulière de dépendance spatiale entre les observations pour capter différents types d'effets de débordement géographiques. Ainsi, les tests et la modélisation de l'autocorrélation spatiale apparaissent être des opérations interdépendantes.

Dans un premier temps, nous présentons les trois spécifications les plus utilisées ainsi que leurs principales propriétés (paragraphe 221), puis nous abordons les méthodes d'estimation adaptées à ces spécifications spatiales : maximum de vraisemblance, variables instrumentales et moments généralisés (paragraphe 222). Nous exposons ensuite les méthodes d'inférence permettant de tester la présence d'autocorrélation spatiale dans les modèles (paragraphe 223). Les trois principes de tests basés sur le maximum de vraisemblance (multiplicateur de Lagrange, Wald et rapport de vraisemblance) peuvent être mobilisés mais les tests du multiplicateur de Lagrange sont particulièrement utilisés en économétrie spatiale car ils permettent, à travers diverses règles de décision, de déterminer la forme prise par l'autocorrélation spatiale. Ces tests ont en particulier été développés par Anselin (1988a, 1988b). En se basant sur les propriétés des différents tests, nous proposons alors une démarche générale visant à déterminer la forme prise par l'autocorrélation spatiale dans les modèles de régression (paragraphe 224). Enfin, les développements précédents étant adaptés aux modèles linéaires en coupe transversale, nous présentons certaines extensions à la prise en compte de l'autocorrélation spatiale dans d'autres types de modèles (paragraphe 225).

221 Autocorrélation spatiale et modèles économétriques

Comme point de départ, considérons le modèle de régression linéaire classique en coupe transversale :

$$y = X\beta + \varepsilon \quad [2.22]$$

Dans le cadre de cette thèse, nous adopterons les conventions d'écriture suivantes :

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

Les hypothèses du modèle linéaire général sont les suivantes :

H_1 : *Sur les variables explicatives.*

X est une matrice non-stochastique de rang complet $K \leq N$ et lorsque la taille de l'échantillon devient infiniment grand, $\lim_{n \rightarrow \infty} (1/N)X'X = Q$ où Q est une matrice finie et non-singulière.

H_2 : *Sur les erreurs.*

Le vecteur d'erreur consiste en des erreurs non-observables qui satisfont les propriétés $E(\varepsilon) = 0$ et $E(\varepsilon\varepsilon') = \sigma^2 I_N$, où $E(\cdot)$ désigne l'espérance mathématique et I_N la matrice identité d'ordre N .

Si ces différentes conditions sont vérifiées, alors l'estimateur des MCO défini par $\hat{\beta} = (X'X)^{-1}X'y$ est un estimateur BLUE (théorème de Gauss-Markov) : il s'agit de l'estimateur centré et à variance minimale dans la classe des estimateurs linéaires.

L'introduction de l'autocorrélation spatiale dans le modèle de régression linéaire conduit au relâchement de certaines de ces hypothèses. Cette introduction peut s'effectuer de plusieurs manières : par des *variables spatiales décalées*, endogènes (paragraphe 221.1), ou exogènes (paragraphe 221.2), ou par une *autocorrélation spatiale des erreurs* (paragraphe 221.3). Enfin, des processus plus complexes dans lesquels ces deux aspects sont combinés sont également envisageables (paragraphe 221.4).

221.1 Le modèle autorégressif spatial : variable endogène décalée

La première façon de prendre en compte l'autocorrélation spatiale peut se faire grâce au modèle autorégressif spatial : une « variable endogène décalée » est incluse dans le modèle [2.22]. Le modèle structurel s'écrit dans ce cas de la façon suivante :

$$y = \rho W y + X\beta + \varepsilon \quad [2.23]$$

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

Dans ce modèle, l'observation y_i est, en partie, expliquée par les valeurs prises par y dans les régions voisines : $(Wy)_i = \sum_{j \neq i} w_{ij} y_j$. En effet, $(Wy)_i$ s'interprète comme la moyenne des valeurs de y sur les observations voisines à i lorsque W est standardisée. Cette standardisation facilite aussi la comparaison de l'ampleur de ρ lorsque [2.23] est estimé pour plusieurs matrices de poids.

L'introduction de Wy dans le modèle [2.22] est un moyen d'apprécier le degré de dépendance spatiale alors que les autres variables sont contrôlées. Symétriquement, il permet de contrôler la dépendance spatiale pour évaluer l'impact des autres variables explicatives. Lorsqu'une variable endogène décalée est ignorée dans la spécification du modèle, mais présente dans le processus générateur des données, les estimateurs des MCO dans le modèle aspatial [2.22] sont biaisés et non convergents.

Cette spécification possède plusieurs propriétés.

1/ Filtrage spatial de la variable dépendante

Le modèle [2.23] se réécrit de la façon suivante :

$$(I - \rho W)y = X\beta + \varepsilon \quad [2.24]$$

Dans ce modèle : $y^* = (I - \rho W)y$ s'interprète comme la variable dépendante « filtrée », dans laquelle les effets de l'autocorrélation spatiale ont été éliminés.

2/ Effets de multiplicateur et de diffusion

Supposons la matrice $(I - \rho W)$ non-singulière. Ceci est vrai lorsque ρ est différent de 0 et lorsque l'inverse de ρ n'est pas une valeur propre de W ³⁰. Dans ce cas, [2.24] se réécrit sous la forme réduite suivante :

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \quad [2.25]$$

L'espérance mathématique de y s'écrit alors : $E(y) = (I - \rho W)^{-1} X\beta$.

³⁰ En effet, $(I - \rho W)^{-1}$ existe si et seulement si $|I - \rho W| \neq 0$. Cette condition est équivalente à : $|\rho| |W - (1/\rho)I| \neq 0$ ou $|\rho| \neq 0$ et $|W - (1/\rho)I| \neq 0$.

La matrice inverse $(I - \rho W)^{-1}$ est une matrice pleine qui, lorsque $w_{ij} < 1$ et $|\rho| < 1$, implique une série infinie pour les variables explicatives et pour le terme d'erreur associé à toutes les localisations :

$$y = (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) X \beta + (I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots) \varepsilon \quad [2.26]$$

A partir de cette expression, on peut définir deux types d'effets : *un effet de multiplicateur spatial* affectant les variables explicatives et *un effet de diffusion spatiale* affectant les erreurs.

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

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

3/ Matrice des variances-covariances

De [2.25], on déduit la matrice des variances-covariances de y :

$$V(y) = (I - \rho W)^{-1} E(\varepsilon \varepsilon') (I - \rho W')^{-1} \quad [2.27a]$$

$$V(y) = \sigma^2 [(I - \rho W)' (I - \rho W)]^{-1} \quad [2.27b]$$

Cette matrice des variances-covariances est une matrice pleine, ce qui implique que chaque localisation est corrélée avec chaque autre localisation du système mais que cette corrélation diminue avec l'ordre de proximité.

4/ Variable endogène décalée et hétéroscédasticité

Posons $u = (I - \rho W)^{-1} \varepsilon$. La matrice de variances-covariances de u s'écrit alors :

$$V(u) = \sigma^2 [(I - \rho W)' (I - \rho W)]^{-1} \quad [2.28]$$

De [2.28], il s'ensuit que la covariance entre chaque paire d'erreurs est non nulle et décroissante, mais décroît avec l'ordre de proximité. Par ailleurs, les éléments de la diagonale de $V(u)$ ne sont pas constants, ce qui implique l'hétéroscédasticité des erreurs u , que ε soit hétéroscédastique ou non (McMillen, 1992).

Ce modèle autorégressif est utilisé, par exemple, pour modéliser les interactions stratégiques et la concurrence fiscale entre communes (Case *et al.*, 1993 ; Brueckner, 1998), les externalités de voisinage dans les modèles hédoniques de prix immobiliers (Can, 1990, 1992 ; Macedo, 1998), les externalités spatiales influençant la distribution de la population dans les espaces urbains (Griffith, 1981 ; Griffith et Can, 1996) ou influençant la production de firmes concurrentes (Griffith, 1999).

221.2 Modèle régressif croisé : variable exogène décalée

Une autre façon de traiter l'interdépendance des observations est d'inclure une ou plusieurs « *variables exogènes décalées* » dans [2.22] :

$$y = X\beta + WZ\delta + \varepsilon \quad [2.29]$$

Z est une matrice de dimension (N, L) contenant les L variables correspondant ou non aux variables incluses dans X , WZ est l'ensemble des variables exogènes décalées pour la matrice de poids W et δ est le vecteur $(L, 1)$ de paramètres spatiaux indiquant l'intensité de la corrélation spatiale existant entre les observations de y et celles de Z .

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

Au contraire du modèle autorégressif [2.23] et des modèles comportant une autocorrélation spatiale des erreurs (voir *infra*), l'estimation du modèle régressif croisé peut être basée sur les MCO si H_2 est vérifiée et si H_1 est vérifiée pour la matrice des variables explicatives $X^* = [X \ WZ]$ de dimension $(N, K+L)$.

221.3 Modèle avec autocorrélation spatiale des erreurs

Une dernière façon d'incorporer l'autocorrélation spatiale dans un modèle de régression est de spécifier un processus spatial pour les erreurs. Différentes possibilités existent qui mènent à différentes covariances des erreurs, avec des implications variables pour la nature et l'ampleur de l'autocorrélation spatiale dans le modèle (Anselin, 2001c).

221.31 Les erreurs suivent un processus autorégressif

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

Le modèle structurel s'écrit alors :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \end{aligned} \quad [2.30]$$

Le paramètre λ reflète l'intensité de l'interdépendance entre les résidus de la régression, u est le terme d'erreur tel que : $u \sim \text{iid}(0, \sigma^2 I)$. Omettre à tort une autocorrélation spatiale des erreurs produit des estimateurs non biaisés mais inefficients et l'inférence statistique basée sur les MCO est biaisée.

Cette spécification possède plusieurs propriétés.

1/ Diffusion spatiale

Tout d'abord, si la matrice $(I - \lambda W)$ est non-singulière, le modèle [2.30] se réécrit sous la forme réduite suivante :

$$y = X\beta + (I - \lambda W)^{-1}u \quad [2.31]$$

Cette expression fait apparaître un effet de diffusion spatiale comme pour le modèle [2.23] mais comme $E(y) = X\beta$, il n'y a pas d'effet de multiplicateur spatial.

2/ Filtrage spatial de la variable dépendante et des variables explicatives

Le modèle [2.30] peut aussi être exprimé en termes de variables spatialement filtrées. Après avoir prémultiplié les deux termes de [2.31] par $(I - \lambda W)$, on obtient l'expression suivante :

$$(I - \lambda W)y = (I - \lambda W)X\beta + u \quad [2.32]$$

Ce modèle de régression est un modèle où la variable expliquée ainsi que les variables explicatives sont filtrées spatialement, le terme d'erreur u étant homoscédastique.

3/ Matrice des variances-covariances

De la formulation [2.31], il s'ensuit que :

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

$$V(y) = V(\varepsilon) = \sigma^2 [(I - \lambda W)'(I - \lambda W)]^{-1} \quad [2.33b]$$

On trouve, pour ε et pour y , une structure identique à celle du modèle spatial autorégressif [2.23]. Par conséquent, la covariance entre chaque paire d'erreurs et d'observations est non nulle et décroissante : elle décroît avec l'ordre de proximité. La structure spatiale de covariance induite par le modèle avec autocorrélation des erreurs est donc *globale*, puisqu'elle relie toutes les localisations du système à toutes les autres (Mur, 1999 ; Anselin, 2001c).

Par ailleurs, la structure d'erreurs [2.33b] induit des éléments de la diagonale de $V(\varepsilon)$ non constants, ce qui implique l'hétéroscédasticité des erreurs ε , que u soit hétéroscédastique ou non (McMillen, 1992).

4/ *Modèle de Durbin spatial*

Le modèle [2.30] peut se réécrire d'une façon qui fait apparaître à la fois une variable endogène décalée et l'ensemble des variables exogènes décalées. En effet, en déplaçant le terme autorégressif de [2.32] à droite, on obtient le modèle de Durbin spatial contraint :

$$y = \lambda W y + X \beta - \lambda W X \beta + u \quad [2.34]$$

L'équivalence entre [2.30] et [2.34] impose un ensemble de contraintes non-linéaires sur les coefficients. Ces contraintes non-linéaires font l'objet d'un test (cf. paragraphe 224.1, test du facteur commun, p. 146).

La détection de l'autocorrélation spatiale des erreurs s'interprète souvent comme un problème dans la spécification du modèle, comme l'omission de variables pertinentes : l'effet qui n'est pas capté dans les variables explicatives se retrouve dans les erreurs sous la forme d'une autocorrélation spatiale. Dans cette perspective, elle est un instrument permettant de repérer l'existence de variables significatives, mais non prises en compte dans le modèle. Si d'autres variables explicatives pertinentes ne peuvent être trouvées, elle est un substitut à ces variables omises. Ce modèle permet d'améliorer les estimations des lois de Kaldor (Fingleton et McCombie, 1998), celles des modèles de prix hédoniques (Pace et Gilley, 1997 ; Dubin, 1998 ; Dubin *et al.*, 1999), ou de convergence conditionnelle (Fingleton, 1999). La propriété de diffusion spatiale a aussi été employée pour modéliser la diffusion d'un choc aléatoire sur le PIB d'une région vers les autres régions aux Etats-Unis (Rey et Montouri, 1999).

221.32 *Deux alternatives*

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

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

$$\varepsilon = \gamma W u + u \quad [2.35]$$

où γ est le coefficient moyenne mobile et u est un terme d'erreur homoscédastique. Contrairement au modèle précédent, le modèle réduit ne contient pas de matrices inverses puisque [2.35] correspond déjà au modèle réduit. Ainsi, la matrice des variances-covariances qui résulte de ce processus est la suivante :

$$E[\varepsilon\varepsilon'] = \sigma^2(I + \gamma W)(I + \gamma W') = \sigma^2[I + \gamma(W + W') + \gamma^2 WW'] \quad [2.36]$$

Contrairement à la matrice des variances-covariances associée au processus autorégressif, [2.36] n'est pas une matrice pleine. Les covariances non nulles n'existent que pour les voisins de premier ordre ($W + W'$) et les voisins de second ordre (WW'). Ce processus implique donc une interaction moins forte par rapport au processus autorégressif. Contrairement au modèle précédent, la structure spatiale de covariance induite par le modèle [2.35] n'est alors que *locale*, puisqu'elle ne relie pas toutes les localisations du système à toutes les autres (Mur, 1999 ; Anselin, 2001c).

Enfin, sauf cas particulier, les éléments de la diagonale de [2.36] ne sont pas constants, ce qui provoque, comme dans le modèle précédent, une hétéroscedasticité induite dans ε , quelle que soit la nature de u .

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

$$\varepsilon = Wu + v \quad [2.37]$$

où u et v sont supposées être des erreurs homoscédastiques et indépendantes. Dans ces conditions, la matrice de variances-covariances de ε est alors :

$$E[\varepsilon\varepsilon'] = \sigma_v^2 I + \sigma_u^2 WW' = \sigma^2 (I + \varphi WW') \quad [2.38]$$

où σ_u^2 et σ_v^2 sont les variances respectivement associées à u et v , $\sigma^2 = \sigma_v^2 > 0$ et $\varphi = \sigma_u^2 / \sigma_v^2$. L'interaction spatiale impliquée par [2.38] est encore plus limitée que dans [2.37] car elle ne concerne que les voisins de premier et de second ordre contenus dans les éléments non nuls de WW' et ceci sans qu'il y ait de décroissance par rapport à l'ordre de proximité (Anselin et Bera, 1998).

221.4 Généralisation

Jusqu'à présent, les modèles présentés n'incluent qu'un seul type de dépendance spatiale (variable endogène décalée ou autocorrélation des erreurs) et qu'un seul ordre de

dépendance. Différents auteurs suggèrent des processus beaucoup plus généraux, comportant à la fois une variable autorégressive et une autocorrélation des erreurs (Brandsma et Kelletaper, 1979 ; Huang, 1984 ; Anselin, 1988a ; Jayet, 1993).

La forme la plus générale est la forme autorégressive moyenne mobile (*Spatial AutoRegressive Moving Average* ou SARMA(p,q)) proposée par Huang (1984) :

$$\begin{aligned} y &= \rho_1 W_1 y + \rho_2 W_2 y + \dots + \rho_p W_p y + \varepsilon \\ \varepsilon &= \lambda_1 W_1 u + \lambda_2 W_2 u + \dots + \lambda_q W_q u + u \end{aligned} \quad [2.39]$$

En général, les différentes matrices de poids W_i sont associées au $i^{\text{ème}}$ ordre de contiguïté. On pourrait envisager de la même façon un processus où les erreurs suivent un processus spatial autorégressif d'ordre q .

Anselin (1980, 1988a) étudie les propriétés du processus contenant à la fois une variable endogène décalée et une autocorrélation spatiale des erreurs. Formellement, le modèle s'exprime comme une combinaison de [2.23] et de [2.30], avec des poids différents ([2.40a] est la forme structurelle alors que [2.40b] est la forme réduite) :

$$\begin{cases} y = \rho W_1 y + X\beta + \varepsilon \\ \varepsilon = \lambda W_2 \varepsilon + u \end{cases} \quad [2.40a]$$

$$\Rightarrow y = (I - \rho W_1)^{-1} X\beta + (I - \rho W_1)^{-1} (I - \lambda W_2)^{-1} u \quad [2.40b]$$

Le premier terme multiplié par λW_2 donne : $\lambda W_2 y = \rho \lambda W_2 W_1 y + \lambda W_2 X\beta + \lambda W_2 \varepsilon$. Si on soustrait l'équation obtenue de l'équation initiale :

$$y = \rho W_1 y + \lambda W_2 y - \rho \lambda W_2 W_1 y + X\beta - \lambda W_2 X\beta + u \quad [2.41]$$

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

$$y = (\rho + \lambda) W y - \rho \lambda W^2 y + X\beta - \lambda W X\beta + u \quad [2.42]$$

Dans ce modèle, le paramètre ρ est suridentifié. En revanche, le modèle SARMA (1,1) ne souffre pas de ce problème.

Pratiquement, ce problème d'identification n'est pas d'une très grande portée car, en général, on retient rarement un modèle avec les deux types d'effets spatiaux. On cherche plutôt à modéliser la dépendance spatiale, soit par l'autocorrélation des erreurs, soit par une variable spatiale autorégressive, et non les deux. Anselin et Bera (1998) considèrent d'ailleurs ces processus d'ordre supérieur comme le résultat d'une matrice de poids mal spécifiée et non de processus générateurs des données réalistes. Par exemple, si la matrice de poids du modèle

autorégressif sous-estime la vraie interaction spatiale dans les données, il y aura une autocorrélation spatiale résiduelle des erreurs. Cela peut mener à estimer un processus d'ordre supérieur alors que seule une matrice de poids bien spécifiée serait nécessaire.

Les modèles présentés dans ce paragraphe intègrent l'autocorrélation spatiale et permettent ainsi de modéliser les effets de débordement géographiques. Cependant, sur le plan économétrique, cette introduction a pour conséquence de rendre la méthode des MCO inadaptée. D'autres méthodes d'estimation ont été proposées et nous les présentons dans le paragraphe suivant.

222 Estimation des modèles spatiaux

En présence d'autocorrélation spatiale, les estimateurs obtenus par les MCO ne sont pas convergents (cas d'une variable endogène décalée) et sont inefficients (cas d'une autocorrélation spatiale des erreurs). Nous le montrons dans le paragraphe 222.1. Pour trouver des estimateurs convergents et asymptotiquement normaux et efficaces, la méthode la plus couramment utilisée est celle du maximum de vraisemblance (paragraphe 222.2). Il est également possible d'avoir recours à la méthode des variables instrumentales ou à celle des moments généralisés (paragraphe 222.3).

222.1 La non-convergence des MCO et ses conséquences

Considérons le modèle spatial le plus général, incluant à la fois une variable endogène décalée et une autocorrélation spatiale des erreurs (modèle [2.40]). Ce modèle se réécrit de la façon suivante :

$$\begin{aligned} y &= (I - \rho W_1)^{-1} X\beta + (I - \rho W_1)^{-1} \varepsilon \\ \varepsilon &= (I - \lambda W_2)^{-1} u \end{aligned} \quad [2.43]$$

Par conséquent, la matrice des variances-covariances de l'erreur ε est :

$$V(\varepsilon) = E(\varepsilon\varepsilon') = \sigma^2[(I - \lambda W_2)'(I - \lambda W_2)]^{-1} \quad [2.44]$$

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

$$E(W_1 y \varepsilon') = E[W_1(I - \rho W_1)^{-1} X\beta \varepsilon' + W_1(I - \rho W_1)^{-1} \varepsilon \varepsilon'] \quad [2.45a]$$

$$E(W_1 y \varepsilon') = W_1(I - \rho W_1)^{-1} E(\varepsilon \varepsilon') \quad [2.45b]$$

$$E(W_1 y \varepsilon') = \sigma^2 W_1(I - \rho W_1)^{-1} (I - \lambda W_2)^{-1} (I - \lambda W_2')^{-1} \quad [2.45c]$$

L'expression [2.45c] est non nulle en général, les éléments de la variable endogène décalée sont corrélés avec ceux des erreurs. Les paramètres du modèle [2.40] ne peuvent donc pas être estimés d'une façon convergente par les MCO.

Pour le modèle spatial autorégressif [2.23], les estimateurs des MCO ne sont pas convergents car la variable endogène décalée Wy est corrélée avec l'erreur ε , quelle que soit la distribution de cette erreur : si $\lambda = 0$ et $W_1 = W$ dans [2.45c], alors $E(Wy\varepsilon') = \sigma^2 W(I - \rho W)^{-1}$, expression non nulle en général. La matrice $(I - \rho W)^{-1}$ est une matrice pleine et non triangulaire, le décalage spatial pour une observation donnée i , $(Wy)_i$ n'est pas seulement corrélé avec le terme d'erreur en i , mais aussi avec tous les termes d'erreurs de toutes les localisations. Ce résultat contraste avec une propriété de séries temporelles. Dans ce cas, en effet, les estimateurs des MCO restent convergents en présence d'une ou de plusieurs variables retardées tant que les erreurs ne sont pas corrélées.

Pour le modèle avec autocorrélation spatiale des erreurs [2.30], l'estimateur de β par les MCO est sans biais mais inefficace puisque les erreurs ne sont pas sphériques. La solution théorique dans ce cas consiste dans l'application des Moindres Carrés Quasi-Généralisés (MCQG). Cependant, cette procédure n'est pas applicable au cas spatial puisque l'estimateur de λ par les MCO n'est pas convergent (Anselin, 1988a, chap. 6). Sous des hypothèses similaires à celles utilisées ici, Kelejian et Prucha (1997) montrent de plus que les estimateurs obtenus par la méthode des variables instrumentales non linéaire, appliquée au modèle spatial de Durbin [2.34] ne sont pas non plus convergents, l'une des conditions données par Amemiya (1985) n'étant pas vérifiée.

Finalement, pour aboutir à des estimateurs des paramètres dans les modèles spatiaux qui sont convergents, asymptotiquement normaux et efficaces, il faut utiliser d'autres méthodes d'estimation. La plus utilisée aujourd'hui est la méthode par le maximum de vraisemblance.

222.2 Estimation par le maximum de vraisemblance

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

Signalons à ce propos que les propriétés asymptotiques des estimateurs des modèles spatiaux peuvent être déterminées de deux façons. Le premier cas correspond aux modèles

d'échantillonnage usuels : c'est-à-dire aux modèles d'échantillonnage extensifs (*increasing domain asymptotics*). L'extension de l'échantillon résulte alors de l'addition de nouvelles observations à celles qui sont déjà disponibles. Dans le cas de données spatiales, un modèle d'échantillonnage extensif conduit à ajouter indéfiniment de nouvelles unités à celles qui ont été sélectionnées antérieurement sans que les unités déjà présentes soient modifiées. Dans le deuxième cas, l'espace étudié est borné et non extensible, le modèle d'échantillonnage intensif (*infill asymptotics*) est plus approprié. Ces modèles correspondent à des échantillons dont la taille croît suite à des découpages de plus en plus fins de la zone étudiée : les unités spatiales sont redécoupées et donc modifiées. De nombreux résultats valables dans le premier cas ne sont pas directement applicables au second cas (Lahiri, 1996 ; Jayet, 1997). Les applications en économétrie spatiale font davantage appel au premier type de modèle.

Sous l'hypothèse de normalité des résidus, on dérive la fonction de log-vraisemblance dans le cas général [2.40] ce qui permet de déterminer l'espace des paramètres spatiaux ρ et de λ (paragraphe 222.21). L'estimation par la méthode du maximum de vraisemblance pose cependant la question de l'évaluation du jacobien (paragraphe 222.22).

222.21 La fonction de vraisemblance dans le cas général

Pour le modèle général [2.40], le point de départ est l'hypothèse de normalité des termes d'erreur. Si u est un vecteur normal multivarié : $u \sim \text{Nid}(0, \sigma^2 I)$, sa fonction de vraisemblance est :

$$L(u) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2} u'u\right\} \quad [2.46]$$

Dans le modèle général [2.40], $u = (I - \lambda W_2)(y - \rho W_1 y - X\beta)$, le jacobien de la transformation s'écrit donc :

$$J = \det(\partial u / \partial y) = |I - \rho W_1| |I - \lambda W_2| \quad [2.47]$$

La fonction de log-vraisemblance de y s'écrit finalement de la façon suivante :

$$\ln L(y|\rho, \lambda, \beta) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) + \ln|I - \rho W_1| + \ln|I - \lambda W_2| - \frac{1}{2\sigma^2} u'u \quad [2.48]$$

De cette expression [2.48], on déduit l'espace des paramètres de ρ et λ . Dans le premier cas, définissons ω_{\max} la valeur propre positive maximale de la matrice de poids et ω_{\min} la valeur propre négative la plus grande en valeur absolue de W_1 .

On démontre alors que le jacobien $|I - \rho W_1|$ est positif si l'inégalité suivante est respectée :

$$1/\omega_{\min} < \rho < 1/\omega_{\max} \quad [2.49]$$

La même condition doit être respectée pour λ . Pour une matrice standardisée en ligne, $\omega_{\max} = 1$ et $\omega_{\min} > -1$ (Anselin, 1982 ; Anselin, 1988a, chap. 6).

Si les erreurs suivent un processus moyenne mobile de type [2.35], γ doit être compris dans l'intervalle : $]-1/\omega_{\max}; -1/\omega_{\min}[$. Dans ce cas, la fonction de log-vraisemblance est modifiée comme suit :

$$\ln L(y|\rho, \gamma, \beta) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) + \ln|I - \rho W_1| - |I + \gamma W_2| - \frac{1}{2\sigma^2} u'u \quad [2.50a]$$

$$\text{avec } u = (I + \gamma W_2)^{-1} (y - \rho W_1 y - X\beta) \quad [2.50b]$$

Les estimateurs du maximum de vraisemblance sont obtenus en résolvant le système d'équations : $S(\theta) = \partial \ln L / \partial \theta = 0$, où S est le vecteur score et $\theta = [\beta \ \rho \ \lambda]$. Lorsque les conditions de régularité sont vérifiées, les estimateurs du maximum de vraisemblance, solutions du système précédent, sont asymptotiquement efficaces : leur matrice des variances-covariances atteint la borne inférieure de Cramer-Rao, donnée par l'inverse de la matrice d'information : $V(\theta) = I(\theta)^{-1}$ où $I(\theta) = -E[\partial^2 \ln L / \partial \theta \theta']$.

On trouvera dans Anselin (1988a, 1988b) les expressions du score et de la matrice d'information pour le modèle général comportant une variable endogène décalée, une autocorrélation spatiale des erreurs et des résidus hétéroscédastiques. Rietveld et Wintershoven (1998) détaillent le score et la matrice d'information pour des modèles comportant plusieurs décalages. Le système d'équations résultant des conditions du premier ordre n'admet pas de solutions analytiques dans ces différents cas. En revanche, les systèmes correspondant aux modèles plus simples [2.23] et [2.30] admettent des solutions issues des conditions du premier ordre, permettant de construire une fonction de log-vraisemblance concentrée. Nous illustrons cette procédure pour ces deux cas dans l'annexe 2.2.

222.22 Le jacobien

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

La solution la plus ancienne et encore la plus utilisée est proposée par Ord (1975). Elle consiste à exploiter la décomposition du jacobien en termes des N valeurs propres de la matrice de poids W :

$$|I - \rho W| = \prod_{i=1}^N (1 - \rho \omega_i) \quad [2.51]$$

$$\text{soit : } \ln|I - \rho W| = \sum_{i=1}^N \ln(1 - \rho \omega_i) \quad [2.52]$$

L'avantage de cette forme simplifiée apparaît lors de la maximisation de la fonction de log-vraisemblance complète ou de la procédure de balayage de la fonction de log-vraisemblance concentrée. Au lieu de calculer le déterminant de $I - \rho W$ à chaque étape, il n'est besoin de le calculer qu'une fois pour toutes les valeurs propres de W et l'évaluation à chaque étape devient plus facile. Cette propriété permet alors d'écrire la fonction de log-vraisemblance complète en une somme d'éléments correspondant aux observations individuelles. Certains auteurs fournissent ainsi des codes pour quelques logiciels économétriques basés sur cette propriété (Bivand, 1992 ; Griffith, 1988b, 1993 ; Anselin et Hudak, 1992 ; Li, 1996) ³¹.

Lorsque le calcul des valeurs propres pour une matrice de poids W asymétrique pose problème, Ord (1975) donne une propriété intéressante dans le cas où la matrice de poids W est le résultat d'une standardisation d'une matrice de poids symétrique, notée W^* . Cet auteur démontre que les valeurs propres de W se trouvent en calculant les valeurs propres de la matrice symétrique, $D^{1/2}W^*D^{1/2}$ où D est une matrice diagonale dans laquelle chaque élément de la diagonale principale est égal à l'inverse de la somme des éléments de la ligne correspondante de W .

Cette méthode est aujourd'hui encore très largement utilisée. Pourtant, pour de très grands échantillons, le calcul des valeurs propres peut devenir numériquement instable. Une variante de cette procédure est suggérée par Smirnov et Anselin (2001). Elle consiste à évaluer directement les coefficients de la fonction caractéristique. Les simulations effectuées par les auteurs indiquent que cette méthode est la seule méthode « directe » capable de calculer les estimations du maximum de vraisemblance pour de très grands échantillons (plus d'un million d'observations).

D'autres techniques sont proposées pour de très grands échantillons. Certaines d'entre elles exploitent la structure particulière des matrices de poids qui contiennent en général beaucoup d'éléments nuls. Il s'agit des méthodes de factorisation de matrices qui s'avèrent très puissantes pour évaluer rapidement le jacobien : la décomposition de Cholesky pour une

³¹ Lorsque W est une matrice des plus proches voisins dans laquelle chaque région ne possède qu'une seule voisine, Pace et Zou (2000) démontrent que le jacobien $\ln|I - \rho W|$ se simplifie à $\ln|I - \rho W| = K \ln(1 - \rho^2)$, où K représente le nombre de paires d'éléments symétriques dans la matrice de poids.

matrice symétrique et la décomposition de LU sinon. Pace (1997) et Pace et Barry (1997a, 1997b, 1998) ont démontré que ces approches donnent des temps de calcul raisonnables pour des échantillons de plusieurs dizaines de milliers d'observations. D'autres auteurs approximent le jacobien par des fonctions polynomiales ou par les puissances successives de la matrice de poids (Griffith, 1992c ; Martin, 1993 ; Griffith et Sone, 1995). Une approche plus récente suggérée par Barry et Pace (1999) est basée sur des simulations de Monte-Carlo et permet de traiter des échantillons de plus d'un million d'observations.

222.3 Autres méthodes d'estimation

Diverses méthodes d'estimation alternatives sont proposées dans la littérature visant à éviter la programmation des fonctions de vraisemblance. Dans ce paragraphe, nous détaillons les méthodes les plus communément utilisées : la méthode des variables instrumentales (paragraphe 222.31) et celle des moments généralisés (paragraphe 222.32). Notons que d'autres méthodes sont possibles : méthode bootstrap appliquée au modèle autorégressif (Anselin, 1988a, 1990b), méthodes d'estimation bayésiennes (Hepple, 1995a, 1995b ; LeSage, 1997), méthode spatiale de filtrage (Getis, 1990, 1995)³².

222.31 La méthode des variables instrumentales

La méthode des variables instrumentales (VI) est utilisée lorsque certaines variables explicatives sont corrélées avec les erreurs. Dans ce cas, les estimateurs fournis par les MCO ne sont pas convergents. Formellement, si dans le modèle de régression linéaire classique [2.22], on a $\text{plim } X' \varepsilon / N \neq 0$, alors :

$$\text{plim } \hat{\beta}_{MCO} = \beta + \text{plim}[(X'X / N)^{-1}] \text{plim}(X' \varepsilon / N) \neq 0 \quad [2.53]$$

Dans le paragraphe 222.1 (p. 129), nous soulignons que dans le modèle spatial autorégressif [2.23], le principal problème est la corrélation entre la variable endogène décalée et le terme d'erreur. L'adaptation de la méthode VI dans ce cas est proposée par Anselin (1980, 1988a), Land et Deane (1992) ou Kelejian et Robinson (1993).

Formellement, la méthode VI est basée sur l'existence d'un ensemble d'instruments Z , fortement corrélés avec l'ensemble des variables explicatives du modèle autorégressif [2.23] : $X^* = [X \ Wy]$, mais asymptotiquement non corrélés avec les erreurs.

³² Cette méthode consiste à « filtrer » les variables spatialement dépendantes pour les transformer en variables indépendantes grâce aux statistiques d'autocorrélation spatiale locales proposées par Getis et Ord (1992) et Ord et Getis (1995). Il est alors possible d'utiliser les MCO sur les variables filtrées ainsi que toutes les autres mesures de régression qui sont biaisées lorsqu'on utilise le maximum de vraisemblance (tel que le coefficient de détermination R^2). Cette approche suppose néanmoins que l'autocorrélation spatiale est systématiquement une nuisance qu'il convient d'éliminer.

Soit $Z_{(N,M)}$ une matrice d'instruments, où M est le nombre d'instruments vérifiant 4 conditions :

- Condition d'identification : le nombre d'instruments doit être supérieur ou égal au nombre de variables explicatives : $M \geq K + 1$.
- Non-corrélation avec les erreurs : $\text{plim } N^{-1}Z' \varepsilon = 0$.
- Corrélations avec les variables explicatives : $\text{plim } N^{-1}Z' X^* = Q_{ZX^*}$.
- La matrice Z doit posséder des moments de second ordre finis : $\text{plim } N^{-1}Z'Z = Q_{ZZ}$ avec Q_{ZZ} définie-positives.

Soit $\beta^* = [\beta' \rho]$ le vecteur des paramètres à estimer. Le modèle [2.23] se réécrit alors de la façon suivante :

$$y = X^* \beta^* + \varepsilon \quad [2.54]$$

Prémultiplions ce modèle par Z' : on obtient le modèle : $Z' y = Z' X^* \beta^* + Z' \varepsilon$ noté $\tilde{y} = \tilde{X} \beta^* + \tilde{\varepsilon}$. Dans ce modèle, $E(\tilde{\varepsilon}) = E(Z' \varepsilon) = 0$ car Z et ε sont indépendants mais $V(\tilde{\varepsilon}) = Z' V(\varepsilon) Z = \sigma^2 Z' Z$. Puisque les erreurs $\tilde{\varepsilon}$ ne sont pas homoscédastiques, ce modèle doit être estimé grâce à la méthode des Moindres Carrés Généralisés (MCG) :

$$\hat{\beta}_{VI} = [\tilde{X}' (Z' Z)^{-1} \tilde{X}]^{-1} \tilde{X}' (Z' Z)^{-1} y \quad [2.55a]$$

$$\hat{\beta}_{VI} = [X^{*'} P_Z X^*]^{-1} X^{*'} P_Z y \quad [2.55b]$$

où $P_Z = Z(Z' Z)^{-1} Z'$ est une matrice idempotente, symétrique de rang M .

On démontre que cet estimateur est convergent et asymptotiquement distribué selon une loi normale d'espérance mathématique nulle et de matrice des variances-covariances asymptotique $\text{AsyV}(\hat{\beta}_{VI}) = \sigma^2 [Q_{ZX^*} Q_{ZZ}^{-1} Q_{ZX^*}]^{-1}$, estimée par les valeurs dans l'échantillon $\text{AsyV}(\hat{\beta}_{VI}) = \hat{\sigma}^2 [X^{*'} Z(Z' Z)^{-1} Z' X^*]^{-1}$.

La méthode VI possède les mêmes propriétés asymptotiques d'efficacité que le maximum de vraisemblance compte tenu d'un choix approprié d'instruments. En pratique, dans les modèles autorégressifs spatiaux, les variables explicatives stochastiques ou non, indépendantes des erreurs par hypothèse, doivent nécessairement figurer dans les instruments. Pour la variable endogène décalée, on pourra prendre WX comme instrument ou des décalages d'ordre supérieurs (Kelejian et Prucha, 1998). Cette approche peut facilement être étendue à des structures d'erreurs plus complexes (Anselin, 1988a).

222.32 La méthode des moments généralisés

Comme la méthode des variables instrumentales ne fournit pas d'estimateurs convergents pour le coefficient spatial dans le modèle à erreurs autocorrélées, Kelejian et Prucha (1998, 1999) développent une approche par la Méthode des Moments Généralisés (GMM). Ils développent un ensemble de conditions sur les moments permettant l'estimation des équations pour les paramètres dans le modèle à erreurs autocorrélées [2.30] : si $u \sim \text{iid}(0, \sigma^2 I)$, les trois conditions à respecter sont les suivantes (Kelejian et Prucha, 1999, p. 514) :

$$E[u' u / N] = \sigma^2 \quad [2.56a]$$

$$E[u' W' W u / N] = \sigma^2 (1 / N) \text{tr}(W' W) \quad [2.56b]$$

$$E[u' W u / N] = 0 \quad [2.56c]$$

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

La méthode VI et la méthode GMM peuvent enfin être combinées pour estimer les paramètres du modèle général [2.40] contenant à la fois une variable autorégressive et une autocorrélation spatiale des erreurs (Kelejian et Prucha, 1998) pour obtenir des estimateurs convergents. Comme le soulignent Bell et Bockstael (2000), l'avantage de la méthode GMM est la simplicité des calculs, même pour de grands échantillons. En outre, les estimateurs restent convergents quelle que soit la distribution des erreurs.

Comme nous l'avons montré dans la section 21, la modélisation des données spatialisées s'effectue de différentes façons : on peut inclure des variables décalées (endogènes ou exogènes), une autocorrélation spatiale des erreurs ou estimer des modèles avec différentes matrices de poids. Le choix entre ces différentes alternatives passe par la mise en œuvre de différents tests de spécification.

223 Les tests en économétrie spatiale

Les origines des tests d'autocorrélation spatiale remontent au test de Moran (1950a, 1950b). Il est resté dans l'obscurité jusqu'à sa redécouverte par Cliff et Ord (1972) et Burridge (1980). A l'heure actuelle, le test de l'autocorrélation spatiale est le plus souvent basé sur les trois grands principes de tests en économétrie : le test du multiplicateur de Lagrange, ou test du score (Rao, 1947), le test du ratio de vraisemblance et le test de Wald (paragraphe 223.1). Les deux derniers types de tests (Brandsma et Kelletaper, 1979 ; Anselin, 1980 ; Cliff et Ord, 1981) nécessitent cependant l'estimation du modèle non-contraint qui doit être estimé

par des méthodes non-linéaires. Au contraire, le test du score est basé uniquement sur les résultats du modèle sous l'hypothèse nulle et il s'agit la plupart du temps du modèle linéaire standard estimé par les MCO. Partant de ces principes, les tests de l'autocorrélation spatiale peuvent se diviser en plusieurs catégories. Tout d'abord, le test de Moran est le test le plus ancien et encore le plus utilisé. Il vise à tester l'autocorrélation spatiale des résidus lorsque les erreurs suivent un processus autorégressif, un processus moyenne mobile ou la spécification de Kelejian et Robinson (paragraphe 223.2). Plus récemment, les tests du multiplicateur de Lagrange ont été développés pour tester une hypothèse simple en supposant une spécification correcte pour le reste du modèle (paragraphe 223.3) ou pour tester plusieurs types d'autocorrélation spatiale simultanément (paragraphe 223.4). La puissance de ces tests basés sur le principe du maximum de vraisemblance est étudiée dans quelques simulations de Monte-Carlo (paragraphe 223.5) ³³. Enfin, des tests basés sur d'autres principes ont été récemment développés (paragraphe 223.6).

223.1 Principes de tests en économétrie

Le test de la présence d'autocorrélation spatiale est le plus souvent basé sur les trois grands principes de tests liés aux propriétés asymptotiques optimales de l'estimation par le maximum de vraisemblance : le test de Wald, du rapport de vraisemblance et du multiplicateur de Lagrange.

Formellement, notons θ le vecteur de paramètres estimés et $I(\theta)$ la matrice d'information associée à ce paramètre. En toute généralité, les tests d'hypothèses sur les paramètres peuvent être formulés de la façon suivante :

$$\begin{aligned} H_0 : g(\theta) &= 0 \\ H_1 : g(\theta) &\neq 0 \end{aligned} \tag{2.57}$$

$g(\theta)$ est un vecteur de p transformations, linéaires ou non, du vecteur de paramètres θ . Par la suite, nous appellerons modèle contraint le modèle où les restrictions sur les paramètres sont respectées par opposition au modèle non-contraint (ou modèle libre) où elles ne le sont pas. Les différents principes de tests utilisés en économétrie nécessitent l'estimation de l'un et/ou de l'autre de ces modèles.

³³ Nous n'évoquerons pas le test de Kelejian-Robinson (1992) en raison de sa faible puissance (Anselin et Florax, 1995b).

Le test de Wald, le plus familier, exige l'estimation du modèle non-contraint. Il s'écrit de la façon suivante :

$$W = g'[G'VG]^{-1}g \xrightarrow{as} \chi_p^2 \quad [2.58]$$

où g est le vecteur des p éléments et G est la matrice des dérivées partielles, tous deux évalués pour les estimations des paramètres au maximum de vraisemblance, V est la matrice asymptotique des variances-covariances associée au vecteur de paramètres θ ³⁴.

Le test du ratio de vraisemblance nécessite à la fois l'estimation du modèle contraint ainsi que l'estimation du modèle non-contraint et il est défini de la façon suivante :

$$LR = 2[L(\theta) - L(\theta_R)] \xrightarrow{as} \chi_p^2 \quad [2.59]$$

où $L(\theta_R)$ et $L(\theta)$ sont respectivement le maximum de la fonction de log-vraisemblance pour le modèle contraint et pour le modèle non-contraint.

Enfin, *le test du multiplicateur de Lagrange* nécessite uniquement l'estimation du modèle contraint. La statistique générale est :

$$LM = d_R' I(\theta_R)^{-1} d_R \xrightarrow{as} \chi_p^2 \quad [2.60]$$

où d_R et $I(\theta_R)$ sont respectivement le vecteur score et la matrice d'information évalués sous l'hypothèse nulle.

Les tests de Wald, du rapport de vraisemblance et du multiplicateur de Lagrange sont asymptotiquement équivalents. En revanche, en échantillon de taille finie, leurs valeurs diffèrent et respectent l'inégalité suivante : $LM \leq LR \leq W$.

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

³⁴ $W \xrightarrow{as} \chi_p^2$ indique que la statistique W converge asymptotiquement vers une loi du chi-deux à p degrés de libertés.

d'hypothèses simples sur les paramètres, est moins complexe à estimer que le modèle non-contraint : il s'agit souvent du modèle de régression simple estimé par les MCO. Ces tests s'avèrent donc particulièrement utiles pour rechercher la spécification du modèle. En effet, ils permettent d'abord de vérifier si l'estimation de modèles plus complexes est nécessaire, sans avoir à mettre en œuvre immédiatement l'estimation de modèles spatiaux. Ensuite, si un modèle spatial s'avère effectivement nécessaire, ils permettent d'orienter le choix d'une spécification lorsqu'on n'a pas *a priori* sur la forme prise par l'autocorrélation spatiale ou de vérifier que le modèle choisi sur une base théorique est correctement spécifié (par exemple, pour vérifier qu'il ne reste pas une autocorrélation spatiale des erreurs si on a choisi un modèle avec variable endogène décalée). Dans les paragraphes suivants, nous nous concentrons donc plus particulièrement sur les tests du multiplicateur de Lagrange. Un résumé de ces tests est fourni dans le tableau 2.3 (p. 147).

223.2 Le test de l'autocorrélation spatiale de Moran

Le test I , développé par Moran (1950a, 1950b), apparaît comme une extension à deux dimensions du test de corrélation temporelle dans les séries temporelles univariées. Il a été adapté aux résidus d'une régression par Cliff et Ord (1972, 1973) et se présente formellement de la façon suivante en notation matricielle :

$$I = \frac{N}{S_0} \left(\frac{\tilde{\varepsilon}' W \tilde{\varepsilon}}{\tilde{\varepsilon}' \tilde{\varepsilon}} \right) \quad [2.61]$$

où $\tilde{\varepsilon} = y - X\tilde{\beta}$ est le vecteur des résidus estimés des MCO et S_0 est un facteur de standardisation égal à la somme de tous les éléments de W .

Sous l'hypothèse nulle d'indépendance spatiale, le test I de Moran est un test localement meilleur invariant (King, 1981) et est asymptotiquement un test du ratio de vraisemblance de $\lambda = 0$ dans [2.30] ou de $\gamma = 0$ dans [2.35] (Burrige, 1980). Sous l'hypothèse nulle, Cliff et Ord (1972, 1973) dérivent les deux premiers moments de I :

$$E(I) = \frac{\text{tr}(MW)}{N - K} \quad [2.62]$$

et
$$V(I) = \frac{\text{tr}(MWMW') + \text{tr}(MW)^2 + \{\text{tr}(MW)\}^2}{(N - K)(N - K + 2)} - [E(I)]^2 \quad [2.63]$$

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

Le test se base alors sur la statistique de Moran centrée et réduite : $Z(I) = [I - E(I)]/V(I)$. Pour des résidus normalement distribués et une matrice de poids « qui se comporte bien », $Z(I)$ suit asymptotiquement une loi normale centrée et réduite. Pinkse (1999) et Kelejian et Prucha (2001) donnent ainsi des conditions formelles pour la

normalité asymptotique du test de Moran dans plusieurs types de modèles. L'inférence statistique peut également être basée sur la distribution exacte de I en échantillon fini. Tiefelsdorf et Boots (1995) et Hepple (1998) dérivent un test exact en utilisant les résultats sur les ratios de formes quadratiques de variables normales (une présentation complète de ces méthodes se trouve dans Tiefelsdorf, 1998, 2000). Le test de Moran est étendu aux modèles comportant des variables explicatives endogènes par Anselin et Kelejian (1997).

223.3 Les tests d'une hypothèse simple

Trois cas sont abordés : celui du modèle avec autocorrélation des erreurs (paragraphe 223.31), celui du modèle avec la spécification de Kelejian et Robinson (paragraphe 223.32) et enfin celui du modèle avec variable endogène décalée (paragraphe 223.33).

223.31 Modèle avec autocorrélation spatiale des erreurs

Considérons tout d'abord le cas où les erreurs suivent un processus spatial autorégressif [2.30] : $\varepsilon = \lambda W\varepsilon + u$ et pour lequel on teste $H_0 : \lambda = 0$. Sous l'hypothèse nulle, on retrouve le modèle linéaire classique [2.22]. La statistique du test du multiplicateur de Lagrange s'écrit de la façon suivante (Burridge, 1980) :

$$LM_{ERR} = \frac{[\tilde{\varepsilon}' W \tilde{\varepsilon} / \tilde{\sigma}^2]^2}{T} \quad [2.64]$$

avec $T = \text{tr}[(W+W)W]$. $\tilde{\varepsilon}$ et $\tilde{\sigma}^2$ sont les estimations de ε et σ^2 obtenues sous l'hypothèse nulle. Puisqu'il n'y a qu'une seule contrainte, cette statistique suit asymptotiquement une loi du chi-deux à 1 degré de liberté : $LM_{ERR} \xrightarrow{as} \chi_1^2$.

La statistique du test est la même si on spécifie comme hypothèse alternative le processus moyenne mobile [2.35] et comme test $H_0 : \gamma = 0$. LM_{ERR} est donc localement optimal pour les deux alternatives (autorégressive et moyenne mobile) et lorsque l'hypothèse nulle est rejetée, le test ne donne pas d'indications quant à la forme du processus des erreurs³⁵.

³⁵ Dans le cas des processus autorégressif et moyenne mobile sans variables explicatives, Mur (1999) a proposé une stratégie de tests permettant de discriminer entre les deux formes d'autocorrélation spatiale. Cette stratégie est également basée sur des tests du multiplicateur de Lagrange.

223.32 *Spécification de Kelejian-Robinson*

Pour la spécification de l'erreur proposée par Kelejian et Robinson (1995), un test du multiplicateur de Lagrange peut être dérivé suivant le même principe. En reprenant les notations du modèle [2.37] et pour le test de l'hypothèse nulle $H_0 : \varphi = 0$, Anselin (2001a) montre que la statistique de test s'écrit :

$$KR = \left[\frac{\tilde{\varepsilon}' WW' \tilde{\varepsilon}}{\tilde{\sigma}^2} - T_1 \right]^2 / 2 \left(T_2 - \frac{T_1}{N} \right) \quad [2.65]$$

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

Un test, basé sur la méthode des moments généralisés, est proposé par Kelejian et Robinson (1997), et il est obtenu à partir de la condition sur les moments pour les éléments de la diagonale principale de la matrice des variances-covariances [2.38].

223.33 *Test d'une variable endogène décalée*

Finalement, le test d'une variable endogène décalée est proposé par Anselin (1988b). Soit l'hypothèse nulle $H_0 : \rho = 0$ dans [2.23]. On trouve la statistique de test suivante :

$$LM_{LAG} = \frac{[\tilde{\varepsilon}' W y / \tilde{\sigma}^2]^2}{\tilde{T}_1} \quad [2.66]$$

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

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

Il est utile de savoir si le modèle correct contient à la fois une autocorrélation des erreurs et une variable autorégressive. Anselin et Bera (1998) remarquent que LM_{ERR} est la statistique de test correspondant à $H_0 : \lambda = 0$ en supposant une spécification correcte pour le reste du modèle, c'est-à-dire $\rho = 0$. En revanche, si $\rho \neq 0$, ce test n'est plus valide, même asymptotiquement et il n'est plus distribué selon un χ^2 centré à 1 degré de liberté. Une inférence statistique valide nécessite donc de prendre en compte la possible variable endogène décalée lorsqu'on teste l'autocorrélation spatiale des erreurs et vice-versa.

Face à ce problème, plusieurs stratégies sont possibles. On peut déjà effectuer un test joint de présence d'une variable décalée et d'une autocorrélation des erreurs. Pourtant, si

l'hypothèse nulle est rejetée, la nature exacte de la dépendance spatiale n'est pas connue. Une autre solution consiste à estimer un modèle avec une variable endogène décalée et tester ensuite s'il y a encore une autocorrélation des erreurs et vice-versa (Anselin, 1988b). Dans ce cas, il est nécessaire d'estimer les modèles par le maximum de vraisemblance. Anselin *et al.* (1996) proposent des tests basés sur les résidus des MCO dans le modèle simple [2.22] mais qui sont capables de prendre en compte une autocorrélation des erreurs lorsqu'on teste la présence d'une variable endogène décalée et vice-versa.

223.41 Test joint

La première approche consiste à tester l'hypothèse nulle jointe $H_0 : \lambda = \rho = 0$ dans le modèle [2.40] grâce au principe du multiplicateur de Lagrange. Ainsi, le test peut être effectué à partir des résidus des MCO dans le modèle simple [2.22]. La statistique est la suivante :

$$SARMA = E^{-1} \left[(\tilde{d}_\lambda)^2 \frac{D}{\tilde{\sigma}^2} + (\tilde{d}_\rho)^2 T_{22} - 2\tilde{d}_\lambda \tilde{d}_\rho T_{12} \right] \quad [2.67a]$$

$$\Rightarrow SARMA = \frac{\tilde{d}_\lambda^2}{T} + \frac{(\tilde{d}_\lambda - \tilde{d}_\rho)^2}{\tilde{\sigma}^{-2}(D - T\tilde{\sigma}^2)} \text{ si } W_1 = W_2 = W \quad [2.67b]$$

où \tilde{d}_λ et \tilde{d}_ρ sont respectivement les scores par rapport à λ et ρ évalués sous l'hypothèse nulle.

$$T_{ij} = \text{tr}[W_i'W_j + W_i'W_j], \quad D = (W_1X\tilde{\beta})'M(W_1X\tilde{\beta}) + T_{11}\tilde{\sigma}^2, \quad E = (D/\tilde{\sigma}^2)T_{22} - (T_{12})^2,$$

$T = \text{tr}[(W'+W)]$. Sous $H_0 : \lambda = \rho = 0$, SARMA converge asymptotiquement vers une loi du chi-deux à deux degrés de libertés. Si l'hypothèse nulle est rejetée, on ne connaît pas la nature exacte de la dépendance spatiale.

223.42 Tests conditionnels

La deuxième approche consiste à faire un test du multiplicateur de Lagrange pour une forme de dépendance spatiale lorsque l'autre forme n'est pas contrainte. Par exemple, on teste l'hypothèse nulle $H_0 : \lambda = 0$ en présence de ρ . Sous l'hypothèse nulle, on retrouve le modèle autorégressif [2.23] alors que sous l'hypothèse alternative, on retrouve le modèle général [2.40]. Le test est alors basé sur les résidus de l'estimation par le maximum de vraisemblance dans le modèle autorégressif [2.23]. La statistique est la suivante :

$$LM_{ERR}^* = \frac{\hat{d}_\rho^2}{T_{22} - (T_{21A})^2 \hat{V}(\hat{\rho})} \quad [2.68]$$

où $T_{21A} = \text{tr}[W_2W_1A^{-1} + W_2'W_1A^{-1}]$, $A = I - \hat{\rho}W_1$ et où le chapeau désigne les estimateurs MV dans le modèle [2.23] obtenus par optimisation non linéaire, $\hat{V}(\hat{\rho})$ est la variance estimée de $\hat{\rho}$ dans le modèle [2.23]. Sous $H_0 : \lambda = 0$, $LM_{ERR}^* \xrightarrow{as} \chi_1^2$.

On peut également tester l'hypothèse nulle $H_0 : \rho = 0$ en présence de λ , le test est alors basé sur les résidus $\hat{\varepsilon}$ estimés par le MV dans le modèle avec erreurs autorégressives [2.30]. La statistique est :

$$LM_{LAG}^* = \frac{[\hat{\varepsilon}' B' B W_1 y]^2}{H_\rho - H_{\theta\rho} \hat{V}(\hat{\theta}) H_{\theta\rho}'} \quad [2.69]$$

où $\theta = (\beta', \lambda, \sigma^2)$, $B = I - \hat{\lambda}W_2$, $\hat{V}(\hat{\theta})$ est la matrice des variances-covariances estimée de $\hat{\theta}$ dans le modèle [2.30]. Les autres termes sont :

$$H_\rho = \text{tr}W_1^2 + \text{tr}(B W_1 B^{-1})' (B W_1 B^{-1}) + \frac{(B W_1 X \hat{\beta})' (B W_1 X \hat{\beta})}{\hat{\sigma}^2} \quad [2.70]$$

$$\text{et : } H_{\theta\rho}' = \begin{bmatrix} \frac{(B X)' B W_1 X \hat{\beta}}{\hat{\sigma}^2} \\ \text{tr}(W_2 B^{-1})' B W_1 B^{-1} + \text{tr}W_2 W_1 B^{-1} \\ 0 \end{bmatrix} \quad [2.71]$$

Sous $H_0 : \rho = 0$, LM_{LAG}^* converge asymptotiquement vers une loi du chi-deux à un degré de liberté.

223.43 Tests robustes

La dernière approche est celle de Bera et Yoon (1993), reprise par Anselin *et al.* (1996). Elle consiste à utiliser des tests robustes à une mauvaise spécification locale. Par exemple, il s'agit d'ajuster LM_{ERR} pour que sa distribution asymptotique reste un χ^2 centré, même en présence locale de ρ . Ce test s'effectue à partir des résidus des MCO du modèle simple [2.22]. La statistique modifiée pour le test de $H_0 : \lambda = 0$ est :

$$RLM_{ERR} = \frac{[\tilde{d}_\lambda - T_{12} \tilde{\sigma}^2 D^{-1} \tilde{d}_\rho]^2}{T_{22} - (T_{12})^2 \tilde{\sigma}^2 D} \quad [2.72a]$$

$$\Rightarrow RLM_{ERR} = \frac{[\tilde{d}_\lambda - T \tilde{\sigma}^2 D^{-1} \tilde{d}_\rho]^2}{T(1 - T \tilde{\sigma}^2 D)} \quad \text{si } W_1 = W_2 = W \quad [2.72b]$$

De la même façon, on teste $H_0 : \rho = 0$ dans la présence locale de λ est :

$$RLM_{LAG} = \frac{[\tilde{d}_\rho - T_{12}T_{22}^{-1}\tilde{d}_\lambda]}{\tilde{\sigma}^{-2}D - (T_{12})^2T_{22}^{-1}} \quad [2.73a]$$

$$\Rightarrow RLM_{LAG} = \frac{[\tilde{d}_\rho - \tilde{d}_\lambda]}{\tilde{\sigma}^{-2}D - T} \quad \text{si} \quad W_1 = W_2 = W \quad [2.73b]$$

223.5 Puissance et robustesse des tests

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

Anselin et Rey (1991) et Anselin et Florax (1995b) comparent les performances des tests de Moran, LM_{ERR} , LM_{LAG} , RLM_{ERR} et RLM_{LAG} pour différentes matrices de poids (sur zonage régulier ou non), différentes distributions des erreurs et différentes tailles d'échantillon. Les principaux résultats sont les suivants.

- Premièrement, LM_{LAG} est le test le plus puissant et le plus robuste à la non-normalité des erreurs.
- Deuxièmement, le test I de Moran est puissant pour les deux alternatives, variable endogène décalée ou autocorrélation des erreurs. Ce test devient ainsi un indicateur général d'une mauvaise spécification du modèle, quelle que soit la forme de la dépendance spatiale omise mais ne peut discriminer entre les différentes formes possibles.
- Troisièmement, les tests LM_{ERR} et LM_{LAG} ont les plus grandes puissances pour leur alternative respective alors que les tests ajustés RLM_{ERR} et RLM_{LAG} possèdent également de bonnes performances en termes de puissance et de taille nominale. Ces tests du score sont donc à privilégier dans la recherche de la spécification du modèle.

Florax et Rey (1995) étudient les conséquences d'une mauvaise spécification de la matrice de poids sur la puissance des tests précédents. Ils distinguent « sur-spécification » et « sous-spécification ». Dans le premier cas, des observations sont rajoutées à tort (par exemple lorsqu'on utilise une matrice de distance au lieu d'une matrice de contiguïté) alors que dans le deuxième cas, des observations sont omises à tort. Les simulations effectuées par ces auteurs suggèrent qu'une sur-spécification entraîne une baisse de puissance alors qu'une sous-spécification augmente la puissance des tests en présence d'autocorrélation spatiale positive et la diminue en cas d'autocorrélation spatiale négative. Le coefficient de Moran est moins affecté que les autres tests par une mauvaise spécification de la matrice de poids. Griffith (1995) indique qu'il est préférable d'employer une matrice de poids légèrement sous-spécifiée plutôt qu'une matrice de poids sur-spécifiée.

Il nous faut encore souligner que l'ensemble des méthodes d'estimation et d'inférence présentées sont soumises à un « effet de frontière » qui apparaît lorsque la dépendance spatiale transcende les frontières des régions étudiées : une variable expliquée dans une zone géographique i est influencée par la variable expliquée en un lieu j alors que j n'appartient pas à l'échantillon de zones géographiques. Ces effets présentent un problème d'estimation (Griffith, 1980 ; Upton et Fingleton, 1985). Ainsi, les travaux de Haining (1977), Griffith (1983, 1985, 1987) et Griffith et Amrhein (1983) montrent que le coefficient associé à la variable endogène décalée dans un modèle autorégressif pur (c'est-à-dire un modèle spatial autorégressif sans variables explicatives autres que la variable endogène décalée) tend à être sous-estimé si aucune correction n'est effectuée. Des techniques de correction ont été proposées visant à modifier la structure de la matrice de poids (par exemple, on introduit une zone artificielle autour des régions étudiées, zone où la variable explicative prend comme valeur la moyenne de toutes les autres valeurs). D'autres approches statistiques ont également été avancées (l'une d'elle consiste à traiter les effets de frontière comme un problème de variables manquantes). Pourtant, les travaux de Griffith et Amrhein (1983) suggèrent que ces corrections n'améliorent que très marginalement les estimations, voire sont sans effet. Florax et Rey (1995) soulignent que le problème des effets de frontière n'a encore aujourd'hui reçu aucune réponse satisfaisante.

223.6 Autres tests d'autocorrélation spatiale

Le test de Moran et les tests basés sur le principe du multiplicateur de Lagrange sont aujourd'hui très largement mobilisés pour tester l'autocorrélation spatiale dans les modèles en coupe transversale. Cependant, d'autres types de tests existent. Ainsi, Baltagi et Li (1998) développent des tests de l'autocorrélation spatiale des erreurs et d'une variable endogène décalée en utilisant la méthode des régressions artificielles proposée par Davidson et McKinnon (1984, 1988, 1993). Des simulations de Monte-Carlo effectuées par les auteurs indiquent que ces tests ont des performances similaires aux tests du multiplicateur de Lagrange équivalents. Saavedra (2000) formule des versions GMM du test de Wald, du ratio de vraisemblance et du multiplicateur de Lagrange, pour le test d'une variable endogène décalée dans un modèle autorégressif comportant une autocorrélation des erreurs sous la forme proposée par Kelejian et Robinson [2.37]. Les simulations de Monte-Carlo effectuées par l'auteur suggèrent que les distributions en échantillon fini de ces tests sont relativement bien approximées par leur distribution asymptotique (chi-deux) et que les puissances empiriques approchent 1 pour des échantillons de petite taille et de faibles valeurs du paramètre spatial. Enfin, Baltagi et Li (2001) proposent des tests du multiplicateur de Lagrange dans le but de tester simultanément l'autocorrélation spatiale des erreurs ainsi qu'une spécification de type Box-Cox.

224 A la recherche de la spécification du modèle

Les tests de spécification présentés permettent de détecter une omission de l'autocorrélation spatiale et la forme prise par cette dernière dans le modèle. Combinés au test du facteur commun (paragraphe 224.1), ils fournissent des règles de décision permettant de rechercher la meilleure spécification du modèle (paragraphe 224.2).

224.1 Le test du facteur commun

Le test du facteur commun (cf. tableau 2.3, p. 147) permet de choisir entre un modèle avec autocorrélation des erreurs et un modèle avec variable endogène décalée et l'ensemble des variables explicatives décalées (BurrIDGE, 1981 ; Bivand, 1984). Le point de départ est la formulation de Durbin : les deux modèles suivants [2.74a] et [2.74b] sont équivalents :

$$\begin{cases} y = X\beta + \varepsilon \\ \varepsilon = \lambda W\varepsilon + u \end{cases} \quad [2.74a]$$

$$\Leftrightarrow y = \lambda Wy + X\beta - \lambda WX\beta + u \quad [2.74b]$$

Le modèle [2.74b] est estimé par :

$$y = \lambda Wy + X\beta - WX\delta + u \quad [2.75]$$

Par conséquent, pour savoir si le modèle [2.75] peut se réduire au modèle [2.74a], il faut tester l'hypothèse suivante : $H_0 : \lambda\beta + \delta = 0$. Au contraire des séries temporelles, un modèle avec erreurs autocorrélées n'est pas plus facile à estimer qu'un modèle avec une variable endogène décalée puisqu'il faut utiliser dans les deux cas la méthode du maximum de vraisemblance. Le seul avantage de la forme [2.74a] sur la forme [2.75] est qu'il y a moins de paramètres à estimer : $K + I$ au lieu de $2K + I$.

Le test sur les contraintes des paramètres se fait avec l'un des trois tests traditionnels : test de Wald, test du rapport de vraisemblance ou test du multiplicateur de Lagrange. Tous trois convergent asymptotiquement vers une loi du chi-deux à χ^2 à $K-I$ degrés de liberté (en ignorant le terme constant).

Le test du facteur commun part donc d'un modèle général sur lequel des tests sont faits pour voir si un modèle plus simple est plus pertinent ou non. Cette procédure peut constituer un désavantage lorsqu'il y a peu d'observations : les degrés de liberté risquent de devenir trop faibles. Ce test a été étendu à plusieurs décalages spatiaux et retards temporels par Blommestein (1983) et Blommestein et Nijkamp (1986).

RESUME DES DIFFERENTS TESTS

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

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

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

Tableau 2.3 : Résumé des différents tests

224.2 Les règles de décision

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

1/ La première étape consiste à estimer le modèle simple [2.22] par les MCO et effectuer le test de Moran [2.61] et le test SARMA [2.67a]. Le rejet de l'hypothèse nulle dans ces deux cas indique une mauvaise spécification du modèle et une omission à tort de l'autocorrélation spatiale mais ne permet pas de connaître la forme prise par l'autocorrélation spatiale.

2/ Si les tests indiquent la présence de dépendance spatiale, il est souvent utile de commencer par inclure dans le modèle, si possible, des variables supplémentaires. Il peut s'agir de variables exogènes supplémentaires qui sont susceptibles d'éliminer ou de diminuer la dépendance spatiale, ou de variables exogènes décalées spatialement, correspondant en totalité ou en partie aux variables explicatives initiales (Florax et Folmer, 1992).

3/ Si l'ajout de variables exogènes supplémentaires n'a pas éliminé l'autocorrélation spatiale, il faut alors estimer un modèle incorporant une variable autorégressive ou une autocorrélation des erreurs. Le choix entre ces deux formes de dépendance spatiale s'effectue en comparant les niveaux de significativité des tests du multiplicateur de Lagrange LM_{ERR} [2.64] et LM_{LAG} [2.66] et leurs version robustes RLM_{ERR} [2.72a] et RLM_{LAG} [2.73a].

Initialement, Anselin et Rey (1991) et Florax et Folmer (1992) suggèrent de choisir l'un ou l'autre modèle [2.23] ou [2.30] en appliquant la règle de décision simple suivante :

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

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

Anselin et Florax (1995b) affinent cette règle de décision en la complétant par l'utilisation des tests robustes : si LM_{LAG} est plus significatif que LM_{ERR} et RLM_{LAG} est significatif mais pas RLM_{ERR} , on inclut une variable endogène décalée. D'une façon similaire, la présence d'une autocorrélation des erreurs peut être identifiée à travers RLM_{ERR} .

4/ Une fois que le modèle spatial adéquat est estimé ([2.23] ou [2.30]), trois tests supplémentaires peuvent encore être mobilisés.

- Pour un modèle autorégressif [2.23], le test LM_{ERR}^* permet de savoir si une autocorrélation spatiale des erreurs est encore nécessaire.
- Pour un modèle avec autocorrélation des erreurs [2.30], le test LM_{LAG}^* permet de savoir si une variable endogène décalée est encore nécessaire. Le test du facteur commun indique si la restriction $\lambda\beta + \delta = 0$ peut être rejetée ou non. Si elle ne l'est pas, le modèle [2.75] se réduit au modèle avec autocorrélation des erreurs [2.74a].

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

$$INF = -2 \ln L + q(K) \quad [2.76]$$

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

L'ensemble des méthodes d'estimation et d'inférence présentées dans les paragraphes 222, 223 et 224 sont valables pour les modèles linéaires en coupe transversale. Nous présentons dans le paragraphe suivant les extensions de ces techniques à d'autres types de modèles.

225 Quelques extensions récentes

De récents développements en économétrie spatiale portent sur l'incorporation de l'autocorrélation spatiale dans les modèles de panel (paragraphe 225.1), ceux à variables qualitatives (paragraphe 225.2) et ceux à équations simultanées (paragraphe 225.3).

225.1 Données de panel et autocorrélation spatiale

L'autocorrélation spatiale est introduite dans des modèles de plusieurs façons. Nous considérons d'abord le cas de modèles spatio-temporels simples où les comportements de tous les individus sont supposés uniformes dans le temps et dans l'espace (paragraphe 225.11). L'autocorrélation spatiale sous la forme d'une variable endogène décalée ou d'une autocorrélation spatiale des erreurs peut également être introduite dans d'autres modèles de panel, tels que le modèle SUR (paragraphe 225.12) ou le modèle à erreurs composées (paragraphe 225.13)³⁶.

225.11 Modèles dynamiques dans le temps et dans l'espace

Supposons qu'il y ait N régions i , T périodes de temps t et donc NT observations y_{it} ³⁷. Considérons tout d'abord le modèle le plus simple (régression ordinaire) dans lequel on suppose des comportements uniformes pour tous les individus :

$$y_{it} = \sum_{k=1}^K x_{k,it} \beta_k + \varepsilon_{it} = x'_{it} \beta + \varepsilon_{it} \quad [2.77]$$

³⁶ La combinaison des effets fixes et de l'autocorrélation spatiale n'est pas possible. En effet, l'estimation de modèles spatiaux nécessite le caractère asymptotique dans le domaine spatial ($N \rightarrow \infty$) alors que les effets fixes (c'est-à-dire une variable muette pour chaque localisation) nécessitent l'estimation de N paramètres. Il n'y a donc pas d'estimateurs convergents et les effets fixes sont incompatibles avec les processus spatiaux.

³⁷ Nous supposons dans tout ce paragraphe que les échantillons sont cylindrés.

où $\varepsilon_{it} \sim iid(0, \sigma^2)$, $\forall i$ et $\forall t$, x_{it}' est le vecteur-ligne des observations des variables explicatives pour une région i au temps t de dimension $(1, K)$ et β est le vecteur des K paramètres inconnus.

Notons y_t le vecteur $(N, 1)$ qui regroupe les N observations de y_{it} pour la période t . X_t désigne la matrice (N, K) des observations des variables explicatives au temps t , ε_t est le vecteur des N erreurs au temps t . Le modèle se réécrit alors de la façon suivante :

$$y_t = X_t \beta + \varepsilon_t \quad [2.78]$$

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

$$y_t = \lambda y_{t-1} + \rho W y_t + \gamma W y_{t-1} + X_t \beta_1 + X_{t-1} \beta_2 + W X_t \beta_3 + W X_{t-1} \beta_4 + \varepsilon_t \quad [2.79]$$

Ce modèle général est estimé grâce à la méthode du maximum de vraisemblance. En outre, ce modèle se simplifie si certaines contraintes de type facteur commun sur les paramètres sont vérifiées (Elhorst, 2001). Nous considérons deux cas particuliers :

- Si $\gamma = -\lambda\rho$, $\beta_2 = -\lambda\beta_1$ et $\beta_4 = -\lambda\beta_3$, alors le modèle [2.79] se réduit à un modèle avec décalage spatial sur la variable expliquée et les variables explicatives avec une autocorrélation temporelle des erreurs :

$$\begin{aligned} Y_t &= \rho W Y_t + X_t \beta_1 + W X_t \beta_3 + u_t \\ u_t &= \lambda u_{t-1} + \varepsilon_t \end{aligned} \quad [2.80]$$

- Si $\gamma = -\lambda\rho$, $\beta_3 = -\rho\beta_1$ et $\beta_4 = -\rho\beta_2$, alors le modèle [2.74] se réduit à un modèle avec décalage temporel sur la variable expliquée et les variables explicatives avec une autocorrélation spatiale des erreurs :

$$\begin{aligned} Y_t &= \lambda Y_{t-1} + X_t \beta_1 + X_{t-1} \beta_2 + u_t \\ u_t &= \rho W u_t + \varepsilon_t \end{aligned} \quad [2.81]$$

Des modèles encore plus simples peuvent être dérivés de [2.79]. Un cas particulier intéressant est le modèle où $\lambda = \rho = \beta_2 = \beta_3 = \beta_4 = 0$:

$$y_t = \gamma W y_{t-1} + X_t \beta_1 + \varepsilon_t \quad [2.82]$$

Au contraire des modèles précédents, l'estimation des paramètres de [2.82] peut être basée sur les moindres carrés ordinaires. Ce modèle est utilisé pour modéliser la diffusion d'un phénomène (Upton et Fingleton, 1985). Par exemple, Dubin (1995, 1997) étudie la diffusion d'une innovation dans le cadre d'un modèle LOGIT en données de panel : la probabilité

d'adoption d'une innovation par une firme dépend entre autres de la proximité de firmes qui ont adopté cette innovation à la période précédente.

Pour tester les formes d'autocorrélation spatiale sur la variable expliquée ou sur le terme d'erreur, on utilise les tests du multiplicateur de Lagrange LM_{LAG} et LM_{ERR} avec la matrice de poids $W^* = I_T \otimes W$.

225.12 Modèle SUR spatial

Considérons dans un premier temps le modèle SUR spatial dans lequel $N > T$ (Anselin, 1988a, 1988c)³⁸. On suppose les paramètres constants dans l'espace mais variables pour chaque période. Le modèle s'écrit alors de la façon suivante :

$$y_{it} = x'_{it} \beta_t + \varepsilon_{it} \quad i = 1, \dots, N \text{ et } t = 1, \dots, T \quad [2.83]$$

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

Pour chaque équation, il est possible d'être confronté à un problème d'autocorrélation spatiale des erreurs :

$$\varepsilon_t = \lambda_t W \varepsilon_t + u_t \quad [2.84]$$

$$\text{avec : } E(u_t) = 0 \quad \forall t \quad [2.85]$$

$$E(u_t u'_t) = \sigma_t^2 I_N \quad [2.86a]$$

$$E(u_t u'_s) = \sigma_{ts} I_N \quad [2.86b]$$

De ces différentes hypothèses, on déduit que les erreurs complètes u possèdent les propriétés suivantes : $E(u) = 0$ et $E(uu') = \Sigma \otimes I_N$ où Σ est la matrice d'élément générique : $[\sigma_{ts}]$.

Dans ces conditions, la matrice des variances-covariances des erreurs ε s'écrit : $\Omega = E(\varepsilon \varepsilon') = B(\Sigma \otimes I_N)B'$ où $B = [I_{NT} - (\Lambda \otimes W)]^{-1}$ et $\Lambda = \text{diag}(\lambda_t)$.

Une application de ce modèle est celle de Rey et Montouri (1999) qui estiment un modèle SUR spatial pour étudier dans quelle mesure le processus de convergence entre les états américains diffère entre les 2 sous-périodes 1929-1945 et 1946-1994.

³⁸ Anselin (1988a) appelle ce modèle le modèle SUR « spatial » par opposition au modèle SUR « classique » (Zellner, 1962) où les coefficients de la régression sont constants dans le temps mais variables selon les individus. Dans ce cas, l'autocorrélation spatiale est captée grâce à l'estimation de la matrice des variances-covariances contemporaine (Arora et Brown, 1977 ; Hordijk et Nijkamp, 1977 ; Hordijk, 1979).

De la même façon, on peut considérer un modèle SUR avec variable endogène décalée. Le modèle s'écrit alors pour les N individus à la période t :

$$y_t = \rho_t W y_t + X_t \beta_t + \varepsilon_t \quad t = 1, \dots, T \quad [2.87]$$

On pose sur l'erreur ε les mêmes hypothèses posées précédemment sur u . Dans ces conditions, $E(\varepsilon) = 0$ et $E(\varepsilon \varepsilon') = \Sigma \otimes I_N$, où $\Sigma = [\sigma_{ts}]$.

Ces différents modèles peuvent être estimés à partir de la méthode du maximum de vraisemblance. En outre, des tests du multiplicateur de Lagrange permettant de tester les hypothèses $H_0 : \lambda_t = 0, t = 1, \dots, T$ dans [2.85] et $H_0 : \rho_t = 0, t = 1, \dots, T$ dans [2.88] sont suggérés par Anselin (1988a, 1988c).

225.13 Modèles à erreurs composées

L'autocorrélation spatiale des erreurs peut enfin être incorporée dans le cadre d'un modèle à erreurs composées dans lequel les coefficients de la régression sont supposés fixes (Anselin, 1988a ; Case, 1991 ; Baltagi et Li, 2002). Dans ce modèle :

$$y_{it} = x_{it}' \beta + \varepsilon_{it} \quad i = 1, \dots, N ; t = 1, \dots, T \quad [2.88]$$

Le terme d'erreur est supposé incorporer les effets non-observables dus à l'espace (Hsiao, 1986 ; Baltagi, 1995) :

$$\varepsilon_{it} = \mu_i + \phi_{it} \quad i = 1, \dots, N ; t = 1, \dots, T \quad [2.89]$$

avec $\phi_{it} \sim iid(0, \sigma^2)$, $\mu_i \sim iid(0, \sigma_\mu^2)$, composantes supposées indépendantes.

En regroupant les N individus pour la période t , l'erreur peut également s'écrire de la façon suivante : $\varepsilon_t = \mu + \phi_t$ avec $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})'$, $\mu = (\mu_1, \dots, \mu_N)'$, $\phi_t = (\phi_{1t}, \dots, \phi_{Nt})'$. On suppose une autocorrélation spatiale des erreurs sur le terme d'erreur résiduel :

$$\phi_t = \lambda W \phi_t + v_t \quad [2.90]$$

où $v_t = (v_{1t}, \dots, v_{Nt})'$, $v_t \sim iid(0, \sigma_v^2)$ et indépendant de ϕ_t et μ . Soit $B = I_N - \lambda W$, alors $\phi_t = (I_N - \lambda W)^{-1} v_t = B^{-1} v_t$. En conséquence, $\varepsilon = (S_T \otimes I_N) \mu + (I_T \otimes B^{-1}) v$ et la matrice de variances-covariances de ε s'écrit :

$$\Omega = E(\varepsilon \varepsilon') = \sigma_\mu^2 (S_T S_T' \otimes I_N) + \sigma_v^2 (I_T \otimes (B' B)^{-1}) \quad [2.91]$$

Ce modèle peut s'estimer grâce à la méthode du maximum de vraisemblance. Un test du multiplicateur de Lagrange pour $H_0 : \lambda = 0$ est suggéré par Anselin (1988a). Plus récemment, Baltagi *et al.* (2001) dérivent des tests du multiplicateur de Lagrange permettant de tester les hypothèses suivantes : $H_0 : \lambda = 0$ et $\sigma_v^2 = 0$, $H_0 : \sigma_v^2 = 0$ lorsque $\lambda = 0$, $H_0 : \lambda = 0$ lorsque $\sigma_v^2 = 0$, $H_0 : \lambda = 0$ lorsque $\sigma_v^2 \geq 0$ et $H_0 : \sigma_v^2 = 0$ lorsque $\lambda \geq 0$. Enfin, Baltagi et Li (2002) dérivent le prévisseur linéaire optimal pour le modèle à erreurs composées et autocorrélation spatiale ³⁹.

225.2 Les modèles à variables qualitatives

La prise en compte de l'autocorrélation spatiale dans les modèles de choix discrets, les modèles censurés ou les modèles à sélection d'échantillon constitue un thème de recherche récent. Nous illustrons les problèmes impliqués par la prise en compte de l'autocorrélation spatiale dans les modèles à variables qualitatives à partir du modèle de choix discret avec la variable latente y_i^* :

$$y_i^* = x_i' \beta + \varepsilon_i \quad [2.92]$$

où ε_i est une variable aléatoire pour laquelle une distribution donnée est supposée (normale pour le modèle PROBIT, logistique pour un modèle LOGIT). On observe :

$$y_i = 1 \text{ lorsque } y_i^* \geq 0 \quad [2.93a]$$

$$y_i = 0 \text{ lorsque } y_i^* < 0 \quad [2.93b]$$

L'autocorrélation spatiale est introduite dans ce modèle sous la forme d'une variable endogène décalée ou d'une autocorrélation spatiale pour le terme d'erreur ε_i . Des problèmes similaires apparaissent pour ces deux cas.

Considérons par exemple une autocorrélation spatiale des erreurs (Anselin, 2001b ; Beron et Vijverberg, 2002) :

$$\varepsilon = (I - \lambda W)^{-1} u \quad [2.94a]$$

$$\Rightarrow \varepsilon_i = \sum_j \delta_{ij} u_j \quad [2.94b]$$

³⁹ Il ne s'agit cependant pas de la seule façon d'introduire de l'autocorrélation spatiale des erreurs dans un modèle à erreurs composées : Bolduc, Laferrrière et Santarossa (1995) proposent, dans le cadre d'un modèle de flux, de supposer que chaque composante de l'erreur suit un processus spatial autorégressif.

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

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

Cette expression n'est pas bloc-diagonale et conduit à des variances hétéroscédastiques. L'autocorrélation spatiale introduit donc des interdépendances entre observations qui interdisent le calcul des probabilités séparément pour chaque localisation : la probabilité $P(\varepsilon_i < x_i' \beta)$ ne peut plus être dérivée de la distribution normale standard univariée mais doit être explicitement dérivée comme la distribution marginale du vecteur multivarié à N dimensions, dont la matrice des variances-covariances contient le paramètre autorégressif λ (McMillen, 1992 ; Jayet, 1993). Des problèmes similaires apparaissent lorsque l'autocorrélation spatiale est introduite dans un modèle de type TOBIT 2 (McMillen, 1995a).

Les problèmes d'estimation et de tests dans ces différents cas ont été abordés par Case (1992), McMillen (1992, 1995a, 1995b) et Pinkse et Slade (1998). Dans ces modèles, l'interdépendance impliquée par l'autocorrélation spatiale produit une fonction de vraisemblance avec de multiples intégrales, rendant l'estimation directe pratiquement impossible. De plus, l'hétéroscédasticité des erreurs, induite par la spécification de l'autocorrélation spatiale sous la forme d'une variable endogène décalée ou d'une autocorrélation spatiale des erreurs, implique que les méthodes d'estimation des modèles PROBIT avec observations dépendantes mais erreurs homoscédastiques proposées par Avery *et al.* (1983) ou Poirier-Ruud (1988) ne sont pas convergentes (Pinkse et Slade, 1998).

Dans le cadre des modèles de choix discrets, McMillen (1992, 1995b) propose d'utiliser la méthode d'estimation « Espérance – Maximisation » ou « EM » pour les modèles PROBIT avec autocorrélation spatiale des erreurs ou variable endogène décalée. Cette procédure d'estimation remplace la variable discrète dépendante par l'espérance de la variable latente continue. Le modèle est ensuite estimé par maximum de vraisemblance en considérant la variable construite comme une variable standard continue et dépendante. La procédure d'estimation – calcul de l'espérance de la variable dépendante et estimation du modèle par maximum de vraisemblance – est répétée jusqu'à convergence et les estimations des paramètres obtenues sont les estimations du maximum de vraisemblance. Cependant, cette méthode ne permet pas de déterminer analytiquement la matrice d'information. LeSage (1999b) propose l'utilisation de méthodes d'estimation bayésiennes qui fournissent les mêmes estimations que celles données par la méthode du maximum de vraisemblance mais ne souffrent pas de ce problème. Enfin, Pinkse (1999, 2002) étend le test de Moran aux modèles PROBIT.

225.3 Les modèles à équations simultanées

L'extension des méthodes de l'économétrie spatiale aux modèles d'équation simultanées est proposée par Rey et Boarnet (2002). Ils étudient un modèle général comportant les deux équations suivantes :

$$\begin{aligned} y_1 &= X\beta_1 + \gamma_{21}y_2 + \rho_{11}Wy_1 + \rho_{21}Wy_2 + \varepsilon_1 \\ y_2 &= X\beta_2 + \gamma_{12}y_1 + \rho_{12}Wy_1 + \rho_{22}Wy_2 + \varepsilon_2 \end{aligned} \quad [2.95]$$

où y_i , $i = 1, 2$ est le vecteur des observations sur les variables dépendantes, X est la matrice (N,K) des observations sur les K variables exogènes avec les vecteurs de paramètres associés β_i , $i = 1,2$. Enfin, ε_i , $i = 1, 2$ sont les vecteurs d'erreurs. Les termes d'erreurs dans chaque équation sont d'espérance nulle, ils ne sont pas spatialement autocorrélés et il n'y a pas de covariance inter-équation.

La simultanéité apparaît dans ce modèle de trois façons possibles :

- y_1 et y_2 sont des variables endogènes qui apparaissent dans les deux équations.
- L'influence de Wy_1 sur y_1 et de Wy_2 sur y_2 : termes autorégressifs spatiaux traditionnels.
- L'influence de Wy_2 sur y_1 et de Wy_1 sur y_2 : termes autorégressifs spatiaux croisés.

En imposant des restrictions de nullité sur tout ou partie des coefficients ρ_{11} , ρ_{21} , ρ_{12} , ρ_{22} , γ_{12} et γ_{21} , les auteurs distinguent 35 spécifications différentes selon que les 3 types de simultanéité sont absents ou présents dans une ou deux équations.

Ces modèles sont estimés à l'aide de la méthode des Doubles Moindres Carrés dans laquelle les décalages spatiaux des variables endogènes ou exogènes sont utilisés comme instruments.

Henry *et al.* (2001) appliquent ces principes pour, d'une part, étudier à l'aide de deux équations les phénomènes d'interdépendance spatiale entre croissance de la population et croissance de l'emploi dans les zones rurales et d'autre part, estimer les effets de la proximité de la croissance urbaine sur les zones rurales.

Conclusion de section

Dans cette section, nous avons présenté la façon dont l'autocorrélation spatiale était introduite dans les modèles économétriques. Dans le cas de modèles linéaires en coupe transversale, nous avons montré que les principales spécifications intégrant cet effet conduisent à une corrélation entre les erreurs et les variables explicatives. Les MCO ne sont donc plus adaptés et il faut se tourner vers d'autres méthodes d'estimation telle la méthode du

maximum de vraisemblance, celle des variables instrumentales ou des moments généralisés. Enfin, nous avons présenté différents tests qui, à travers les règles de décision, permettent de déterminer la forme prise par l'autocorrélation spatiale et son interprétation.

D'une façon générale, cela signifie que la présence d'autocorrélation spatiale doit être systématiquement testée dans les spécifications que l'on cherche à estimer. Bien sûr, le modélisateur peut disposer *a priori* de références théoriques ou empiriques orientant la spécification qu'il souhaite estimer, en particulier lorsqu'il pense que les phénomènes économiques dépendent des valeurs prises par la variable endogène ou par certaines variables explicatives dans les unités voisines. Cependant, même dans ce cas, la présence d'autocorrélation spatiale « résiduelle » doit être testée afin de vérifier la justesse de la spécification définie *a priori*. On peut par ailleurs préférer opter pour la spécification avec autocorrélation spatiale des erreurs en substitution des variables omises lorsqu'on ne souhaite pas intégrer de variables explicatives ou lorsqu'on ne dispose pas de ces variables.

Nous abordons maintenant la façon dont le second effet, l'hétérogénéité spatiale, est introduit dans les modèles économétriques.

SECTION 23 : HETEROGENEITE SPATIALE

Dans cette section, nous abordons le traitement de l'hétérogénéité spatiale. Dans une régression économétrique, ce phénomène se traduit de deux façons : par des coefficients différents ou par des variances des termes d'erreurs différentes selon la localisation. Dans le premier cas, on parle d'*instabilité spatiale des paramètres* de la régression qui varient systématiquement avec la localisation. Dans le second cas, on est confronté à un problème d'*hétéroscédasticité*, problème par ailleurs fréquent en coupe transversale. Ces deux cas, instabilité des paramètres et hétéroscédasticité, peuvent apparaître conjointement.

Notre objectif ici est double. D'une part, nous présentons les principales spécifications économétriques permettant de capter ces phénomènes d'hétérogénéité spatiale. D'autre part, nous examinons dans quelle mesure les tests traditionnels d'hétéroscédasticité ou d'instabilité doivent être amendés pour tenir compte de l'autocorrélation spatiale. Nous nous limitons ici aux spécifications valables pour les données en coupe transversale.

Au contraire de l'autocorrélation spatiale, l'hétérogénéité spatiale est habituellement traitée par les méthodes économétriques standards. Ainsi, les modélisations en coupe transversale relatives aux paramètres variables, aux coefficients aléatoires ou aux changements de régimes, peuvent être adaptées aux spécifications particulières nécessaires à l'utilisation de données localisées. Nous présentons d'abord les différentes spécifications permettant de prendre en compte l'instabilité des paramètres dans l'espace (paragraphe 231) ainsi que l'hétéroscédasticité (paragraphe 232). Ensuite, nous explicitons les liens entre autocorrélation spatiale et hétérogénéité spatiale. Ces deux aspects sont, en effet, souvent présents conjointement dans une régression, ce qui conduit à développer des tests particuliers permettant de vérifier la présence de l'un des effets en tenant compte de l'autre ou de vérifier la présence de ces deux effets simultanément (paragraphe 233).

231 Modélisation et estimation de l'instabilité spatiale des paramètres

Reprenons comme point de départ le modèle de régression linéaire classique en coupe transversale :

$$y = X\beta + \varepsilon \quad [2.96]$$

où y est le vecteur $(N,1)$ des observations de la variable dépendante, N est le nombre total d'observations, X est la matrice (N,K) des observations des K variables explicatives, β est le

vecteur $(K,1)$ des K coefficients inconnus à estimer. Sauf indication contraire, ε désigne dans tout ce paragraphe le vecteur $(N,1)$ des termes d'erreurs tel que $\varepsilon \sim \text{iid}(0, \sigma^2 I_N)$. Comme précédemment, nous considérons que les observations sont localisées.

L'instabilité spatiale des paramètres se réfère à l'absence de stabilité dans l'espace des comportements ou des relations économiques. Ce phénomène nécessite des modélisations permettant la prise en compte des caractéristiques particulières de chaque zone géographique : les paramètres de la régression [2.96] varient selon leurs localisations. Dans le cas le plus général, on suppose une relation différente pour chaque zone géographique i de l'échantillon :

$$y_i = x_i' \beta_i + \varepsilon_i \quad i = 1, \dots, N \quad [2.97]$$

où y_i représente l'observation de la variable dépendante pour la zone i , x_i' représente le vecteur de dimension $(1,K)$ comportant les observations des K variables explicatives pour l'observation i . Ce vecteur est associé à un vecteur de dimension $(K,1)$: β_i . Il existe alors un vecteur de paramètres différent pour chaque observation i .

Compte tenu de cette spécification, il n'est pas possible d'estimer un ensemble de N vecteurs de K paramètres inconnus β_i avec un échantillon de N observations. Il faut donc imposer une structure spatiale aux données, c'est-à-dire fournir une spécification parcimonieuse pour étudier la variation dans l'espace. La variabilité spatiale de la moyenne d'une variable ou des coefficients d'une régression peut être *discrète*, lorsque des différences systématiques entre différents régimes sont observées (paragraphe 231.1), ou *continue* sur l'ensemble de la zone étudiée (paragraphe 231.2).

231.1 Variation spatiale discrète

231.11 ANOVA spatiale

Dans un premier temps, nous nous demandons dans quelle mesure la moyenne d'une variable varie entre différents sous-groupes d'observations localisées. Par exemple, on voudrait savoir si la moyenne des prix des logements est la même à l'Est et à l'Ouest d'une ville. Une réponse possible passe par une régression où les variables explicatives sont des variables muettes représentant l'appartenance aux différents sous-groupes.

Formellement, considérons sans perte de généralité un échantillon d'observations localisées partitionné en 2 sous-groupes de localisations ⁴⁰, G_1 et G_2 . Supposons N_1 observations dans le premier sous-groupe et N_2 observations dans le second sous-groupe avec $N_1 + N_2 = N$. Le but est de tester l'hypothèse si α_1 , l'espérance de y dans G_1 est égale à α_2 , l'espérance de y dans G_2 .

⁴⁰ La généralisation à G sous-groupes est immédiate.

Il suffit alors d'estimer par les MCO le modèle suivant :

$$y_i = \alpha_1 d_{i1} + \alpha_2 d_{i2} + \varepsilon_i \quad [2.98]$$

d_{ij} sont les variables muettes égales à 1 si la région i appartient à la zone G_j et égales à 0 sinon, pour $i = 1, \dots, N$ et $j = 1, 2$.

Sous l'hypothèse de normalité des erreurs, un test d'égalité des moyennes s'effectue alors en testant l'égalité des coefficients α_1 et α_2 .

231.12 Régimes spatiaux

Dans le modèle [2.98], seule la constante varie entre les différents sous-groupes spatiaux. Plus généralement, si l'on considère l'ensemble des coefficients d'une régression, l'hétérogénéité spatiale peut être présente sous la forme de différentes constantes et/ou de différentes pentes. Dans ce cas général, on parle alors d'instabilité structurelle dans l'espace ou de régimes spatiaux (Anselin, 1988a).

Formellement, considérons le cas où deux régimes, indiqués par 1 et 2, sont considérés et où *tous* les paramètres varient selon le régime :

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad [2.99]$$

où y_1 et y_2 sont les vecteurs des observations des variables dépendantes, X_1 et X_2 sont les matrices des observations des variables explicatives, β_1 et β_2 sont les vecteurs de coefficients et ε_1 et ε_2 sont les vecteurs des erreurs respectivement des sous-groupes 1 et 2. On suppose à nouveau que le nombre total d'observations dans le sous-groupe 1 est N_1 et que celui du sous-groupe 2 est N_2 ($N_1 + N_2 = N$). Les observations dans chaque régime sont supposées être suffisamment nombreuses pour estimer β_1 et β_2 .

Les notations sont simplifiées de la façon suivante :

$$y = X^* \beta + \varepsilon \quad [2.100]$$

où y , X^* , β et ε correspondent à la structure [2.99].

Nous supposons dans ce paragraphe que la variance des termes d'erreurs est constante sur l'ensemble des observations : $E[\varepsilon\varepsilon'] = \sigma^2 I_N$. Dans ces conditions, le test d'homogénéité des coefficients, c'est-à-dire le test de changement structurel spatial $\beta_1 = \beta_2$, s'effectue grâce

au test de Chow traditionnel. Ce test exact est construit à partir des erreurs estimées sous l'hypothèse nulle et sous l'hypothèse alternative :

$$F = \frac{(\hat{e}'_c \hat{e}_c - \hat{e}'_L \hat{e}_L) / K}{\hat{e}'_L \hat{e}_L / N - 2K} \sim F(K, N - 2K) \quad [2.101]$$

où \hat{e}_c est le vecteur des erreurs estimées du modèle contraint et \hat{e}_L celui du modèle libre. De la même manière, ce test peut s'effectuer pour chacun des coefficients du modèle ou seulement pour un sous-ensemble de coefficients.

231.13 Fonctions spline

L'hétérogénéité spatiale est enfin susceptible de se manifester sous la forme de paramètres de la régression qui diffèrent, non pas selon le régime spatial, mais selon l'intervalle des valeurs prises par une ou plusieurs variables explicatives dans l'aire géographique. Par exemple, si l'on s'intéresse à l'estimation d'une fonction de densité de population sur une aire urbaine, on peut considérer, conformément aux modèles de la nouvelle économie urbaine, que la densité de population décroît lorsqu'on s'éloigne du centre de cette aire urbaine. Cependant, dans la réalité, cette décroissance n'est pas forcément monotone. Il est alors préférable de considérer qu'elle peut décroître de façon différente dans certains intervalles de valeurs de cette distance au centre.

La technique des variables muettes peut à nouveau être mobilisée pour modéliser une fonction dans laquelle les paramètres de la régression diffèrent par intervalle. L'estimation d'une telle fonction, appelée *fonction spline*, nécessite de déterminer le nombre de segments qui divisent l'axe de l'une des variables explicatives x et d'estimer une fonction linéaire à chaque segment. Supposons que l'axe des x soit divisé en M segments. Les valeurs seuils entre les segments, notées \bar{x}_j , $j = 0, \dots, M$, sont appelées « nœuds ». Pour une observation i , la fonction *spline* ou linéaire par morceaux s'écrit de la façon suivante :

$$y_i = \sum_{j=1}^M (a_j + b_j(x_i - \bar{x}_{j-1}))D_j + \varepsilon_i \quad [2.102]$$

où y est la variable dépendante, x est la variable explicative divisée en M segments, D_j est la variable muette prenant la valeur 1 pour les observations x telles que $\bar{x}_{j-1} \leq x < \bar{x}_j$ et 0 sinon. Sous les hypothèses habituelles sur les erreurs, cette fonction est estimée par les MCO.

En général, la fonction linéaire par morceaux [2.102] est discontinue aux valeurs seuils. Pour s'assurer de la continuité aux nœuds, il est nécessaire d'imposer la contrainte suivante sur les coefficients : $a_{j+1} = a_j + b_j(\bar{x}_j - \bar{x}_{j-1})$, $\forall j = 1, \dots, M - 1$.

Suits *et al.* (1978) généralisent la formulation [2.102] en posant une fonction *polynomiale* dont les paramètres diffèrent par intervalle. Utiliser une fonction polynomiale par

intervalle permet d'assurer la continuité des dérivées de la fonction aux nœuds et d'employer une forme fonctionnelle plus flexible que des fonctions linéaires. Pour des fonctions cubiques, la fonction est appelée *spline cubique*. Lorsque les segments sont de longueurs égales, Suits *et al.* (1978) montrent que cette fonction peut être obtenue en estimant la régression suivante par les MCO :

$$y_i = a_1 + b_1(x_i - \bar{x}_0) + c_1(x_i - \bar{x}_0)^2 + d_1(x_i - \bar{x}_0)^3 + \sum_{j=1}^{M-1} (d_{j+1} - d_j)(x_i - \bar{x}_j)^3 D_j^* + \varepsilon_i \quad [2.103]$$

où D_j^* est la variable muette égale à 1 pour les observations telles que $x \geq \bar{x}_j$.

Les différentes approches présentées dans le paragraphe 231.1 sur la variation spatiale discrète nécessitent la définition préalable des sous-groupes ou des intervalles qui, dans l'idéal, doivent être définis pour correspondre à un ou plusieurs changements structurels géographiques du modèle, connus *a priori*. Ainsi, dans le cas des régimes spatiaux, ces modélisations sont particulièrement adaptées pour des découpages spatiaux bien définis (« Centre » *versus* « Périphérie », rural *versus* urbain, etc.).

En l'absence d'informations précises sur les changements structurels, Anderson (1982, 1985), McMillen (1994) et McMillen et McDonald (1998) choisissent d'estimer des fonctions spline cubique avec des segments de longueur égale et des intervalles relativement courts entre les nœuds. Brueckner (1981, 1986) détermine le nombre de ruptures en effectuant un test du rapport de vraisemblance et en calculant des critères d'information pour chaque nouvelle rupture ajoutée.

Dans tous les cas, il n'existe pas à l'heure actuelle de techniques permettant à la fois une estimation endogène des régimes et une prise en compte de l'autocorrélation spatiale. Cependant, l'analyse exploratoire des données spatiales, ensemble de techniques permettant d'étudier les schémas globaux et locaux d'autocorrélation spatiale (cf. paragraphe 213, p. 105), s'avère utile pour suggérer les concentrations locales et servir de base à la définition des régimes spatiaux (Bernat, 1996).

Lorsqu'on ne dispose pas d'information sur d'éventuels régimes spatiaux, ou si l'on pense que la moyenne d'une variable ou les coefficients de la régression ne changent pas brutalement d'un régime à l'autre, il est préférable d'utiliser des spécifications permettant de capter une variation spatiale continue sur l'ensemble de la zone étudiée.

231.2 Variation spatiale continue

Deux types d'analyse sont possibles : étude de la variabilité de la moyenne ou de la variabilité des coefficients de régression. Dans le premier cas, on utilise la méthode TSA (*Trend Surface Analysis*) alors que dans le second cas, c'est la méthode VE (*Variable*

Expansion) et son extension, la méthode non-paramétrique GWR (*Geographically Weighted Regression*) qui sont les plus adaptées.

231.21 La méthode Trend Surface Analysis (TSA)

Dans cette méthode, on régresse la variable étudiée sur une expansion polynomiale des coordonnées (telles que la latitude et la longitude) de chaque localisation. Formellement, le modèle suivant est estimé par les MCO :

$$y_i = \sum_{r+s \leq p} a_{rs} u_i^r v_i^s + \varepsilon_i \quad [2.104]$$

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

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

$$y_i = a_{00} + a_{10}u_i + a_{01}v_i + \varepsilon_i \quad [2.105a]$$

$$y_i = a_{00} + a_{10}u_i + a_{01}v_i + a_{11}u_i v_i + a_{02}u_i^2 + a_{20}v_i^2 + \varepsilon_i \quad [2.105b]$$

Cette méthode permet de décrire les grandes tendances caractéristiques de la surface de régression, comme des tendances simples de type « Nord-Sud » ou « Est-Ouest » pour la fonction linéaire ou des tendances plus complexes pour les fonctions d'ordre supérieur. Elle sert donc à « lisser » les données (Ripley, 1981 ; Agterberg, 1984). Par ailleurs, comme la variable expliquée est uniquement une fonction des coordonnées des points, des valeurs prévues pour cette variable peuvent facilement être obtenues pour chaque localisation et représentées sur une carte (Bertazzon et Zaninotto, 1996).

Ce type de modélisation a, entre beaucoup d'autres applications, servi dans l'étude des fonctions de densité urbaine (Schroeder et Sjoquist, 1976), dans les modèles hédoniques de prix fonciers (Johnson et Ragas, 1987) ou immobiliers (Olmo, 1995 ; Des Rosiers et Thériault, 1995, 1996). En effet, elle permet de représenter graphiquement les surfaces de régression, leurs « pics » ou leurs « vallées », à l'aide de représentations en 3 dimensions dans lesquelles les coordonnées des points constituent l'abscisse et l'ordonnée et la valeur estimée (ou prévue) la cote (figure 2.6).

Cependant, cette méthode présente plusieurs limites (Ripley, 1981 ; Pace *et al.*, 1998). Tout d'abord, elle provoque des distorsions à la marge de la zone afin de pouvoir mieux s'ajuster aux points situés au centre. Ensuite, les variables des polynômes sont très souvent

fortement corrélées, ce qui est source de multicollinéarité. Enfin, l'autocorrélation spatiale des erreurs n'est pas systématiquement éliminée dans ce type de modèle ⁴¹.

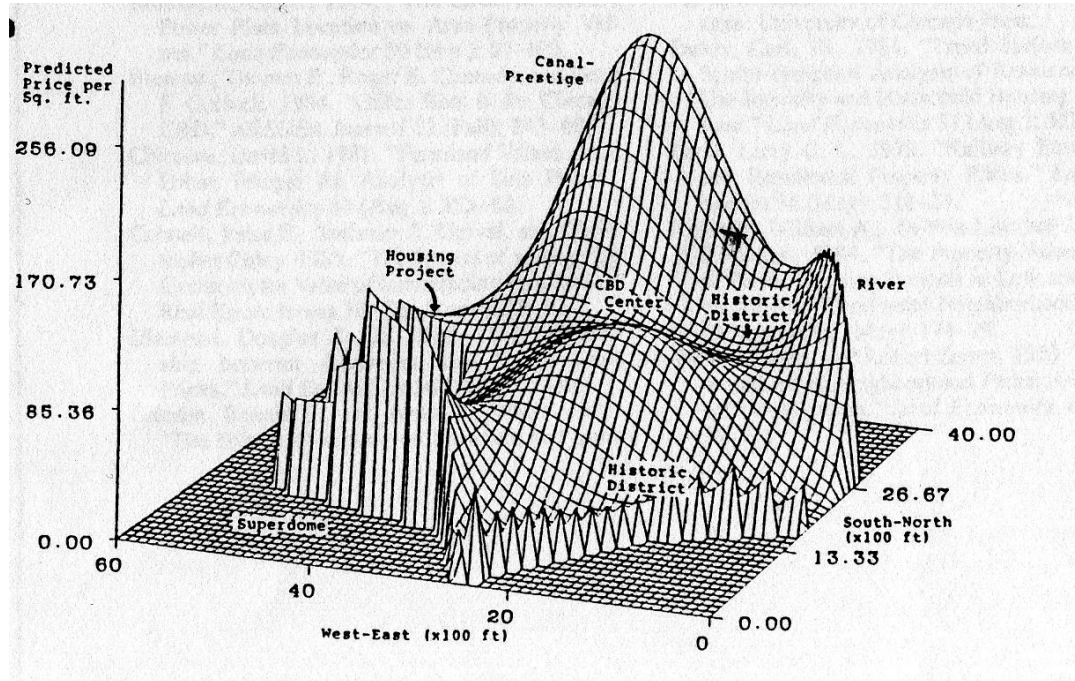


Figure 2.6 : Un exemple d'utilisation de la méthode TSA

Valeurs foncières prévues à New Orleans (source : Johnson et Ragas, 1987, p. 345)

231.22 La méthode Variable Expansion (VE)

La méthode VE, développée par Casetti (1972, 1997), généralise la méthode précédente dans la mesure où on suppose que l'hétérogénéité spatiale se traduit par des coefficients de la régression variables pour chaque localisation i . Cette différenciation est spécifiée par une fonction dépendant d'un certain nombre de variables auxiliaires.

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

$$y_i = \beta_0 + \beta_{1i}x_i + \varepsilon_i \quad [2.106]$$

⁴¹ Dans le cadre des modèles hédoniques de prix immobiliers, DesRosiers et Thériault (1999) et Thériault *et al.* (1999) développent une stratégie d'estimation et de tests combinant autocorrélation spatiale et variables polynomiales tout en évitant les problèmes de multicollinéarité.

où β_0 et β_{1i} sont les coefficients de régression et x_i correspond à la $i^{\text{ème}}$ observation de la variable explicative. On suppose que le second coefficient est variable pour chaque observation ⁴² : β_1 est une fonction exacte de variables d'« expansion » ou variables augmentées, notées $z_i, i = 1, \dots, P$:

$$\beta_{1i} = f(z_i, \gamma) \quad [2.107]$$

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

Ainsi, pour deux variables augmentées et une relation fonctionnelle linéaire, la relation [2.107] s'écrit de la façon suivante :

$$\beta_{1i} = \gamma_0 + \gamma_1 z_{1i} + \gamma_2 z_{2i} \quad [2.108]$$

L'introduction des variables augmentées dans le modèle initial donne *le modèle terminal* suivant, estimé par les MCO sous les hypothèses habituelles :

$$y_i = \beta_0 + \gamma_0 x_i + \gamma_1 (z_{1i} x_i) + \gamma_2 (z_{2i} x_i) + \varepsilon_i \quad [2.109]$$

Sous forme matricielle, le modèle [2.109] s'écrit de la façon suivante :

$$y = S\beta_0 + \gamma_0 x + \gamma_1 (z_1 x) + \gamma_2 (z_2 x) + \varepsilon \quad [2.110]$$

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

Si le modèle terminal est la spécification correcte, les estimateurs des paramètres du modèle initial sont biaisés. En effet, il s'agit d'un cas particulier du problème traditionnel de variables omises.

Soient $b' = [\beta_0 \ \gamma_0]$, $X = [S \ x]$, $Z = [z_1 x \ z_2 x]$ et $\gamma' = [\gamma_1 \ \gamma_2]$. On vérifie immédiatement que l'espérance de l'estimateur de b du modèle initial, si le modèle final est vrai, est $E(\hat{b}) = b + (X'X)^{-1} X'Z\gamma$. L'estimateur de b dans le modèle initial est biaisé (Anselin, 1988a ; Anselin et Griffith, 1988).

Comme dans la méthode TSA, les variables z_i correspondent le plus souvent aux coordonnées du point i (latitude et longitude). Dans le cas linéaire le plus simple, seules des tendances simples dans la variation des paramètres sont captées (tendance du type « Nord/Sud » ou « Est/Ouest ») et des expansions quadratiques ou d'ordre supérieur doivent

⁴² Comme précédemment, on peut également supposer la constante variable pour chaque observation.

être spécifiées pour capter des tendances plus complexes. A nouveau, plus l'ordre de l'expansion est grand, plus la multicolinéarité risque de poser problème ⁴³.

Outre les coordonnées des observations, les variables augmentées peuvent plus généralement correspondre à toutes les variables susceptibles d'être à l'origine de la différenciation spatiale des coefficients. Par exemple, LeSage (1999a) utilise la distance au centre de la ville pour modéliser la variation spatiale des prix marginaux des caractéristiques dans un modèle hédonique de prix immobiliers. Toujours dans le même cadre, Can (1990, 1992) préfère utiliser une variable composite de la « qualité » du voisinage de chaque logement : revenu moyen, composition raciale de la population, taux de chômage, etc. ⁴⁴

La méthode VE permet de modéliser des paramètres variables dans l'espace mais elle souffre de deux limites principales (Fotheringham *et al.*, 2000). La première est que cette technique ne peut capter que des *tendances* dans les relations dans l'espace, la complexité de ces tendances étant dépendante de la complexité des équations d'expansion spécifiées (tendances linéaires, quadratiques, etc.). Les estimations des paramètres obtenues à partir de cette méthode ne permettent donc pas toujours de capter d'éventuelles importantes variations locales. La seconde est que la forme des équations d'expansion doit être supposée *a priori*. Pour pallier ces problèmes, trois auteurs, Brundson, Fotheringham et Charlton, développent la méthode d'estimation non-paramétrique GWR (*Geographically Weighted Regression*).

231.23 La méthode non-paramétrique *Geographically Weighted Regression* (GWR)

La méthode GWR est présentée dans un ouvrage (Fotheringham *et al.*, 2000) et dans une série d'articles (Brundson *et al.*, 1996, 1998a, 1998b, 1999 ; Fotheringham et Brundson, 1999 ; Fotheringham *et al.*, 1996a, 1996b, 1997, 1998). Cette méthode non-paramétrique est destinée à capter au niveau de chaque observation les variations des coefficients d'une régression dans l'espace. Pour cela, un paramètre différent est estimé pour chaque observation en utilisant les valeurs des caractéristiques prises par les observations voisines.

Formellement, le point de départ est à nouveau la formulation générale [2.97] dans laquelle il existe un vecteur de K paramètres inconnus à estimer pour chaque observation i :

$$y_i = x_i' \beta_i + \varepsilon_i \quad [2.111a]$$

$$\Leftrightarrow y_i = \sum_k \beta_{ki} x_{ki} + \varepsilon_i \quad [2.111b]$$

Pour estimer les paramètres β_{ki} du modèle [2.111b], on suppose que les données observées près du point i ont plus d'influence dans l'estimation de β_{ki} que les données situées

⁴³ Casetti et Jones (1988) proposent dans ce cas de remplacer les variables augmentées initiales par leurs composantes principales (méthode VE orthogonale).

⁴⁴ Voir Casetti et Jones (1992) pour les différentes applications de cette méthode dans le cadre spatial ou non.

loin de ce point i . On utilise alors une pondération variable avec le point i . Notons $\hat{\beta}_i$ l'estimateur des Moindres Carrés Pondérés du vecteur des K coefficients pour l'observation i . Il s'écrit :

$$\hat{\beta}_i = (X'V_iX)^{-1} X'V_iY \quad [2.112]$$

où X est la matrice des observations des variables explicatives et Y représente le vecteur des observations de la variable dépendante. Enfin, V_i est une matrice diagonale (N,N) , variable pour chaque observation i , dont les éléments de la diagonale principale représentent le poids géographique accordé aux données autour du point i : $V_i = \text{diag}(v_{i1}, \dots, v_{in})$, où v_{ij} dénote le poids de la $j^{\text{ème}}$ observation sur l'estimation du modèle autour du $i^{\text{ème}}$ point.

Cette méthodologie diffère de la régression non-paramétrique habituelle du noyau (Cleveland, 1979 ; Cleveland et Devlin, 1988, Cleveland *et al.*, 1988) dans laquelle les pondérations se réfèrent à l'espace des *attributs* de X : les données plus similaires à x_i sont davantage pondérées que les données moins similaires et le résultat est un ensemble d'estimations de paramètres dans l'espace de X . En revanche, la méthode GWR utilise des pondérations basées sur l'espace *géographique*. Selon les auteurs, différents schémas de pondération V_i , ou « noyaux », sont utilisés (McMillen, 1996 ; McMillen et McDonald, 1997 ; Fotheringham *et al.*, 2000 ; Pavlov, 2000). Une possibilité consiste par exemple à retenir des poids de la forme suivante :

$$v_{ij} = \exp(-d_{ij}^2 / h^2) \quad [2.113]$$

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

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

$$v_{ij} = [1 - (d_{ij} / d_i)^3]^3 \quad I(d_{ij} < d_i) \quad [2.114]$$

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

Quel que soit le noyau utilisé, comme dans toute estimation non-paramétrique, il existe un arbitrage entre le biais et la variance. Dans [2.113] par exemple, lorsque h tend vers l'infini, les poids tendent vers 1 pour toutes les paires de points, les estimations GWR deviennent équivalentes à celles des MCO et le biais augmente. Quand la fenêtre devient plus petite, les estimations des paramètres dépendent de plus en plus des observations proches de i et ont une variance accrue. La fenêtre détermine donc le degré de « lissage » des données ainsi que le biais et la variance de l'estimateur non-paramétrique associé. Le problème est ainsi de sélectionner un noyau et une fenêtre appropriés. Le choix d'un noyau est moins important que celui de la fenêtre qui a une plus grande incidence sur les propriétés des estimateurs.

Pour choisir la valeur de h dans [2.113], Cleveland (1979) et Bowman (1984) proposent de minimiser la quantité suivante (procédure de *cross-validation*) :

$$z = \sum_{i=1,n} [y_i - y_{\neq i}(h)]^2 \quad [2.115]$$

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

Les auteurs présentent la méthode GWR comme une extension de la méthode VE. Par ailleurs, elle peut également s'interpréter comme une méthode essentiellement exploratoire permettant d'identifier la nature et les schémas d'hétérogénéité spatiale sur l'ensemble de la zone étudiée. En effet, le résultat d'une GWR est un ensemble d'estimations localisées des paramètres, ainsi que des versions localisées de mesures de qualité de la régression comme le R^2 . Ces estimations locales étant toutes associées à des localisations spécifiques, chaque ensemble de paramètres peut être cartographié pour illustrer les variations spatiales de la relation mesurée. Pour savoir si ces estimations locales des paramètres sont significativement différentes, des tests paramétriques sont proposés par Brundson *et al.* (1999) et Leung *et al.* (2000).

232 Modélisation et estimation de l'hétéroscédasticité

L'hétéroscédasticité se manifeste par la variabilité des variances des termes d'erreurs selon la localisation. Elle constitue un problème courant en coupe transversale et provient généralement de variables manquantes ou de toute autre forme de mauvaise spécification du modèle. En outre, lorsqu'on travaille sur des données localisées, les unités spatiales utilisées ne sont généralement ni de formes régulières, ni homogènes : elles peuvent avoir des formes et des aires différentes, des niveaux de développement technologique variables, des populations plus ou moins importantes, etc.

Dans ces conditions, deux stratégies sont possibles. White (1980) fournit un estimateur convergent de la matrice des variances-covariances de l'estimateur des MCO en présence d'hétéroscédasticité de forme inconnue pour que l'inférence statistique basée sur les MCO soit asymptotiquement fiable (paragraphe 232.1). Alternativement, lorsqu'on choisit une forme particulière pour l'hétéroscédasticité, des tests d'hétéroscédasticité plus puissants que le test de White peuvent être obtenus (paragraphe 232.2).

232.1 La méthode de White

Reprenons le modèle de régression suivant général [2.96] avec les mêmes notations que précédemment :

$$y = X\beta + \varepsilon$$

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

$$V(\varepsilon_i) = \sigma_i^2 \tag{2.116a}$$

soit $E(\varepsilon\varepsilon') = \Omega = \text{diag}(\sigma_i^2)$ [2.116b]

En présence d'hétéroscédasticité, les estimateurs des MCO sont non biaisés mais inefficients. Les tests de significativité doivent dans ce cas être basés sur la *vraie* matrice des variances-covariances de l'estimateur des MCO, qui est égale à $(X'X)^{-1}X'\Omega X(X'X)^{-1}$, au lieu de $\sigma^2(X'X)^{-1}$. Lorsqu'on ne connaît pas la vraie forme de l'hétéroscédasticité, White (1980) démontre que, même si la matrice Ω ne peut pas être estimée de manière convergente, on peut tout de même estimer de manière convergente (sous des conditions de régularité générales) $X'\Omega X$ par $X'GX$ où G est la matrice diagonale des erreurs estimées élevées au carré ⁴⁵. Par conséquent, un test de significativité pour les coefficients de régression est alors basé sur les MCO et la matrice des variances-covariances estimée $(X'X)^{-1}X'GX(X'X)^{-1}$.

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

$$\begin{aligned} H_0 : \sigma_i^2 &= \sigma^2, \forall i \\ H_1 : &\text{non } H_0 \end{aligned} \tag{2.117}$$

⁴⁵ Pour les petits échantillons, MacKinnon et White (1985) proposent d'ajuster les résidus élevés au carré de la matrice G par le terme $1/(1 - k_{ii})$ où k_{ii} est le $i^{\text{ème}}$ élément de la matrice idempotente $X(X'X)^{-1}X'$.

Soit $\hat{\varepsilon}_i$ l'erreur estimée pour la $i^{\text{ème}}$ observation. Pour trouver la statistique du test, on calcule NR^2 dans la régression de $\hat{\varepsilon}_i^2$ sur une constante, les régresseurs de l'équation estimée, leurs carrés et leurs produits croisés. Cette statistique est asymptotiquement distribuée selon une loi du chi-deux à $S - 1$ degrés de liberté, où S est le nombre de régresseurs, sans la constante.

Le test de White est un test très général qui ne nécessite pas de connaissances *a priori* sur la forme de l'hétéroscédasticité. Si l'échantillon est assez grand, il est suffisamment puissant pour détecter toutes les formes d'hétéroscédasticité pour lesquelles la matrice des variances-covariances usuelle n'est pas convergente. Cependant, même si le test de White est convergent contre une grande variété d'alternatives hétéroscédastiques, il peut ne pas être très puissant en échantillon fini. Enfin, il faut souligner que ce test peut également identifier un problème dans la spécification du modèle ou une dépendance entre les régresseurs et les erreurs.

232.2 Spécification de l'hétéroscédasticité

Lorsque l'hétéroscédasticité est correctement spécifiée, les tests d'hétéroscédasticité associés sont alors plus puissants que le test de White (Greene, 2000). Deux formes possibles particulièrement adaptées au cas spatial sont présentées dans ce paragraphe ⁴⁶.

232.21 La variance est fonction d'un ensemble de variables

La première possibilité consiste à exprimer la variance des termes d'erreurs comme une fonction d'un ensemble de variables :

$$\sigma_i^2 = \sigma^2 f(z_i' \alpha) \quad [2.118]$$

où σ^2 est un facteur d'échelle, f une forme fonctionnelle, z_i est un vecteur de P variables, $z_i' \alpha = \alpha_0 + \alpha_1 z_{1i} + \dots + \alpha_p z_{pi}$. Enfin, α_i , $i = 1 \dots P$ sont des paramètres inconnus à estimer.

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

⁴⁶ Certaines modélisations visant à capter l'instabilité structurelle produisent aussi une certaine forme d'hétéroscédasticité (voir *infra*, paragraphe 232. 3, p. 172).

Le modèle est homoscédastique si $\alpha_i = 0, i = 1 \dots P$ dans [2.118]. Cette forme des erreurs est à la base du test de Breusch-Pagan (1979). Il s'agit d'un test du multiplicateur de Lagrange relativement puissant dès lors qu'il est possible d'identifier les variables susceptibles d'être à la source de l'hétéroscédasticité.

Soit Z la matrice $(N, P + 1)$ comportant les observations sur $(1, z_i)$ et soit g le vecteur $(N, 1)$ comportant les observations sur $g_i = \hat{\varepsilon}_i^2 / (\hat{\varepsilon}' \hat{\varepsilon} / N)$. La statistique du test pour l'hypothèse $H_0 : \alpha_i = 0, i = 1 \dots P$ s'écrit alors :

$$BP = \frac{1}{2} (g' Z (Z' Z)^{-1} Z' g - N) \quad [2.119]$$

Sous l'hypothèse nulle d'homoscédasticité, BP converge asymptotiquement vers une loi du chi-deux à P degrés de liberté⁴⁷. Une procédure asymptotiquement équivalente consiste à régresser $\hat{\varepsilon}_i^2$ sur Z et à calculer NR^2 . Cette quantité converge asymptotiquement vers une loi du chi-deux à P degrés de liberté sous l'hypothèse nulle.

232.22 Hétéroscédasticité en groupes

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

$$\Omega = \begin{bmatrix} \sigma_1^2 I_{N_1} & 0 \\ 0 & \sigma_2^2 I_{N_2} \end{bmatrix} \quad [2.120]$$

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

L'estimation de ce modèle peut s'effectuer par la méthode du maximum de vraisemblance (MV). Dans le cas le plus général, considérons G groupes différents. Il y a N_g observations dans chaque groupe tel que $N_1 + N_2 + \dots + N_G = N$. σ_g^2 représente la variance de l'erreur dans le $g^{\text{ème}}$ groupe.

⁴⁷ Ce test étant sensible à l'hypothèse de normalité, Koenker (1981) et Koenker et Basset (1982) proposent un test plus robuste à la non-normalité des erreurs.

Dans ce cas, la fonction de log-vraisemblance s'écrit :

$$\ln L_1(y|\beta, \sigma_1^2, \dots, \sigma_G^2) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \sum_{g=1}^G (N_g \ln \sigma_g^2) - \frac{1}{2} \sum_{g=1}^G \sum_{i=1}^{N_g} \left(\frac{1}{\sigma_g^2} \varepsilon_{ig}^2 \right) \quad [2.121]$$

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

$$\hat{\beta} = \left[\sum_g \frac{1}{\sigma_g^2} X_g' X_g \right]^{-1} \left[\sum_g \frac{1}{\sigma_g^2} X_g' y_g \right] \quad [2.122]$$

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

Dans ces conditions, l'estimateur du maximum de vraisemblance de β s'obtient en appliquant une méthode itérative. Dans un premier temps, on estime β par les MCO. Ensuite, les variances sont estimées séparément pour chaque groupe : $\hat{\sigma}_g^2 = (\hat{\varepsilon}_g' \hat{\varepsilon}_g) / N_g$. Dans un troisième temps, on recalcule $\hat{\beta}$ avec [2.122]. On recommence les deux dernières étapes tant que le critère de convergence numérique pour $\hat{\beta}$ n'est pas atteint.

Dans le cas particulier d'une hétéroscédasticité en groupes, sous l'hypothèse de normalité, le test du ratio de vraisemblance est un test uniformément plus puissant. Sous l'hypothèse nulle d'homoscédasticité (variances identiques), la fonction de log-vraisemblance du modèle de régression linéaire classique est :

$$\ln L_0(y|\beta, \sigma^2) = -\frac{N}{2} [(\ln(2\pi) + \ln \sigma^2)] - \frac{1}{2\sigma^2} \sum_i \varepsilon_i^2 \quad [2.123]$$

L'estimation MV de σ^2 est donnée par $\hat{\sigma}^2 = (\hat{\varepsilon}' \hat{\varepsilon}) / N$.

Sous l'hypothèse alternative d'hétéroscédasticité entre les G groupes, cette fonction est donnée par [2.121]. Dans ces conditions, si on évalue $\ln L_0$ et $\ln L_1$ à l'optimum, le test du ratio de vraisemblance est donné par :

$$-2(\ln L_0 - \ln L_1) = N \ln \hat{\sigma}^2 - \sum_g N_g \ln \hat{\sigma}_g^2 \quad [2.124]$$

Sous l'hypothèse nulle, cette statistique est asymptotiquement distribuée selon une loi du chi-deux à $G - 1$ degré de liberté.

232.3 Hétéroscédasticité et instabilité des paramètres

L'instabilité spatiale peut être modélisée, par exemple, par la spécification de régimes spatiaux sous la forme [2.99]. Pourtant, lorsque l'instabilité spatiale n'est pas complètement prise en compte de cette façon, il peut subsister une hétéroscédasticité qui représente l'hétérogénéité non observée. Ainsi, il est possible qu'une hétéroscédasticité résiduelle se manifeste dans le cadre d'un modèle avec régimes spatiaux. Nous présentons ce point dans le paragraphe 232.31. En revanche, la situation est différente dans le cadre du modèle VE [2.109] dans lequel l'hétéroscédasticité des erreurs est induite par la structure même du modèle. Ce cas est abordé dans le paragraphe 232.32.

232.31 Régimes spatiaux et hétéroscédasticité

Reprenons les notations du modèle [2.99] où deux régimes spatiaux sont considérés :

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

Quandt (1958) suppose en outre que les termes d'erreurs ont une variance différente pour chaque sous-ensemble (hétéroscédasticité en blocs, modèle [2.120]). Dans ce cas, l'estimation peut être basée sur la fonction de log-vraisemblance du modèle global qui est de la forme (en ignorant le terme constant) :

$$\ln L = -\left(\frac{N_1}{2}\right) \ln(\sigma_1)^2 - \left(\frac{N_2}{2}\right) \ln(\sigma_2)^2 - \frac{(y_1 - X_1\beta_1)'(y_1 - X_1\beta_1)}{2(\sigma_1)^2} - \frac{(y_2 - X_2\beta_2)'(y_2 - X_2\beta_2)}{2(\sigma_2)^2} \quad [2.125]$$

Cette fonction de log-vraisemblance peut être modifiée pour tenir compte de structures d'erreurs plus complexes (comme par exemple, l'incorporation d'une autocorrélation spatiale des erreurs, voir *infra*, paragraphe 233, p. 174).

232.32 Variation spatiale continue et hétéroscédasticité

Certains auteurs (Anselin, 1988a, 1992 ; Casetti, 1982 ; Casetti et Can, 1999) soulignent que la méthode VE est également susceptible d'être affectée par un problème d'hétéroscédasticité. Reprenons en effet ce modèle (équations [2.106] à [2.109]) :

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

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

Jusqu'à présent, nous avons supposé que la forme exacte de l'expansion est *a priori*. Pourtant, en pratique, l'hypothèse d'une relation exacte (déterministe) entre les coefficients et leurs variables augmentées est difficile à tenir (Anselin, 1988a, 1992). Il est donc plus pertinent de rajouter un terme aléatoire dans l'expansion linéaire [2.108] :

$$\beta_{1i} = \gamma_0 + \gamma_1 z_{1i} + \gamma_2 z_{2i} + v_i \quad [2.126]$$

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

$$y = S\beta_0 + \gamma_0 x + \gamma_1(z_1 x) + \gamma_2(z_2 x) + vx + \varepsilon \quad [2.127]$$

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

Le modèle final induit donc une certaine forme d'hétéroscédasticité dont il faut tenir compte pour l'inférence statistique, en particulier si l'on cherche à tester la significativité de γ_1 et γ_2 . Cette forme particulière d'hétéroscédasticité peut être testée avec le test de Breusch-Pagan. Cette approche revient alors à tester la version aléatoire contre la version déterministe de la méthode VE : sous l'hypothèse nulle d'homoscédasticité, on retrouve le modèle [2.108] alors que sous l'hypothèse alternative d'hétéroscédasticité, on retrouve le modèle [2.127].

Dans le cas où l'hétéroscédasticité impliquée par la méthode des variables augmentées est liée à la distance à un point du plan (par exemple la distance à un point central), Casetti (1982) et Casetti et Can (1999) proposent la modélisation DARP (*Drift Analysis of Regression Parameters*). Les auteurs supposent que la variance du terme d'erreur associée à chaque observation i dans [1.26] est une fonction de la distance entre cette observation et le point de référence. Il s'agit alors d'un modèle hétéroscédastique et l'estimation des paramètres de ce modèle requiert soit les MCQG, soit la méthode du maximum de vraisemblance.

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

Formellement, pour chaque observation i , on pose :

$$y_i = x_i' \beta_i \quad \text{avec} \quad \beta_i = \beta + u_i \quad [2.128]$$

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

$$y_i = x_i' \beta + v_i \quad \text{avec} \quad v_i = x_i' u_i \quad [2.129]$$

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

Le modèle [2.129] s'estime par la méthode du maximum de vraisemblance ou par la méthode des moindres carrés généralisés. Avant l'estimation du modèle plus complexe, la présence de coefficients aléatoires peut être testée grâce à un test d'hétéroscédasticité, comme celui de Breusch-Pagan (1979). Sous l'hypothèse nulle d'homoscédasticité, on retrouve alors le modèle de régression simple.

Les méthodes d'estimation et de test de l'hétérogénéité, présentées dans les paragraphes précédents, permettent de traiter différentes formes d'hétérogénéité, dont l'hétérogénéité spatiale lorsque l'on soupçonne que l'instabilité des paramètres et/ou l'hétéroscédasticité sont liées à la structure spatiale des données. En revanche, des problèmes spécifiques d'estimation et de tests apparaissent lorsque l'autocorrélation et l'hétérogénéité spatiales sont présentes simultanément.

233 L'interaction entre autocorrélation et hétérogénéité spatiales

L'autocorrélation et l'hétérogénéité spatiales sont souvent présentes conjointement dans une régression. Nous avons déjà souligné que ces deux effets sont liés dans le modèle avec variable endogène décalée ainsi que dans le modèle avec autocorrélation spatiale des erreurs (cf. paragraphe 221.31, p. 124). Plus généralement, nous montrons dans le paragraphe 233.1 (p. 175) que les liens unissant les deux effets spatiaux sont nombreux et complexes. Dès lors, les tests d'autocorrélation spatiale présentés dans la section 22 doivent être adaptés en présence d'une hétéroscédasticité de forme connue ou non (paragraphe 233.2). Symétriquement, les tests d'hétéroscédasticité (paragraphe 233.3) et d'instabilité structurelle (paragraphe 233.4) doivent également être ajustés à la présence d'autocorrélation spatiale.

233.1 Les liens entre hétérogénéité et autocorrélation spatiales

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

Premièrement, il n'est pas toujours facile de distinguer les conséquences de l'autocorrélation de celles de l'hétérogénéité spatiale dans une coupe transversale : une équivalence observationnelle entre ces deux effets peut exister (Anselin et Bera, 1998 ; Anselin, 2001b). Par exemple, l'observation d'un regroupement spatial de valeurs très élevées des erreurs estimées peut provenir d'une hétérogénéité (qui se manifeste notamment sous la forme d'une hétéroscédasticité en blocs) ou d'une autocorrélation spatiale (c'est-à-dire un processus stochastique impliquant le regroupement de points extrêmes).

Deuxièmement, les tests pour l'hétéroscédasticité et pour l'instabilité structurelle ne sont pas fiables en présence d'autocorrélation spatiale. Des simulations de Monte-Carlo effectuées par Anselin et Griffith (1988) montrent, par exemple, que la présence d'autocorrélation spatiale des erreurs affecte le seuil et la puissance des tests de White (1980) et de Breusch-Pagan (1979). En particulier, sous l'hypothèse nulle d'absence d'hétéroscédasticité, les fréquences empiriques de rejet du test de White et de Breusch-Pagan sont inférieures au niveau de signification choisi en présence d'autocorrélation spatiale. Les puissances des tests sont également affectées : elles diminuent pour des valeurs positives du paramètre spatial. Réciproquement, les tests d'autocorrélation spatiale voient leurs propriétés affectées lorsqu'il y a hétéroscédasticité de forme inconnue.

Le test de Chow, permettant de tester l'instabilité structurelle et basé sur les résidus de l'estimation par les MCO, est également affecté par la présence d'autocorrélation spatiale des erreurs, particulièrement dans le cas de régimes correspondant à deux ensembles compacts d'unités spatiales contiguës (configuration de type Est-Ouest ou Nord-Sud par exemple). Les simulations de Monte-Carlo réalisées par Anselin (1990a) montrent, en effet, que dans ce cas, l'hypothèse nulle d'homogénéité régionale est trop souvent rejetée pour des valeurs positives du paramètre spatial alors que la puissance du test augmente presque exponentiellement.

La présence jointe d'hétéroscédasticité et de dépendance spatiale nécessite donc des tests spécialisés et des méthodes d'estimation adaptées (Anselin, 1988a, 1988b).

Troisièmement, l'autocorrélation spatiale est parfois le résultat d'une instabilité des paramètres non modélisée (Anselin et Griffith, 1988 ; Brundson *et al.*, 1999). En d'autres termes, si des relations variables dans l'espace sont modélisées à l'aide d'une régression « globale » (spécifiée de la même façon pour toutes les observations), les termes d'erreurs peuvent être spatialement autocorrélés. Par exemple, supposons une zone géographique pour laquelle y est régressé sur x et que le modèle global estimé qui en résulte est :

$$\hat{y}_i = 1,5 + 2x_i \quad [2.130]$$

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

$$\text{région 1 : } y_i = 1,5 + 2,5x_i \quad [2.131a]$$

$$\text{région 2 : } y_i = 1,5 + 1,5x_i \quad [2.131b]$$

Dans ces conditions, appliquer l'équation [2.130] à la partie de la zone pour laquelle l'équation [2.131a] est valable conduit à la sous-estimation des valeurs de y_i dans cette zone et des erreurs estimées positives (pour des valeurs positives de y_i). Réciproquement, appliquer l'équation [2.130] à la partie de la zone pour laquelle l'équation [2.131b] est valable conduit à la surestimation des valeurs de y_i dans cette zone et des erreurs estimées négatives. Une forte autocorrélation spatiale des erreurs apparaît donc, résultant de l'incapacité du modèle global à traiter l'instabilité structurelle de la relation.

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

Enfin, certaines formes de mauvaises spécifications du modèle mais aussi des erreurs de mesure, causes d'hétéroscédasticité, peuvent aussi être à la source d'une autocorrélation spatiale des erreurs.

Tous ces éléments suggèrent des liens entre hétérogénéité spatiale et dépendance spatiale forts et complexes. Les deux aspects ne peuvent pas être considérés indépendamment l'un de l'autre et ils doivent être spécifiés très précisément pour pouvoir identifier les paramètres du problème et tester chaque type de mauvaise spécification. D'une part, la dépendance spatiale ne se teste plus de la même façon en présence d'hétéroscédasticité, connue ou non, et d'autre part, il faut adapter les tests d'hétéroscédasticité et d'instabilité structurelle à la présence d'autocorrélation spatiale.

233.2 Le test de la dépendance spatiale en présence d'hétéroscédasticité

Les tests de l'autocorrélation spatiale varient suivant que la forme précise de l'hétéroscédasticité est connue (paragraphe 233.21) ou non (paragraphe 233.22).

233.21 *Le test de l'autocorrélation spatiale en présence d'hétéroscédasticité de forme connue*

Dans ce premier cas, il s'agit des modèles de régression linéaire avec une forme pré-spécifiée d'hétéroscédasticité. Ainsi, une hypothèse particulière est parfois faite sur le modèle. Ce cas de figure peut intervenir dans diverses situations, comme par exemple lorsque des coefficients aléatoires sont inclus.

Formellement, considérons le modèle avec autocorrélation spatiale des erreurs (sans variable autorégressive) et hétéroscédasticité sur le terme résiduel :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \quad u \sim N(0, \Omega) \end{aligned} \quad [2.132]$$

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

$$LM = [\hat{\varepsilon}'\Omega^{-1}W\hat{\varepsilon}]^2 / T^* \quad [2.133]$$

$\hat{\varepsilon}$ est le vecteur contenant les erreurs estimées du modèle sous l'hypothèse nulle, estimé par la méthode du maximum de vraisemblance et $T^* = \text{tr}\{(W + W')W\}$. Cette statistique converge asymptotiquement vers une loi du chi-deux à 1 degré de liberté.

Dans le modèle avec variables augmentées et hétéroscédasticité [2.127], la forme de l'hétéroscédasticité est également connue. Dans ce cas, un test de l'autocorrélation spatiale (test du multiplicateur de Lagrange) est proposé par Anselin (1992).

233.22 *Le test de l'autocorrélation spatiale en présence d'hétéroscédasticité de forme inconnue*

Souvent, l'hétéroscédasticité est susceptible d'être présente sans que sa forme précise soit connue. Par conséquent, le test précédent n'est pas applicable. De plus, un mauvais choix de la spécification de l'hétéroscédasticité diminue la puissance du test du multiplicateur de Lagrange. Il faut donc avoir un test robuste de l'autocorrélation spatiale en présence d'une hétéroscédasticité dont la forme précise est inconnue. Anselin (1990b) propose deux tests robustes à l'hétéroscédasticité, l'un pour l'autocorrélation spatiale des erreurs et l'autre pour la variable endogène décalée, en s'appuyant sur le test de significativité des coefficients robuste à l'hétéroscédasticité proposé par Davidson et MacKinnon (1985) dans le cadre non spatial.

Davidson et MacKinnon (1985) développent, comme extension de l'approche de White, un test de significativité des coefficients de régression robuste à l'hétéroscédasticité. Formellement, ils testent l'hypothèse nulle $\gamma = 0$ dans le modèle suivant :

$$\begin{aligned} H_0 : y &= X\beta + u \\ H_1 : y &= X\beta + Z\gamma + u \end{aligned} \quad [2.134]$$

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

$$DM_1 = y' MZ(Z' M\Omega(u)MZ)^{-1} Z' My \quad [2.135]$$

M est la matrice de projection habituelle ($M = I - X(X'X)^{-1}X'$) et $\Omega(u)$ est la matrice diagonale contenant les erreurs estimées élevées au carré sous H_0 . Cette statistique converge asymptotiquement vers une loi du chi-deux à R degrés de liberté.

Il apparaît que cette statistique est en fait égale à N moins la somme des carrés des erreurs estimées de la régression auxiliaire suivante : $S = UMZ\gamma + \text{erreur}$, où S est le vecteur somme et U est la matrice diagonale des erreurs estimées des MCO.

La présence de l'autocorrélation spatiale des erreurs ou d'une variable endogène décalée étant liée à la significativité des paramètres spatiaux dans les régressions, deux tests robustes à l'hétéroscédasticité peuvent donc être proposés dans ce cadre.

1) Test de l'autocorrélation spatiale des erreurs robuste à l'hétéroscédasticité

Anselin (1990b) dérive un test de l'autocorrélation spatiale des erreurs $\lambda = 0$ en utilisant le modèle de Durbin contraint sous l'hypothèse alternative :

$$\begin{aligned} H_0 : y &= X\beta + u \\ H_1 : y &= X\beta + \lambda Wy - \lambda WX\beta + u \end{aligned} \quad [2.136]$$

Sous l'hypothèse alternative, le modèle doit être estimé par : $y = X\beta + \lambda Wy + WX\gamma + u$ avec la contrainte $\lambda\beta = -\gamma$. Cette estimation peut se faire par la méthode des variables instrumentales.

Soit Q la matrice des instruments et Z la matrice : $[Wy \quad WX]$ de dimension $(N, K + 1)$. Il faut prendre en compte les contraintes sur les paramètres explicitement dans une formulation

non linéaire, estimée par la méthode des variables instrumentales non linéaire $y = f(\beta, \lambda)$, avec les dérivées :

$$\partial f / \partial \beta = X - \lambda WX \quad [2.137a]$$

et $\partial f / \partial \lambda = Wy - WX\beta \quad [2.137b]$

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

$$DM_2 = (y - f)' MPF(\lambda) [F(\lambda)' PM\Omega(u) MPF(\lambda)]^{-1} F(\lambda)' PM(y - f) \quad [2.138]$$

où $y - f$ sont les erreurs estimées des MCO, $\Omega(u)$ la matrice diagonale des erreurs estimées des MCO élevées au carré, $M = I - PF(\beta)[F(\beta)' PF(\beta)]^{-1} F(\beta)' P$, $P = Q(Q'Q)^{-1} Q'$, $F(\lambda)$ et $F(\beta)$ sont les matrices des dérivées partielles, évaluées sous l'hypothèse nulle $\lambda = 0$. Cette statistique converge asymptotiquement vers une loi du chi-deux à 1 degré de liberté.

La statistique DM_2 est égale à N moins la somme des carrés des erreurs estimées de l'équation auxiliaire suivante : $S = U MPF(\lambda)\gamma + erreur$, où U est la matrice diagonale des erreurs estimées des MCO, $PF(\lambda)$ sont les erreurs estimées de la régression de $PF(\lambda)$ sur $PF(\hat{\beta})$ ou de $P(Wy - WX\hat{\beta})$ sur PX .

2) Test d'une variable endogène décalée robuste à l'hétéroscédasticité

Le test porte sur la nullité de ρ dans le modèle suivant sous l'hypothèse alternative (Anselin, 1990b) :

$$H_0 : y = X\beta + u \quad [2.139]$$

$$H_1 : y = X\beta + \rho Wy + u$$

La présence d'une variable endogène décalée nécessite une estimation par le maximum de vraisemblance ou les variables instrumentales. Dans ce dernier cas, l'extension de la statistique de Davidson et McKinnon donne :

$$DM_3 = y' MPWy [y' W' PM\Omega(u) MWy]^{-1} y' W' PMy \quad [2.140]$$

où $P = Q(Q'Q)^{-1} Q'$, Q est une matrice d'instruments, $\Omega(u)$ est la matrice diagonale des erreurs estimées élevées au carré, sous l'hypothèse nulle. Cette statistique converge asymptotiquement vers une loi du chi-deux à 1 degré de liberté.

Il s'avère que cette statistique est équivalente à N moins la somme des carrés des erreurs estimées de la régression auxiliaire : $S = U MPWy\gamma + erreur$, où U est la matrice diagonale des erreurs estimées des MCO.

Aucune simulation établissant les propriétés de biais et de puissance de ces tests en échantillon fini n'a encore été réalisée.

233.3 Le test de l'hétéroscélasticité en présence de dépendance spatiale

En présence d'autocorrélation spatiale, les propriétés des tests d'hétéroscélasticité ne sont plus valables et face à ce problème, deux stratégies sont possibles (Anselin, 1988a). La première stratégie consiste tout d'abord, à faire un test joint d'hétéroscélasticité et d'autocorrélation spatiale des erreurs et ensuite, si l'hypothèse nulle est rejetée, à tester séparément ces deux effets (paragraphe 233.31). La seconde stratégie consiste à effectuer un test d'hétéroscélasticité en présence d'autocorrélation spatiale des erreurs (paragraphe 233.32)

233.31 Test joint d'hétéroscélasticité et d'autocorrélation spatiale des erreurs

Formellement, considérons le modèle avec autocorrélation spatiale des erreurs et hétéroscélasticité sur le terme d'erreur résiduel :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \quad u \sim N(0, \Omega) \end{aligned} \quad [2.141]$$

où les éléments de la diagonale principale de Ω sont de la forme [2.118] :
 $\sigma_i^2 = \sigma^2 f(\alpha_0 + \alpha_1 z_{1i} + \dots + \alpha_p z_{pi})$.

Le modèle [2.141] peut s'estimer avec la méthode du maximum de vraisemblance et la fonction de log-vraisemblance s'écrit :

$$\ln L(y|\beta, \lambda, \Omega) = -\frac{N}{2} \ln(2\pi) - \frac{1}{2} \ln|\Omega| + \ln|I - \lambda W| - \frac{1}{2} v'v \quad [2.142a]$$

$$v'v = (y - X\beta)'(I - \lambda W)' \Omega^{-1} (I - \lambda W)(y - X\beta) \quad [2.142b]$$

La première étape consiste à tester l'hypothèse nulle jointe $H_0 : \lambda = 0$ et $\alpha_i = 0, i = 1 \dots P$ dans [2.141]. La statistique du multiplicateur de Lagrange est la suivante (Anselin, 1988a) :

$$JLM = (1/2) f' Z (Z' Z)^{-1} Z' f + (1/T^*) [\hat{\varepsilon}' W \hat{\varepsilon} / \hat{\sigma}^2]^2 \quad [2.143]$$

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

Cette statistique JLM est la somme de la statistique de Breusch-Pagan et de la statistique LM_{ERR} . Elle converge asymptotiquement vers une loi du chi-deux à $P + 1$ degré de liberté.

Kelejian et Robinson (1998) proposent une autre formulation de ce test qui n'est pas basé sur le principe du maximum de vraisemblance et qui n'exige pas *a priori* un modèle linéaire, des erreurs normales et une forme d'hétéroscédasticité connue : seules les variables susceptibles d'être à l'origine de l'hétéroscédasticité doivent être connues à l'avance. Ce test est noté $KR-SPHET$. Si l'hypothèse nulle est rejetée, un test pour chaque cas doit être effectué afin de déterminer l'origine du rejet.

Des simulations de Monte-Carlo sont effectuées par Kelejian et Robinson (1998) sur les performances de ces différents tests. Ces simulations suggèrent qu'en l'absence d'hétéroscédasticité, le test de Moran est le meilleur pour détecter l'autocorrélation spatiale même si les erreurs ne sont pas normales. Si le paramètre d'autocorrélation spatiale est élevé, le test de Moran, les tests LM_{ERR} et JLM et le test $KR-SPHET$ sont puissants qu'il y ait ou non de l'hétéroscédasticité. En revanche, si de l'hétéroscédasticité est présente, JLM et le test $KR-SPHET$ sont beaucoup plus puissants que les tests de Moran et LM_{ERR} pour détecter au moins un des deux problèmes : autocorrélation spatiale ou hétéroscédasticité.

En conclusion, si l'on pense que la présence d'hétéroscédasticité est peu probable, il est préférable d'utiliser le test de Moran et/ou le test LM_{ERR} . En revanche, s'il y a possibilité d'hétéroscédasticité, alors il vaut mieux utiliser les deux autres tests JLM et $KR-SPHET$.

233.32 Le test de l'hétéroscédasticité en présence d'autocorrélation spatiale des erreurs

Il s'agit de tester $H_0 : \alpha_i = 0, i = 1 \dots P$ dans le modèle [2.141]. Sous l'hypothèse nulle, ce modèle devient donc un modèle homoscédastique avec autocorrélation spatiale des erreurs :

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \lambda W\varepsilon + u \quad u \sim N(0, \sigma^2 I) \end{aligned} \quad [2.144]$$

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

$$BPS = (1/2) fZ^{-1} [Z' DZ]^{-1} Z' f \quad [2.145]$$

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

Cette statistique converge asymptotiquement vers une loi du chi-deux à P degrés de liberté.

Ce test s'effectue après avoir estimé le modèle avec autocorrélation des erreurs [2.144] alors que le test joint [2.143] peut s'effectuer à partir des erreurs estimées des MCO du modèle de régression standard. Un test de Breusch-Pagan spatialement ajusté peut aussi être dérivé pour le modèle avec variable endogène décalée et avec hétéroscédasticité. Cependant, comme pour les tests présentés dans le paragraphe 233.22, les propriétés de ces deux tests en échantillon fini ne sont pas connues.

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

Lorsque les hypothèses d'homoscédasticité et de non-corrélation ne sont pas vérifiées dans les modèles ANOVA [2.98], TSA [2.104] ou VE [2.109], les MCO fournissent des estimateurs sans biais mais inefficients et l'inférence statistique n'est pas fiable. Par exemple, s'il y a autocorrélation spatiale positive des erreurs dans un modèle ANOVA spatial [2.98], la variance intra-groupe est sous-estimée et l'hypothèse nulle d'égalité entre les moyennes des différentes zones géographiques est rejetée trop souvent (Cliff et Ord, 1981 ; Legendre *et al.*, 1990 ; Sokal *et al.*, 1993). Il faut donc estimer le modèle [2.98] grâce à une méthode adaptée dans le but d'ajuster les sommes des carrés associés aux variances intra- et inter-groupe (Griffith, 1978). D'une façon générale, il faut tester dans les différents modèles d'instabilité des paramètres la présence d'hétéroscédasticité et d'autocorrélation spatiale et si nécessaire estimer par des méthodes appropriées des modèles avec hétéroscédasticité et/ou autocorrélation spatiale (Griffith, 1978, 1992a).

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

233.41 Les différents cas possibles

Formellement, reprenons les notations du modèle avec deux régimes [2.99] :

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

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

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

On veut tester l'hypothèse d'homogénéité régionale, soit : $H_0 : \beta_1 = \beta_2$. La forme de la statistique du test dépend des hypothèses faites sur les erreurs. Dans le paragraphe 231.12, nous avons vu que lorsque la variance des termes d'erreurs est constante sur l'ensemble des observations, le test traditionnel de Chow s'applique. En revanche, lorsque les erreurs ne sont pas homoscédastiques, le test d'instabilité structurelle doit être établi sur des bases asymptotiques. Différents cas sont envisageables.

1^{er} cas : hétéroscélasticité en blocs

Tout d'abord, on peut, comme dans la première section, envisager une structure d'hétéroscélasticité en blocs (modèle [2.120]). La matrice des variances-covariances de ε s'écrit alors :

$$\Psi = \begin{bmatrix} \sigma_1^2 I_{N_1} & 0 \\ 0 & \sigma_2^2 I_{N_2} \end{bmatrix} \quad [2.146]$$

Lorsque l'autocorrélation spatiale est présente dans les erreurs, quatre situations différentes peuvent être distinguées selon la structure de la variance Ψ dans le modèle [2.99]. Dans les deux premières, la dépendance des erreurs est reliée au système spatial complet alors que dans les deux dernières, la structure de la dépendance spatiale est différente pour chaque régime.

2^{ème} cas : Homoscélasticité et autocorrélation spatiale

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

$$\varepsilon = \lambda W \varepsilon + u \quad [2.147]$$

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

Si les erreurs u sont homoscédastiques, c'est-à-dire si $E(uu') = \sigma^2 I$, alors la matrice des variances-covariances de ε prend la forme :

$$\Psi = \sigma^2 (B' B)^{-1} \quad [2.148]$$

3^{ème} cas : Hétéroscélasticité en blocs et autocorrélation spatiale

Reprenons la structure [2.147] pour ε , mais supposons le modèle hétéroscélastique où chaque régime a une variance des erreurs différentes.

Dans ce cas, la matrice des variances-covariances pour u devient :

$$E[u u'] = \Omega = \begin{bmatrix} \sigma_1^2 I_1 & 0 \\ 0 & \sigma_2^2 I_2 \end{bmatrix} \quad [2.149]$$

$$\Rightarrow \Psi = B^{-1} \Omega (B')^{-1} \quad [2.150]$$

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

Lorsque la structure de dépendance spatiale est différente pour chaque sous-régime, les erreurs ε_1 et ε_2 dans [2.120] suivent chacune un processus spatial différent. Ces processus sont reflétés respectivement par les matrices de poids W_1 et W_2 :

$$\begin{aligned} \varepsilon_1 &= \lambda_1 W_1 \varepsilon_1 + u_1 \\ \varepsilon_2 &= \lambda_2 W_2 \varepsilon_2 + u_2 \end{aligned} \quad [2.151]$$

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

Supposons tout d'abord que les erreurs des deux sous-régimes sont indépendantes : $E[u_{h_1} u_{h_2}] = 0$ pour tout (h, k) appartenant aux deux sous-régimes. La matrice des variances-covariances de chaque sous-régime est alors de la forme :

$$\Psi_i = \sigma_i^2 (B_i' B_i)^{-1} \text{ pour } i = 1, 2 \quad [2.152]$$

La matrice des variances-covariances totale est :

$$\Psi = \begin{bmatrix} \sigma_1^2 (B_1' B_1)^{-1} & 0 \\ 0 & \sigma_2^2 (B_2' B_2)^{-1} \end{bmatrix} \quad [2.153]$$

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

Reprenons la structure [2.151] pour les erreurs. Lorsqu'il y a autocorrélation spatiale à la fois à l'intérieur et entre les différents régimes, alors les erreurs entre les deux sous-régimes ne sont plus indépendantes : $E[u_{h_1} u_{h_2}] = \sigma_{12}$. Dans ce cas, la variance totale est de la forme :

$$\Psi = \begin{bmatrix} B_1^{-1} & 0 \\ 0 & B_2^{-1} \end{bmatrix} \cdot \begin{bmatrix} \sigma_1^2 I_1 & \sigma_{12} S S' \\ \sigma_{12} S S' & \sigma_2^2 I_2 \end{bmatrix} \cdot \begin{bmatrix} (B_1')^{-1} & 0 \\ 0 & (B_2')^{-1} \end{bmatrix} \quad [2.154]$$

où S est le vecteur somme d'ordre N .

233.42 *Le test*

Lorsque les erreurs ne sont pas homoscédastiques : $\Psi = \sigma^2 \Omega$, l'expression du test d'instabilité structurelle est :

$$C_G = \frac{\hat{e}'_c \hat{\Omega}^{-1} \hat{e}_c - \hat{e}'_L \hat{\Omega}^{-1} \hat{e}_L}{\hat{\sigma}^2} \stackrel{\text{as}}{\sim} \chi^2(K) \quad [2.155]$$

où \hat{e}_c est le vecteur des erreurs estimées du modèle contraint et \hat{e}_L celui du modèle libre. Cette statistique suit asymptotiquement une loi du chi-deux à K degrés de liberté, K étant le nombre de paramètres du modèle. De la même manière, on peut effectuer ce test sur chacun des coefficients du modèle.

Ainsi, pour toutes les formes de modèles présentés dans le paragraphe précédent, le test de l'instabilité structurelle en présence d'autocorrélation spatiale peut être construit de cette façon, avec des estimations convergentes des paramètres du modèle. Par exemple, le test de Chow spatial pour le modèle avec autocorrélation spatiale [2.147] et erreurs homoscédastiques [2.148] est de la forme :

$$C_s = \frac{\hat{e}'_c (I - \hat{\lambda}W)' (I - \hat{\lambda}W) \hat{e}_c - \hat{e}'_L (I - \hat{\lambda}W)' (I - \hat{\lambda}W) \hat{e}_L}{\hat{\sigma}^2} \stackrel{\text{as}}{\sim} \chi^2(K) \quad [2.156]$$

Comme mentionné précédemment, les simulations de Monte-Carlo effectuées par Anselin (1990a) indiquent qu'ignorer la présence de l'autocorrélation spatiale affecte le seuil et la puissance du test de Chow traditionnel. Il apparaît donc nécessaire de tester la présence d'une autocorrélation spatiale des erreurs avant d'effectuer un test d'instabilité structurelle et de la prendre en compte si elle est présente. En effet, la présence d'autocorrélation spatiale dans les analyses de régression invalide les résultats des tests de stabilité structurelle. On pourrait, par exemple, conclure que deux régions appartenant à des régimes différents répondent différemment à une politique économique alors que ce n'est pas le cas.

En pratique, Anselin (1990a) suggère de commencer par effectuer un test destiné à détecter la présence d'autocorrélation spatiale. Différents tests et différentes règles de décision sont proposés permettant de déterminer la forme prise par l'autocorrélation spatiale : autocorrélation des erreurs ou variable endogène décalée (Anselin *et al.*, 1996). S'il existe une forte indication d'autocorrélation spatiale, et plus particulièrement pour des valeurs positives et des régimes correspondant à des observations compactes et contiguës, les techniques « standards » ne sont pas fiables et d'autres méthodes d'estimation doivent être mobilisées (méthode du maximum de vraisemblance, méthode des variables instrumentales ou des moments généralisés). Cette approche empirique peut cependant poser des problèmes de

pretest, qui invalident la distribution asymptotique usuelle des différents tests. Dans ces conditions, il est prudent de choisir des niveaux de significativité restrictifs pour les tests.

Conclusion de section

Dans cette section, nous avons présenté la façon dont l'hétérogénéité spatiale était introduite dans les modèles économétriques et quelles étaient les conséquences de la présence conjointe de l'autocorrélation et de l'hétérogénéité spatiales.

L'hétérogénéité spatiale revêt deux aspects. D'une part, dans le cas d'une instabilité des paramètres, les coefficients de la régression ne sont plus constants sur l'espace où se produisent les phénomènes économiques étudiés. La variation spatiale des paramètres d'une régression peut être discrète ou continue. D'autre part, dans le cas d'une hétéroscédasticité, la variance des termes d'erreurs varie avec chaque observation localisée, cette structure spatiale pouvant être de forme connue ou inconnue.

La présence simultanée d'autocorrélation et d'hétérogénéité spatiale est un phénomène complexe mais qui semble relativement fréquent. Il nécessite des méthodes d'estimation et d'inférence adaptées. Nous avons proposé une synthèse de ces méthodes d'estimation et de tests et nous avons présenté, lorsque cela était possible, la manière de conduire les études empiriques dans les différentes configurations possibles. Cependant, des règles de décision permettant de choisir entre les différents modèles combinant les deux effets spatiaux restent encore à définir.

Conclusion du chapitre 2

L'objectif de ce chapitre était d'examiner pourquoi l'autocorrélation et l'hétérogénéité spatiales doivent être modélisées et comment elles sont introduites dans les modèles économétriques.

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

L'hétérogénéité spatiale a deux facettes. Dans le cas d'une instabilité des paramètres, les coefficients de la régression ne sont plus constants dans l'espace. La variation géographique des paramètres d'une régression peut être discrète, auquel cas il faut spécifier des régimes spatiaux, ou continue, les paramètres sont alors supposés être une fonction des coordonnées de chaque localisation. En présence d'une hétéroscédasticité, la variance des erreurs varie avec chaque observation localisée. Deux stratégies différentes peuvent alors être adoptées. La première est l'approche de White (1980) dans laquelle on utilise pour l'inférence statistique une estimation convergente de la matrice des variances-covariances basée sur les MCO en présence d'une hétéroscédasticité de forme inconnue. La seconde est de donner à l'hétéroscédasticité une forme spécifique, telle qu'une hétéroscédasticité en blocs. Lorsque cette spécification est adéquate, les tests associés sont plus puissants.

La présence simultanée d'autocorrélation spatiale et d'hétérogénéité spatiale dans un modèle est très fréquente compte tenu des nombreux liens qui unissent ces deux effets. Ce cas de figure nécessite des méthodes d'estimation et d'inférence spécifiques. Concernant l'instabilité structurelle, le test de Chow traditionnel doit être modifié pour tenir compte d'une autocorrélation spatiale dans les erreurs. Concernant l'hétéroscédasticité, deux aspects sont à souligner. D'une part, s'il est difficile de spécifier la forme de l'hétéroscédasticité, il faut alors utiliser les tests d'autocorrélation spatiale robustes à une forme inconnue d'hétéroscédasticité. D'autre part, si l'on utilise une forme particulière d'hétéroscédasticité, il est possible de tester conjointement ou séparément les deux effets spatiaux. Le test de Breusch-Pagan (1979) doit également être amendé pour tenir compte d'une éventuelle autocorrélation spatiale des erreurs.

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

Longtemps ignorées dans les articles contenant des études empiriques, les techniques de l'économétrie spatiale sont aujourd'hui de plus en plus appliquées. Elles permettent d'améliorer les estimations produites par les MCO, de détecter la présence des effets de débordement géographiques sur les comportements économiques et d'évaluer l'impact de ces effets à travers l'estimation des paramètres spatiaux.

Les études empiriques conduites dans les chapitres suivants s'appuient sur ces différentes techniques. En effet, notre terrain d'application est l'analyse des disparités économiques de PIB par tête entre les régions européennes et la façon dont la localisation de chaque région au sein de l'espace européen influence ces disparités. En optant pour l'étude de la convergence des régions européennes, nous nous plaçons alors dans une perspective de renouvellement de ce débat en mettant l'accent sur l'influence de l'espace dans la réduction des disparités économiques. A cet effet, nous suivons la méthodologie présentée dans ce chapitre qui mène de la détection de la structure spatiale des données en termes d'autocorrélation et/ou d'hétérogénéité spatiales à l'estimation des modèles incorporant ces deux effets.

Annexe 2.1

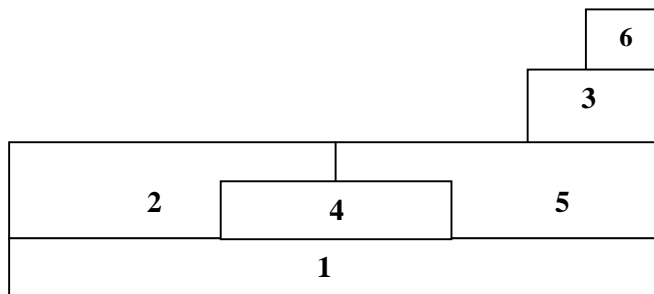
Algorithmes de construction d'une matrice générale des distances à partir d'une matrice de contiguïté d'ordre 1

Le but de cette annexe est de présenter les algorithmes permettant de passer d'une matrice binaire de contiguïté d'ordre 1 à la matrice générale des distances où chaque élément (i,j) représente le nombre de frontières qu'il faut traverser pour aller de la région i à la région j . L'existence de ces algorithmes se justifie par le fait que la matrice de contiguïté d'ordre s n'est pas égale à la matrice de contiguïté d'ordre 1 élevée à la puissance s , comme nous le verrons dans le premier point. Ensuite, nous présentons les différents algorithmes dont celui de Blommestein et Koper (1992) qui comporte des imperfections et qui a été amendé par Anselin et Smirnov (1996). Ces derniers introduisent d'autres algorithmes qui convergent plus rapidement vers la solution.

1. Les matrices de contiguïté d'ordre supérieur à 1

On dit que deux régions i et j sont contiguës à l'ordre s à la condition de ne pas être déjà contiguës à un ordre inférieur à s . En d'autres termes, deux régions i et j sont contiguës à l'ordre s si s est le nombre *minimal* de frontières à traverser pour aller de i à j .

Considérons la figure suivante (Anselin et Smirnov, 1996, p. 71) :



Il existe différents problèmes lorsqu'on élève la matrice de contiguïté à une puissance quelconque. Cette opération ne permet pas d'aboutir à une matrice de contiguïté d'ordre supérieur à un.

En effet, la matrice W de contiguïté d'ordre 1 associée à la figure précédente et la matrice W^2 sont :

$$W = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{et :} \quad W^2 = \begin{bmatrix} 3 & 2 & 1 & 2 & 2 & 0 \\ 2 & 3 & 1 & 2 & 2 & 0 \\ 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 2 & 1 & 3 & 2 & 0 \\ 2 & 2 & 0 & 2 & 4 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

La matrice W^2 n'est pas une matrice de contiguïté adéquate.

Tout d'abord, les éléments de la diagonale principale ne sont pas nuls. Evidemment, cela peut être facilement corrigé en mettant ces éléments à 0. En outre, les éléments hors diagonale principale donnent le nombre de chemins de longueur 2 entre chaque paire de régions. Par exemple, il y a deux chemins de longueur 2 qui lient la région 1 et la région 5. Ces chemins ne sont pas donnés par la matrice mais nous les voyons sur la figure : il s'agit des chemins 1 – 4 – 5 et 1 – 2 – 5 (cette information n'est cependant pas nécessaire puisque seule l'existence de tels chemins est importante et non leur nombre).

Ensuite, cette matrice contient des sentiers redondants (*redundant paths*), c'est-à-dire des chemins de longueur 2 concernant des régions qui sont déjà contiguës à l'ordre 1. Ainsi, par exemple, pour avoir une matrice de contiguïté adéquate, il faudrait éliminer les chemins 1 – 4 – 5 et 1 – 2 – 5 puisque les régions 1 et 5 sont déjà contiguës à l'ordre 1.

Enfin, les matrices élevées à une puissance quelconque sont également susceptibles de contenir des routes circulaires (*circular routes*), c'est-à-dire des chemins dans lesquels au moins un indice de région (par exemple, le nombre avec lequel la région est désignée) est répété dans une séquence d'indices de régions (Blommestein et Koper, 1992). Par exemple, la région 3 peut être reliée à la région 5 par le chemin suivant : 3 – 5 – 2 – 5. Cette séquence est une des 5 séquences qui se trouve dans la case (5,3) (ou la case (3,5)) de la matrice de contiguïté d'ordre 1 élevée à la puissance 3 :

$$W^3 = \begin{bmatrix} 6 & 7 & 2 & 7 & 8 & 1 \\ 7 & 6 & 2 & 7 & 8 & 1 \\ 2 & 2 & 0 & 2 & 5 & 2 \\ 7 & 7 & 2 & 6 & 8 & 1 \\ 8 & 8 & 5 & 8 & 6 & 0 \\ 1 & 1 & 2 & 1 & 0 & 0 \end{bmatrix}$$

Ces chemins doivent aussi être éliminés pour obtenir des matrices de contiguïté adéquates (Blommestein et Koper, 1997).

Finalement, l'élévation de la matrice de contiguïté d'ordre 1 à une puissance quelconque ne donne pas des matrices de contiguïté adéquates. En effet, cette opération aboutit à des sentiers redondants et des chemins circulaires qui doivent être éliminés. Dans l'exemple précédent, il faut donc trouver des moyens qui permettent de passer de la matrice de contiguïté d'ordre 1 à la matrice générale des distances suivante :

$$S = \begin{bmatrix} 0 & 1 & 2 & 1 & 1 & 3 \\ 1 & 0 & 2 & 1 & 1 & 3 \\ 2 & 2 & 0 & 2 & 1 & 1 \\ 1 & 1 & 2 & 0 & 1 & 3 \\ 1 & 1 & 1 & 1 & 0 & 2 \\ 3 & 3 & 1 & 3 & 2 & 0 \end{bmatrix}$$

L'ordre de contiguïté le plus élevé est égal à 3. Le passage de cette matrice des distances aux différentes matrices de contiguïté est alors immédiat. Par exemple, si l'on veut construire la matrice correspondant à l'ordre 2, il suffit de remplacer dans la matrice le nombre 2 par 1 et les nombres 1 et 3 par 0.

2. Les algorithmes

Nous verrons tout d'abord l'algorithme proposé par Blommestein (1985) et repris par Blommestein et Koper (1992). Il permet l'élimination des routes circulaires mais il ne détecte pas les sentiers redondants. Il est donc corrigé par Anselin et Smirnov (1996), ces deux auteurs ayant également proposé trois autres algorithmes plus efficaces.

2.1. L'algorithme de Blommestein et Koper

L'algorithme de Blommestein et Koper (qu'on notera « BK ») est basé sur le calcul des puissances successives de la matrice de contiguïté d'ordre 1, puissances dans lesquelles les sentiers redondants sont éliminés en imposant des contraintes sur les indices.

La formule récursive pour la contiguïté à l'ordre p est (Blommestein et Koper, p. 97):

$$l_{ij}^p = \sum_{k_{p-1}=1}^n (w_{ik_{p-1}} \cdot l_{k_{p-1}j}^{p-1})$$

Dans la formulation précédente, l_{ij}^p sont les éléments du décalage spatial « corrigé » à l'ordre p et w_{ij} sont les éléments de la matrice de contiguïté d'ordre 1. Le sous-indice $(p - 1)$ se réfère à l'étape précédente dans l'itération.

Pour éliminer les sentiers redondants, les conditions suivantes sont nécessaires :

$$i \neq j$$

$$i \neq k_m \quad m = 1, \dots, p-1$$

$$j \neq k_m \quad m = 1, \dots, p-1$$

$$k_n \neq k_m \quad m = 1, \dots, p-1 \text{ et } n = 1, \dots, p-1, m \neq n$$

Cet algorithme élimine les routes circulaires mais pas les sentiers redondants. C'est pour cela qu'Anselin et Smirnov amendent l'algorithme BK en considérant une fonction $S(i, j)$ définie pour toutes les paires i et j , et qui est égale à l'ordre de contiguïté entre ces régions. L'essence de l'algorithme BK peut être vue comme la solution du problème suivant, pour tout i et tout ordre de contiguïté p :

$$S(i, j) = p \Leftrightarrow \exists k \text{ tel que } S(i, k) = p-1, S(k, j) = 1 \text{ et } k \neq i, j$$

L'algorithme a une structure récursive puisqu'il découle immédiatement de l'équation précédente :

$$S(i, k) = p-1 \Leftrightarrow \exists t \text{ tel que } S(i, t) = p-2, S(t, k) = 1 \text{ et } t \neq i, j, k$$

Et ainsi de suite jusqu'à ce que :

$$S(i, u) = 2 \Leftrightarrow \exists v \text{ tel que } S(i, v) = 1, S(v, u) = 1 \text{ et } v \neq i, j, k, t, \dots, u$$

La première partie de ces trois équations est simplement la définition de la contiguïté alors que la deuxième partie est la condition nécessaire à l'élimination des routes circulaires. Pour corriger l'algorithme et lui permettre d'éliminer les sentiers redondants, il suffit d'ajouter à ces trois équations la condition de vérification de la présence d'un ordre de contiguïté inférieur. Vérifier alors l'existence d'une contiguïté d'ordre inférieur élimine automatiquement également les routes circulaires. L'algorithme BK corrigé devient donc :

$$S(i, j) = p \Leftrightarrow S(i, j) \neq q < p \text{ et } \exists k \text{ tel que } S(i, k) = p-1, S(k, j) = 1 \text{ et } k \neq i, j$$

soit : $S(i, k) = p-1 \Leftrightarrow S(i, k) \neq q < p \text{ et } \exists t \text{ tel que } S(i, t) = p-2, S(t, k) = 1 \text{ et } t \neq i, j, k$

Et ainsi de suite jusqu'à ce que :

$$S(i, u) = 2 \Leftrightarrow S(i, u) \neq 1 \text{ et } \exists v \text{ tel que } S(i, v) = 1, S(v, u) = 1 \text{ et } v \neq i, j, k, t, \dots, u$$

22. Deuxième algorithme

Le principe de cet algorithme est très simple puisqu'il consiste à calculer les puissances successives de la matrice de contiguïté d'ordre 1 et à les comparer à une matrice booléenne qui contient des informations sur les contiguïtés d'ordre inférieur. Seuls les éléments de la première matrice différents de ceux de la seconde sont conservés pour aboutir à une matrice de contiguïté adéquate.

La première étape de l'algorithme consiste à poser une matrice « d'accumulation » $A = W$ où W est la matrice de contiguïté d'ordre 1. On calcule ensuite $W^2 = W.W$. La matrice de contiguïté d'ordre 2, notée L^2 , est obtenue en ne retenant que les éléments positifs d'une matrice auxiliaire E^2 , cette dernière étant la différence entre la forme booléenne de W^2 , notée B^2 , et A : $E^2 = B^2 - A$. Comme mentionné précédemment, les éléments de B^2 sont tels que $b_{ij}^2 = 1$ si $w_{ij}^2 > 0$, $b_{ij}^2 = 0$ si $w_{ij}^2 = 0$ (et $b_{ij}^2 = 0$). Dans l'exemple :

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 0 \end{bmatrix}$$

Dans l'équation précédente, la première matrice est la forme booléenne de W^2 , la deuxième matrice est W . Le résultat est une matrice dont l'élément (i,j) est égal à 1 si i et j sont contiguës à l'ordre 2 et -1 si i et j sont contiguës à l'ordre 1 mais pas à l'ordre 2. On en déduit les éléments de L^2 qui sont donc $l_{ij}^2 = 1$ si $e_{ij}^2 > 0$, $l_{ij}^2 = 0$ sinon :

$$L^2 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

La nouvelle matrice d'accumulation est alors actualisée : $A = A + L^2$ et la procédure est poursuivie pour l'ordre suivant.

23. Troisième algorithme

Cet algorithme utilise l'information contenue dans la matrice W et construit pour chaque région i (c'est-à-dire pour chaque ligne i de W) un vecteur auxiliaire s , dont chaque élément s_j ($j = 1, \dots, n$) indique l'ordre de contiguïté entre i et j . Ainsi, pour un vecteur s associé à une région i , le décalage spatial à l'ordre p est :

$$l_{ij}^p = 1 \text{ pour } s_j = p$$

$$l_{ij}^p = 0 \text{ sinon.}$$

Formellement, la première étape de l'algorithme consiste à initialiser les éléments du vecteur s pour la ligne i de la façon suivante :

$$s_j = w_{ij}, \forall j \neq i$$

et $s_i = -1$

La valeur -1 est arbitraire, il suffit simplement que cette valeur soit différente d'un des ordres de contiguïté utilisé.

Pour chaque ordre de contiguïté c (avec $c > 1$) et pour chaque élément j du vecteur s étudié (sauf pour $i = j$), on liste les éléments de s , notés s_k , pour lesquels l'ordre de contiguïté est égal à $c - 1$. On vérifie ensuite la présence d'une contiguïté d'ordre 1 entre la région j et la région k dans la ligne correspondante de la matrice W . Si pour la région j , on a $s_j = 0$, l'ordre de contiguïté à i est actualisé à c :

$$s_j \leftarrow c \Leftrightarrow \exists k, j : s_k = c - 1, w_{kj} = 1 \text{ et } s_j = 0$$

Par exemple, nous allons effectuer cet algorithme pour la première région. Le vecteur s pour la première ligne est initialisé :

$$s = [-1 \ 1 \ 0 \ 1 \ 1 \ 0]$$

Pour construire l'opérateur spatial du second ordre, les éléments des lignes $k = 2, 4$ et 5 de W sont retenus puisque s_2, s_4, s_5 sont différents de zéro. Pour $k = 2$, $w_{21} = w_{24} = w_{25} = 1$ mais ces éléments ne donnent pas lieu à une contiguïté d'ordre 2 puisque s_1, s_4, s_5 sont différents de zéro. De même pour $k = 4$. En revanche, pour $k = 5$, $w_{53} = 1$ et $s_3 = 0$. Par conséquent, s_3 est actualisé à une nouvelle valeur 2.

Concernant l'opérateur du troisième ordre, seul $k = 3$ est susceptible de conduire à un chemin de longueur 3 (puisque $s_3 = 2$). Dans la troisième ligne de W , seul $w_{36} = 1$. Comme $s_3 = 0$, on peut actualiser s_6 : $s_6 = 3$. L'opération se termine ici pour la première région

puisque'il ne reste plus d'élément nul dans s . Pour la première région, le vecteur s final est donc :

$$s = [-1 \ 1 \ 2 \ 1 \ 1 \ 3]$$

On vérifie bien que la première région est contiguë à l'ordre 1 avec les régions 2, 4 et 5, à l'ordre 2 avec la région 3 et à l'ordre 3 avec la région 6.

24. Quatrième algorithme

Cet algorithme est très similaire à l'algorithme précédent mais il utilise plus efficacement l'information liée à la structure des données en évitant la recherche de tous les voisins potentiels (c'est-à-dire les éléments du vecteur s du troisième algorithme) et en permettant à l'algorithme de se mouvoir directement vers les régions pertinentes. Pour chaque région i , cet algorithme s'effectue en plusieurs étapes.

Première étape

On regarde sur la $i^{\text{ème}}$ ligne de la matrice de poids quelles sont les régions contiguës à l'ordre 1 avec la région i . Les indices de ces régions sont consignés dans un vecteur Q_1 .

Dans l'exemple, on voit sur la matrice que la première région est contiguë avec les régions 2, 4 et 5. Par conséquent, le vecteur Q_1 est : $Q_1 = [2 \ 4 \ 5]$.

Deuxième étape

Pour chaque région du vecteur Q_1 , on regarde si elle est contiguë à l'ordre 1 avec une région qui n'est pas déjà consignée dans le vecteur Q_1 . En effet, si la région i est contiguë à l'ordre 1 avec une région k du vecteur Q_1 , région qui est elle-même contiguë à l'ordre 1 avec une autre région l (qui n'appartient pas à Q_1), cela signifie que les régions i et l sont contiguës à l'ordre 2. Ces régions contiguës à l'ordre 2 sont consignées dans un vecteur Q_2 .

Dans l'exemple, la première région de Q_1 est la région 2 qui est contiguë à l'ordre 1 avec les régions 1, 4 et 5 (on lit ces régions en regardant la deuxième ligne de la matrice de poids). Toutes ces régions appartiennent déjà au vecteur Q_1 (puisque 1 est la région étudiée, on ne la prend pas en compte). On peut donc passer à l'élément suivant du vecteur Q_1 . Pour la région 4, on procède de la même manière et de nouveau, aucun élément n'est trouvé. Enfin, la région 5 est contiguë à l'ordre 1 avec les régions 1, 2, 3, et 4. La région 3 n'est pas un élément de Q_1 et est donc contiguë à l'ordre 2 avec la région 1. Le vecteur Q_2 est le suivant : $Q_2 = [3]$.

Troisième étape

On répète la deuxième étape en prenant Q_2 au lieu de Q_1 .

Dans l'exemple, la région 3 est contiguë à l'ordre 1 avec les régions 5 et 6 (on le voit sur la troisième ligne de la matrice). La région 5 est déjà contiguë à l'ordre 1 avec la première région puisque 5 est un élément de Q_1 . En revanche, la région 6 n'appartient ni à Q_2 ni à Q_1 , elle est donc contiguë à l'ordre 3 avec la région 1 et : $Q_3 = [6]$.

Si cet ordre est l'ordre désiré, l'algorithme passe à la région suivante. Sinon, il faut continuer de la même façon jusqu'à l'ordre de contiguïté désiré. Dans l'exemple, la région 6 est contiguë à l'ordre 1 avec la région 3 qui est déjà contiguë à l'ordre 2 avec la région 1 (puisque 3 est le seul élément de Q_2). On retrouve bien le fait que la première région est contiguë à l'ordre 1 avec les régions 2, 4 et 5, à l'ordre 2 avec la région 3 et à l'ordre 3 avec la région 6.

Les essais effectués par Anselin et Smirnov sur les 48 états contigus des Etats-Unis, les 921 *Kreise* allemands et les 3111 comtés US montrent que le quatrième algorithme est le plus rapide dans tous les cas, suivi par le troisième, le second et enfin le premier. Ce dernier est de plus particulièrement peu performant puisque ses temps de calcul augmentent exponentiellement avec l'ordre de contiguïté. Le deuxième algorithme n'est pas non plus performant pour les ordres de contiguïté élevés.

Annexe 2.2

Estimation par le maximum de vraisemblance : étude de deux cas particuliers

Le but de cette annexe est de détailler la procédure d'estimation pour le modèle autorégressif et le modèle à erreurs autocorrélées. Dans ces deux cas, en effet, l'estimation peut se réaliser à partir de la fonction de log-vraisemblance concentrée. Le principe de cette méthode est de résoudre une partie des équations associées aux conditions du premier ordre et d'introduire ensuite les solutions obtenues dans la fonction de log-vraisemblance. On obtient alors une fonction de log-vraisemblance concentrée qui est une fonction simplifiée ne dépendant plus que du paramètre spatial. Dès lors, on peut trouver des estimations de ce paramètre par balayage, c'est-à-dire en évaluant la fonction de log-vraisemblance pour un petit intervalle des valeurs de ces paramètres.

1. Le modèle spatial autorégressif

Dans le cas du modèle autorégressif (équation [2.23]), la fonction de log-vraisemblance est construite en appliquant à ce cas particulier la formule générale (donnée par l'équation [2.27]) :

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

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

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

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

Conditionnellement à ρ , ces deux expressions s'interprètent comme les estimateurs des MCO appliqués au modèle filtré (équation [2.24]). La substitution de ces estimateurs dans la fonction de log-vraisemblance aboutit à la fonction de log-vraisemblance concentrée qui ne dépend plus alors que du paramètre spatial ρ :

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

où $C = -(N/2) \ln(2\pi) - (N/2)$ et où $\hat{\varepsilon}_0$ et $\hat{\varepsilon}_L$ sont respectivement les résidus estimés des régressions de y sur X et de Wy sur X .

La procédure d'estimation s'effectue alors sur la base des étapes suivantes :

1/ Dans un premier temps, on estime par les MCO y sur X et Wy sur X et on calcule les résidus estimés, notés respectivement $\hat{\varepsilon}_0$ et $\hat{\varepsilon}_L$.

2/ Compte tenu de $\hat{\varepsilon}_0$ et $\hat{\varepsilon}_L$, on trouve ρ qui maximise la fonction de log-vraisemblance concentrée $\ln L_c$.

3/ Compte tenu de $\hat{\rho}_{ML}$, les estimations de β et de σ^2 sont calculées grâce aux formules obtenues à partir des conditions de premier ordre.

Finalement, en notant $W_A = W(I - \rho W)^{-1}$, la matrice des variances-covariances asymptotique est donnée par l'inverse de la matrice d'information :

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

La covariance entre β et la variance est nulle comme dans le modèle standard, mais la covariance entre ρ et les deux autres paramètres n'est pas nulle. Cette matrice n'est donc pas bloc-diagonale.

2. Le modèle avec autocorrélation des erreurs

Dans le cas de l'autocorrélation spatiale des erreurs (équation [2.30]), la fonction de log-vraisemblance prend la forme suivante :

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

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

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

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

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

Si λ est connu, l'estimateur du maximum de vraisemblance est équivalent à celui des MCO appliqué au modèle avec variables filtrées (équation [2.32]).

Si l'on considère la forme moyenne mobile de l'autocorrélation spatiale, l'expression de l'estimateur des MCG (équation [2.36]) nécessite l'inversion d'une matrice des variances-covariances de dimension (N, N) : $\Omega(\lambda)^{-1} = [I + \gamma(W + W') + \gamma^2 WW']^{-1}$. Cette complication a sans doute contribué à limiter l'application pratique de ce modèle.

La substitution de ces estimateurs dans la fonction de log-vraisemblance aboutit à la fonction de log-vraisemblance concentrée suivante :

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

Une procédure itérative est nécessaire pour l'estimation des paramètres dans le cas du modèle avec autocorrélation spatiale des erreurs :

- 1/ Dans un premier temps, on estime y sur X on calcule les résidus estimés, notés $\hat{\varepsilon}_{MCO}$.
- 2/ Compte tenu de $\hat{\varepsilon}_{MCO}$, on trouve λ qui maximise la fonction de log-vraisemblance concentrée $\ln L_c$.
- 3/ Compte tenu de $\hat{\lambda}$, obtenu à l'étape 2, on calcule $\hat{\beta}_{MCQG}$ et on en déduit les nouveaux résidus estimés $\hat{\varepsilon}_{MCQG} = y - X \hat{\beta}_{MCQG}$.
- 4/ Si le critère de convergence est atteint, on passe à l'étape suivante, sinon on retourne à la seconde étape.
- 5/ Compte tenu de $\hat{\varepsilon}_{MCQG}$ et de $\hat{\lambda}$, on calcule la variance de la façon suivante :

$$\sigma_{ML}^2 = (1/N) \hat{\varepsilon}_{MCQG}' (I - \hat{\lambda} W)' (I - \hat{\lambda} W) \hat{\varepsilon}_{MCQG}$$

Finalement, en notant $W_B = W(I - \lambda W)^{-1}$, la matrice des variances-covariances asymptotique pour les estimateurs est similaire à la forme générale de Magnus (1978) et Breusch (1980) et est bloc-diagonale entre les paramètres de la régression β et les paramètres de variance des erreurs σ^2 et λ :

$$\text{Asy}V(\beta, \lambda, \sigma^2) = \begin{bmatrix} \frac{X'(I - \lambda W)'(I - \lambda W)X}{\sigma^2} & 0 & 0 \\ 0 & \text{tr}(W_B)^2 + \text{tr}(W_B'W_B) & \frac{\text{tr}(W_B)}{\sigma^2} \\ 0 & \frac{\text{tr}(W_B)}{\sigma^2} & \frac{N}{2\sigma^4} \end{bmatrix}^{-1}$$

La propriété de bloc-diagonalité de la matrice des variances-covariances est intéressante dans le cas de très grands échantillons, pour lesquels le calcul de la matrice W_B d'ordre N , peut se révéler malaisé. En effet, dans ce cas, même si le calcul du ratio des coefficients à leurs écarts-types est impossible, l'inférence asymptotique reste réalisable. La bloc-diagonalité implique que les statistiques asymptotiques pour $\hat{\beta}$ peuvent être calculées sans que la précision du paramètre autorégressif λ soit connue (l'inférence du paramètre autorégressif doit être basée sur le test du ratio de vraisemblance).

CHAPITRE 3

Analyse exploratoire des disparités régionales de PIB par tête en Europe

Une partie des résultats empiriques présentés dans ce chapitre est tirée des articles suivants :

Ertur C., Le Gallo J. (2002) An exploratory spatial data analysis of European regional disparities, 1980-1995, dans : Fingleton B. (ed.), *European Regional Growth*, Springer, Berlin, à paraître.

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

Chapitre 3

Analyse exploratoire des disparités régionales de PIB par tête en Europe

Dans un contexte d'intégration du marché européen, qui se consolide au cours des années 80 par l'entrée au sein de l'Union Européenne de l'Espagne, du Portugal et de la Grèce, de nombreux auteurs se sont intéressés à la répartition géographique de la richesse au sein des régions européennes (Armstrong, 1995b ; Neven et Gouyette, 1995 ; Martin, 2001). En effet, cette analyse est intéressante à plusieurs titres.

D'une part, elle sert à souligner l'importance des liens existant entre disparités économiques et localisation géographique. En effet, elle est révélatrice de la persistance des schémas inégalitaires de répartition de la richesse au sein de l'Europe, où s'illustre de façon marquante un schéma de type centre-périphérie décrit par les théories de l'économie géographique. Pourtant, la diminution des coûts d'interaction, suite à la baisse des coûts de transport et à la consolidation du processus d'intégration régionale, mais aussi le développement rapide des technologies de la communication et les restructurations économiques de la dernière décennie, auraient pu avoir pour effet de diminuer plus clairement l'emprise de conditions nationales et locales dans la croissance économique et de réduire davantage l'isolement des régions périphériques (Rodríguez-Posé, 1998).

D'autre part, elle permet d'évaluer les liens existant entre ce schéma de disparités et le dynamisme économique des régions pauvres de l'Europe communautaire impulsé notamment par les politiques de développement structurel et de cohésion, depuis la réforme des fonds structurels en 1988 et sa mise en œuvre en 1989. Nous précisons que notre objectif n'est pas d'évaluer ces politiques, mais simplement de garder à l'esprit que sur notre période d'étude, les régions européennes éligibles à l'objectif 1 de la réforme des fonds structurels étaient globalement en retard de développement par rapport aux autres régions et qu'elles ont bénéficié à ce titre d'importantes mesures visant à dynamiser leur économie.

Dans un tel contexte, notre analyse reste donc essentiellement empirique et se focalise sur les liens entre la localisation géographique des régions au sein de l'espace européen et leur situation économique, que nous apprécions par les PIB par tête. Plus précisément, notre objectif est d'analyser la dynamique spatio-temporelle de la distribution des PIB par tête en Europe sur la période 1980-1995 pour détecter la forme de l'autocorrélation et de l'hétérogénéité spatiales au sein de notre échantillon. En effet, dans ce contexte particulier, l'autocorrélation spatiale positive se caractériserait par la concentration spatiale de régions

possédant des PIB par tête régionaux similaires et conduirait à distinguer les concentrations de régions « pauvres » des concentrations de régions « riches ». Une hétérogénéité spatiale semblerait alors la plus adaptée pour décrire les schémas de développement économique propres à chaque concentration régionale. Ceci nous amènerait alors à raisonner en termes de régimes spatiaux où le centre (c'est-à-dire une concentration de régions riches) se distinguerait de la périphérie (c'est-à-dire une concentration de régions pauvres).

Nous nous appuyons pour cela sur les méthodes récemment développées de l'analyse exploratoire des données spatiales (Anselin, 1995, 1996).

Nous présentons d'abord la base de données Eurostat-Regio, notre échantillon de 138 régions européennes sur la période 1980-1995 et les matrices de poids, exprimant la structure spatiale des données, que nous utilisons dans ce chapitre et les suivants (section 31).

Nous nous intéressons ensuite aux disparités en termes de niveaux de PIB par tête (section 32). Nous montrons déjà que cette distribution est effectivement caractérisée par une forte autocorrélation spatiale positive. Les statistiques locales de Getis-Ord (Ord et Getis, 1995) sont calculées afin de détecter les concentrations de PIB par tête élevés ou faibles. En effet, l'existence de telles concentrations durant la période est une indication de la persistance des disparités spatiales entre les régions européennes. Les statistiques locales de Moran (diagramme de Moran et LISA : Anselin, 1995, 1996) sont également calculées et révèlent l'existence de quelques régions atypiques, c'est-à-dire des régions qui se démarquent des schémas de concentrations dominants. Ces régions atypiques sont cependant peu nombreuses. Nous proposons enfin une méthode originale permettant d'étudier la robustesse de nos résultats au choix des matrices de poids.

Enfin, nous examinons les disparités spatiales sur la croissance régionale (section 33). Ceci nous conduit à étudier le dynamisme des régions européennes en relation avec leur niveau de PIB par tête initial et à montrer que les deux effets spatiaux - autocorrélation et hétérogénéité – sont potentiellement présents dans cette relation. Les implications principales de ces résultats sur l'estimation et les tests de β -convergence absolue et conditionnelle sont soulignées dans la conclusion ⁴⁸.

⁴⁸ Tous les résultats des tests donnés dans ce chapitre sont effectués à l'aide du logiciel SpaceStat 1.90 (Anselin, 1999). Les cartes et les figures ont été réalisées à l'aide d'Arcview 3.2 (Esri) et de Statistica (StatSoft).

SECTION 31 :

LES DONNEES ET LA MATRICE DE POIDS

Dans l'étude des disparités économiques en termes de PIB par tête entre les régions européennes, la banque de données Eurostat-Regio a souvent été utilisée (Neven et Gouyette, 1995 ; Quah, 1996b ; López-Bazo *et al.*, 1999).

Pour l'ensemble des études empiriques réalisées dans la thèse, nous avons également travaillé sur un échantillon spécifique de régions européennes, pour lesquelles nous disposons de données sur les PIB par tête extraites de la base de données Eurostat-Regio. L'objectif de cette section est d'abord de rappeler quelques précautions méthodologiques concernant l'utilisation de cette base de données (paragraphe 311) et ensuite de présenter l'échantillon utilisé (paragraphe 312). Compte tenu des spécificités de l'échantillon, nous exposons enfin les différentes matrices de poids qui serviront pour les études empiriques présentées dans ce chapitre et les suivants (paragraphe 313).

311 La base de données Eurostat-Regio

Eurostat est l'office statistique des Communautés européennes. Son but est de fournir à l'Union Européenne des statistiques au niveau européen qui permettent les comparaisons entre les pays et les régions. Ces statistiques sont utilisées par la Commission Européenne et les autres institutions européennes pour qu'elles puissent définir, mettre en œuvre et analyser les politiques européennes. Eurostat-Regio est la source officielle de données annuelles harmonisées au niveau régional pour l'Union Européenne.

L'utilisation d'Eurostat-Regio pour l'étude des disparités de PIB par tête requiert quelques précautions méthodologiques concernant le découpage régional adopté par Eurostat (paragraphe 311.1), ainsi que les procédures de correction des parités de pouvoir d'achat (paragraphe 311.2).

311.1 La Nomenclature des Unités Territoriales Statistiques

Le découpage régional adopté par Eurostat est la Nomenclature des Unités Territoriales Statistiques (« NUTS »). Cette nomenclature a été établie dans le but de fournir une ventilation unique et uniforme des unités territoriales pour la production de statistiques régionales et communautaires. En outre, même si cette classification n'a pas de valeur légale en soi, elle constitue depuis 1988 la base de l'identification des régions pouvant bénéficier d'une aide de la Communauté.

Les NUTS ont été créés suivant trois principes (Eurostat, 1999). Premièrement, les NUTS favorisent les découpages institutionnels ou « régions normatives » qui sont l'expression de la volonté politique. Ainsi, leurs limites sont fixées suivant les tâches allouées aux communautés territoriales, suivant la taille de la population nécessaire pour mettre en œuvre ces tâches d'une façon efficiente et économique ou suivant d'autres facteurs historiques ou culturels. Les régions normatives ont pour avantage d'être le plus souvent définies clairement, reconnues universellement et relativement stables. Ces régions sont généralement reconnues par les systèmes statistiques nationaux pour la collecte de données. En revanche, Eurostat exclut les régions analytiques (ou fonctionnelles) qui regroupent les zones suivant des critères géographiques ou socio-économiques, même si ces régions analytiques semblent définir des périmètres pertinents pour les actions de politiques économiques et sociales. En effet, pour des raisons pratiques liées à la disponibilité des données et à la mise en œuvre des politiques régionales par une autorité politique reconnue, la nomenclature NUTS se fonde essentiellement sur les divisions institutionnelles actuellement en vigueur dans les Etats membres.

Deuxièmement, la nomenclature NUTS favorise les unités régionales d'un caractère général, bien que des unités géographiques spécifiques à certaines zones d'activité (régions minières, aires d'emploi, aires agricoles, aires urbaines, etc.) soient définies par certains états membres. Par exemple, en Europe, plusieurs auteurs utilisent les *Functional Urban Regions* ou FUR définies par Hall et Hay (1980) et Cheshire et Hay (1989)⁴⁹. Ces régions sont définies en termes de concentration d'emplois. Pour chaque concentration d'emplois, on ajoute toutes les unités spatiales à partir desquelles les travailleurs migrent quotidiennement plus souvent vers cette concentration que vers les autres. Chaque FUR consiste alors en une ville-centre et sa sphère d'influence en termes de marché du travail. Le découpage, effectué selon cet indicateur, semble ainsi plus pertinent pour l'étude des sphères d'influence des zones urbaines et leurs spécificités en termes de mobilité et de marchés locaux du travail que pour d'autres études portant davantage sur les politiques sociales. D'une façon générale, le découpage régional effectué suivant un certain indicateur sera moins pertinent, voire non approprié, pour un indicateur différent. Par ailleurs, les périmètres des régions définies de cette façon sont susceptibles d'être variables dans le temps et posent donc des problèmes de comparaison temporelle. Pour cette raison, Eurostat exclut ces unités au profit d'unités plus générales. Nous sommes cependant conscients des limites imposées par Eurostat-Regio sur l'utilisation de régions administratives.

Troisièmement, les NUTS constituent une classification hiérarchique à 3 niveaux des régions pour chaque Etat membre. Ces niveaux sont nommés : NUTS 1, NUTS 2 et NUTS 3. Ainsi, la nomenclature NUTS subdivise chaque état membre en un nombre entier de régions de niveau 1 dont chacune à son tour est subdivisée en un nombre entier de régions de niveau 2 qui sont elles-mêmes subdivisées en un nombre entier de régions de niveau 3.

⁴⁹ C'est le cas, par exemple, de Cheshire et Carbonaro (1995) Magrini (1999) ou de Cheshire et Magrini (2000).

Au niveau régional, la structure administrative des états membres comprend généralement deux niveaux principaux (*Länder* et *Kreise* en Allemagne, *régions et départements* en France, *Comunidades autonomas* et *provincias* en Espagne, *regioni* et *provincia* en Italie, etc.). Le niveau NUTS correspondant peut être 1 et 3 (Allemagne), 2 et 3 (France) ou 1 et 2 (Belgique). Pour que chaque état membre possède des unités régionales comparables aux 3 niveaux NUTS, un niveau régional supplémentaire a été créé par Eurostat pour chaque état. Ce niveau correspond à des structures administratives moins importantes, voire non-existantes, et leur niveau de classification dépend de l'état membre. Par exemple, les régions NUTS 1 en France, en Italie, en Grèce et en Espagne correspondent à des « méta-régions ». Enfin, en raison de leurs aires ou populations réduites, certains pays ne possèdent pas tous les niveaux régionaux. Ainsi, le Luxembourg et le Danemark ne possèdent pas de région NUTS 1 ou NUTS 2 et sont définis comme pouvant être de niveau 0, 1 ou 2.

Les niveaux NUTS correspondent à des réalités politiques très différentes selon les Etats membres. Le tableau 3.1 donne la correspondance entre les niveaux NUTS et les divisions administratives nationales.

	NUTS 1		NUTS 2		NUTS 3	
Belgique	Régions	3	Provinces	11	Arrondissements	43
Danemark	-	1	-	1	-	1
Allemagne	Länder	16	Regierungsbezirke	38	Kreise	445
Grèce	Groups of development regions	4	Development regions	13	Nomoi	51
Espagne	Agrupacion de comunidades autonomas	7	Comunidades autonomas + Ceuta y Mellila	17 + 1	Provincias + Ceuta y Mellila	50 + 2
France	Z.E.A.T. + DOM	8 + 1	Régions + DOM	22 + 4	Départments + DOM	96 + 4
Irlande	-	1	-	1	Regional Authorities	8
Italie	Gruppi di regioni	11	Regioni	20	Regions	103
Luxembourg	-	1	-	1	Provincie	1
Pays-Bas	Landselen	4	Provincies	12	COROP regio's	40
Autriche	Gruppen von Bundesländern	3	Bundesländer	9	Gruppen von Politischen Bezirken	35
Portugal	Continente + Regioes autonomas	1 + 2	Commissaoes de coordenação regional + Regioes	5 + 2	Grupos de Concelhos	30
Finlande	Manner – Suomi / Ahvenanmaa	2	Suuralueet	6	Maakunnat	19
Suède	-	1	Riksområden	8	Län	24
Royaume-Uni	Standard regions	11	Groups of counties	35	Counties/Local authority	65
EUR 15		77		206		1031

Source : Regional Statistics reference guide (2001)

Tableau 3.1 : Correspondance entre les niveaux NUTS et les divisions administratives nationales pour les 15 pays de l'Union Européenne

Malgré le souci d'Eurostat de situer à un même niveau de la classification NUTS des régions de dimension comparable, il subsiste à chaque niveau des régions très différentes par leur superficie, leur population, leur poids économique ou leurs compétences administratives. Ainsi, Casellas et Galley (1999) soulignent qu'au niveau NUTS 2, des régions très diverses sont considérées comme étant « comparables » :

- des villes et des aires métropolitaines (Hamburg, Bremen, Darmstadt en Allemagne, la région Ile-de-France...)
- des pays (Luxembourg, Irlande...)
- des agglomérations de petites îles (Voreio Aigaio, Ionia Nisia en Grèce...)
- des territoires lointains (Guyane, Guadeloupe, La Réunion en France, Madeira au Portugal...)
- de grandes régions rurales (Calabria en Italie, Extremadura en Espagne, Alentejo au Portugal...)

Un autre problème lié au découpage spatial concerne les flux de travailleurs migrants journaliers qui faussent les comparaisons entre les pays et surtout entre les régions. Dans le cas du Luxembourg ou des régions-états, comme Hamburg, l'excédent de migrants alternants se traduit généralement par une activité productrice dans ces régions, supérieure à celle qu'elle serait avec la seule population active résidente. En conséquence, l'indicateur PIB par tête dans ces régions a tendance à surestimer la productivité de ces régions et à sous-estimer celle des régions dans lesquelles vivent les migrants alternants. Cet effet est significatif dans peu de cas seulement (Bruxelles, Luxembourg, Hamburg, Darmstadt, Bremen, Ile-de-France). Les régions pauvres ne sont généralement pas affectées par ce problème.

311.2 Ecus et parités de pouvoir d'achat

La comparaison internationale d'agrégats économiques, tel que le PIB, exige d'une part, que les agrégats soient comparables du point de vue de la méthodologie statistique (comparabilité interne) et d'autre part, que les agrégats comparables exprimés en unités monétaires nationales soient convertis en une monnaie commune (comparabilité externe).

La cohérence dans la base de mesure des agrégats est obtenue grâce à l'adhésion au système européen des comptes nationaux. Dans Eurostat-Regio, en effet, le PIB est établi selon un système de comptabilité nationale, qui, pour les états membres de l'Union Européenne, est le système européen de comptes économiques intégrés (SEC). Jusqu'en 1995, les données sont relevées selon le système SEC 1979.

L'unité de mesure commune est l'Ecu⁵⁰. La conversion est effectuée à l'aide des taux de change officiels. Cependant, ces derniers ne reflètent pas le pouvoir d'achat réel d'une monnaie donnée d'un pays. En effet, les taux de change sont très fluctuants, ils sont déterminés par de nombreux facteurs qui influent sur l'offre et la demande des monnaies et ils reflètent généralement d'autres éléments que les seules différences de prix. En conséquence, leur emploi comme facteur de conversion ne permet pas une comparaison véritable des volumes de biens et de services produits et consommés dans les différents pays. Ainsi, l'utilisation des taux de change sous-estime généralement les volumes réels pour les pays qui ont des niveaux de prix relativement bas et vice-versa.

Puisque les cours de change ne compensent pas automatiquement les différences de pouvoir d'achat au sein des pays et des régions en Europe, il importe d'utiliser les parités de pouvoir d'achat entre les monnaies nationales. Ces dernières ont été conçues pour servir d'une manière plus appropriée de facteurs de conversion et pour tenir compte des différences de prix qui ne sont pas reflétées par les cours de change. Les parités de pouvoir d'achat sont exprimées dans une unité de référence arbitraire, un numéraire, appelé Standard de Pouvoir d'Achat (SPA). Le facteur de conversion de l'Ecu en SPA pour les pays plus pauvres, c'est-à-dire généralement les pays qui ont un niveau des prix plus bas, est supérieur à un. En revanche, les pays se caractérisant par un niveau de prix plus élevé ont un facteur de conversion inférieur à un (Eurostat, 2000).

La construction de comptes régionaux en SPA comparables dans l'espace et dans le temps est néanmoins relativement compliquée et pose de sérieux problèmes dans le contexte des régions européennes.

Premièrement, la conversion de l'Ecu en SPA devrait normalement être basée sur les parités de pouvoir d'achat régionales. Cependant, suite à un manque de disponibilité des données au niveau régional, la conversion est effectuée à l'aide de parités de pouvoir d'achat nationales. Cet ajustement ne permet donc pas de prendre en compte des différences dans les prix régionaux, qui peuvent être significatives, particulièrement lorsqu'il y a de fortes variations de revenus entre les régions. Ainsi, le coût de la vie est beaucoup plus élevé au nord de l'Italie qu'au sud, et, en principe, il faudrait en tenir compte dans les chiffres de PIB régional par habitant.

Deuxièmement, les données obtenues à l'aide de cette méthode sont principalement destinées à comparer les pays et les régions pour une même année. Utiliser ces données dans des comparaisons en séries temporelles peut être problématique pour trois raisons. Tout d'abord, elles conduisent à des résultats irréalistes pour les taux de croissance annuels moyens (Vanhoudt *et al.*, 2000). Ensuite, le PIB par habitant en termes de SPA peut changer dans une économie par rapport à une autre, non seulement à cause d'une différence dans le taux de croissance du PIB par tête en termes réels mais aussi en raison d'un changement dans les niveaux de prix relatifs. Cela peut compliquer l'analyse des changements intervenus dans le temps, dans la mesure où une augmentation relative du PIB par tête, qui est due à une diminution du niveau relatif des prix, ou d'une ré-estimation de l'ajustement par les SPA,

⁵⁰ Anciennement *European Currency Unit* remplacée par l'Euro depuis 1999.

pourrait avoir des implications légèrement différentes de celle qui résulterait d'une croissance relative du PIB par tête réel (Commission Européenne, 2001). Cet aspect n'est pas pris en compte dans les calculs de parités de pouvoir d'achat. Enfin, il convient de souligner que l'ajustement des parités de pouvoir d'achat a changé lors de chaque élargissement de l'Union Européenne, ce qui complique encore les comparaisons entre des périodes différentes.

Finalement, assurer à la fois la cohérence spatiale et temporelle est un objectif qu'il est difficile d'atteindre dans la pratique. Les parités de pouvoir d'achat n'ont pas été conçues à cet effet, mais elles ont pour but de donner la meilleure estimation possible des relativités spatiales à une date donnée. En conséquence, les comparaisons entre des périodes de temps différentes sont difficiles à cause des changements dans les données et dans les méthodes (Eurostat, 2000).

L'indicateur alternatif le plus simple au PIB par tête mesuré en SPA est le PIB par habitant mesuré en Ecus. Il indique la valeur marchande du produit de chaque région et a pour effet d'accroître l'amplitude des différences entre régions. En effet, comme le niveau des prix est corrélé positivement à la richesse d'une région, un faible PIB par habitant dans des régions pauvres tend à être partiellement compensé par un coût de la vie plus bas. En revanche, dans les pays où le coût de la vie est très élevé, le PIB par tête est nettement plus élevé lorsqu'il est mesuré en Ecus (Commission Européenne, 2001). Nous illustrons ces principales différences entre Ecus et SPA à l'aide de notre échantillon de régions européennes.

312 L'échantillon

Nous présentons d'abord notre échantillon en justifiant le découpage spatial choisi (paragraphe 312.1). Ensuite, nous illustrons avec quelques exemples les différences existant entre la mesure du PIB en Ecus et celle en SPA (paragraphe 312.2).

312.1 Le choix du découpage spatial

Notre échantillon comprend les données sur les PIB par tête pour 138 régions européennes sur la période 1980-1995 : le Royaume-Uni (11) au niveau NUTS 1 et la Belgique (11), le Danemark (1), la France (21), l'Allemagne (30), la Grèce (13), le Luxembourg (1), l'Italie (20), les Pays-Bas (9), le Portugal (5) et l'Espagne (16) au niveau NUTS 2. La liste des régions, les codes associés et les cartes représentant la localisation géographique de ces régions dans leurs pays sont présentés dans l'annexe 3.1.

Pour le Royaume-Uni, l'utilisation du niveau NUTS 1 est due au fait qu'il n'y a pas d'équivalent officiel aux unités NUTS 2. Comme mentionné précédemment, le Luxembourg et le Danemark peuvent être considérés comme des régions NUTS 2 selon Eurostat. Nous

excluons la région Groningen (Pays-Bas) en raison de certaines anomalies liées aux revenus de la mer du Nord, qui accroissent significativement son PIB par tête au début de la période (Neven et Gouyette, 1994, 1995 ; Maurseth, 2001). Nous excluons également les Iles Canaries et Ceuta y Mellila (Espagne), qui sont géographiquement isolées. Les DOM-TOM et la Corse (France), l'Autriche, la Finlande, l'Irlande et la Suède sont exclues à cause de la non-disponibilité des données sur l'ensemble de période 1980-1995 dans Eurostat-Regio. Berlin (Allemagne) et l'ensemble des régions de l'Allemagne de l'Est sont également exclues pour les raisons historiques et politiques bien connues qui se sont déroulées au cours de cette période.

Comme dans tout travail empirique, nos résultats sont conditionnés par les caractéristiques de l'échantillon utilisé. Dans notre cas particulier, ils peuvent être affectés par la géographie particulière de l'échantillon (dans le cas des régions manquantes par exemple) et par le choix du niveau d'agrégation. En effet, ce choix du découpage régional peut produire un problème de type « MAUP » (*Modifiable Areal Unit Problem*, évoqué dans le chapitre 2, paragraphe 213.32, p. 118). Ainsi, le choix du niveau NUTS 2 peut apparaître arbitraire et est susceptible d'influencer nos résultats d'inférence statistique. Si, par exemple, les régions au niveau NUTS 2 sont trop grandes pour notre variable d'intérêt (PIB par tête), il existerait alors dans notre échantillon une hétérogénéité non observée et il aurait été préférable d'utiliser le niveau NUTS 3. Inversement, les régions peuvent être trop petites. Dans ce cas, l'autocorrélation spatiale détectée ne serait qu'un artéfact provenant du découpage de zones homogènes pour la variable considérée. Il aurait alors été plus pertinent d'utiliser les régions au niveau NUTS 1. Cependant, même si, idéalement, le choix de l'unité spatiale devrait être basé sur des considérations théoriques, nous sommes contraints par la disponibilité des données.

Dans ces conditions, notre choix de préférer le niveau NUTS 2 au niveau NUTS 1, lorsque les données sont disponibles, est dicté par plusieurs raisons. Premièrement, c'est le niveau le plus désagrégé qui comporte des données sur le PIB pour un plus grand nombre de régions. En second lieu, le niveau NUTS 2 est considéré par les autorités européennes comme le cadre généralement utilisé par les Etats membres pour l'application de leurs politiques régionales. Enfin, il s'agit du niveau auquel l'éligibilité à l'objectif 1 des fonds structurels est déterminée depuis leurs réformes en 1989. L'objectif 1 des fonds structurels a pour but de promouvoir le développement et l'ajustement structurel des régions en retard de développement. Les régions éligibles à cet objectif sont les régions dont le PIB par habitant est, sur la base des données des trois dernières années, inférieur à 75% de la moyenne communautaire. Plus de 60% des ressources de l'Union Européenne sont utilisées pour mettre en œuvre les politiques structurelles assignées à l'Objectif 1. Pour les années 1980 à 1993, les régions de notre échantillon éligibles à l'objectif 1 sont les suivantes :

- Espagne : Andalous, Asturias, Castille y Léon, Castille-La-Manche, Communauté Valencienne, Extremadure, Galicie, Canaries et Murcie
- Grèce : en totalité
- Italie : Abruzzi, Basilicate, Calabrie, Campanie, Molise, Pouille, Sardaigne et Sicile

- Portugal : en totalité
- Royaume-Uni : Northern Ireland

A titre comparatif, le tableau 3.2 présente les échantillons des auteurs ayant utilisé la base de données Eurostat-Regio pour étudier les disparités de PIB par tête entre les régions européennes. Parmi les échantillons provenant exclusivement d'Eurostat, notre échantillon apparaît être l'un des plus complets pour la période étudiée. La fin de période est l'année 1995. En effet, après cette date, les données sont collectées selon un nouveau système de comptabilité, le SEC95, incompatible avec le SEC79 car aucune rétropolation des données avec le système SEC95 pour les années antérieures à 1995 n'est prévue par Eurostat. En outre, notre échantillon a l'avantage de provenir d'une source unique, au contraire de certains auteurs qui utilisent différentes sources hétérogènes. Par exemple, Barro et Sala-I-Martin (1991, 1995), Armstrong (1995a), Sala-I-Martin (1996a) ou encore Cappelen *et al.* (1999) utilisent les données Eurostat-Regio à partir des années 1970 et pour les années antérieures à 1970, ils utilisent les données collectées par Molle (1980). Dans ces dernières, les données pour l'Espagne proviennent de divers numéros de la Banco de Bilbao.

312.2 Disparités économiques et mesure du PIB par tête

Compte tenu des avantages et des inconvénients liés à l'utilisation des parités de pouvoir d'achat dans le contexte européen, nous utilisons dans ce chapitre et les suivants la série E2GDP des PIB par tête mesurés en Ecus et en SPA. Les disparités régionales sont atténuées dans le second cas. Nous illustrons ce fait par quelques exemples.

Pour l'année 1995, il apparaît que le PIB par habitant varie considérablement selon les régions. Le niveau le moins élevé (33% de la moyenne européenne de l'échantillon en Ecus et 44% de la moyenne européenne en SPA) est enregistré dans la région grecque d'Ipeiros alors que le niveau le plus élevé (240% de la moyenne européenne en Ecus et 201% de la moyenne européenne en SPA) est enregistré dans la région allemande Hamburg. En Ecus, le niveau observé à Hamburg est ainsi plus de 7 fois supérieur à celui observé à Ipeiros alors qu'en SPA, ce rapport se réduit à moins de 5. Notons que dans la région de Hamburg, la migration pendulaire joue un rôle non négligeable.

Plus généralement, en 1995, le PIB par tête exprimé en Ecus de 39 des 138 régions de l'échantillon est inférieur à 75% de la moyenne européenne. Ce chiffre s'établit seulement à 27 lorsque les PIB par tête des régions sont exprimés en SPA. Dans ce dernier cas, il s'agit de 4 des 5 régions portugaises, de 5 des 20 régions italiennes, de 11 des 13 régions grecques, de 7 des 16 régions espagnoles. En Ecus, toutes les régions grecques et portugaises sont concernées, ainsi que 3 régions espagnoles et italiennes supplémentaires et 3 régions anglaises.

Auteurs	Pays	Niveau NUTS	Période
Barro et Sala-I-Martin (1991)	Allemagne, Royaume-Uni, Italie, France, Pays-Bas, Belgique, Danemark	NUTS 1 – NUTS 2	1950 - 1985*
Armstrong (1995a)	UE 12	NUTS 1-NUTS2	1950 - 1990*
Barro et Sala-I-Martin (1995)	Allemagne, Royaume-Uni, Italie, France, Pays-Bas, Belgique, Danemark, Espagne	NUTS 1-NUTS2	1950 – 1990*
Button et Pentecost (1995)	UE 12	NUTS 1	1975 – 1990**
Dewhurst et Mutis-Gaitan (1995)	UE 12	NUTS 1	1981 – 1991
Neven et Gouyette (1994, 1995)	UE 12	NUTS 2	1975 – 1989**
Quah (1996b)	Allemagne, Royaume-Uni, Italie, Pays-Bas, Belgique, Espagne	NUTS 2	1980 – 1989
Sala-I-Martin (1996a)	Allemagne, Royaume-Uni, Italie, France, Pays-Bas, Belgique, Danemark, Espagne	NUTS 1	1950 – 1990*
Beine <i>et al.</i> (1999)	UE 12 sauf Irlande	NUTS 1	1980 – 1994
Button et Pentecost (1999)	UE 12	NUTS 1 – NUTS 2	1977 – 1990**
Cappelen <i>et al.</i> (1999)	UE 12	NUTS 1	1960 – 1995*
Jean-Pierre (1999)	UE 12	NUTS 2	1980 – 1995
López-Bazo <i>et al.</i> (1999)	UE 12	NUTS 2	1980 – 1992
Beine et Jean-Pierre (2000)	UE 12 sauf Irlande	NUTS 1	1980 – 1995
Capron (2000)	UE 12	NUTS 1 – NUTS 2 - NUTS 3***	1980 – 1996
Dunford et Smith (2000)	UE 12 + Slovaquie, Hongrie et Russie	NUTS 3***	1995
Maurseth (2001)	UE 12 sauf Irlande	NUTS 1 – NUTS 2	1980 – 1994
Baumont <i>et al.</i> (2002a)	UE 12 sauf Royaume-Uni, Irlande et Grèce	NUTS 2	1980 – 1995

Notes : UE 9 = France, Allemagne, Italie, Luxembourg, Belgique, Pays-Bas, Irlande, Royaume-Uni, Danemark.

UE 12 = France, Allemagne, Italie, Luxembourg, Belgique, Pays-Bas, Irlande, Royaume-Uni, Danemark, Grèce, Espagne et Portugal.

(*) Les auteurs ont utilisé d'autres sources qu'Eurostat-Regio.

(**) Certains calculs sont effectués sur une partie des pays ou une partie de la période à cause de la non-disponibilité des données pour certains pays sur la totalité de la période.

(***) Les données au niveau NUTS 3 ne sont que partiellement disponibles.

Tableau 3.2 : Auteurs ayant utilisé Eurostat-Regio pour étudier les disparités de PIB par tête entre les régions européennes

Le tableau 3.3 présente la moyenne des 10 et des 25 régions les plus riches et les plus pauvres en Ecus et en SPA pour 1980 et 1995. Les écarts entre les données exprimées en Ecus et les données en SPA apparaissent de façon claire. Les indices fondés sur les données en SPA réduisent considérablement les différences entre pays quant au niveau de vie : les pays les plus riches le sont en fait moins que ne le laisseraient penser les données converties à l'aide du taux de change en vigueur, les pays les plus défavorisés ne le sont pas autant que ne le font croire les calculs en écus fondés sur ces taux de change.

Année	Ecus		SPA	
	1980	1995	1980	1995
Moyenne des 10 régions les plus pauvres	35	42	51	57
Moyenne des 10 régions les plus riches	178	189	160	165
Moyenne des 25 régions les plus pauvres	45	48	60	63
Moyenne des 25 régions les plus riches	153	160	140	143

Tableau 3.3 : Disparités de PIB par tête entre 138 régions européennes (moyenne de l'échantillon = 100)

Les figures 3.1a et 3.2a représentent la distribution des PIB par tête en Ecus relatifs à la moyenne européenne en 1980 et 1995 alors que les figures 3.1b et 3.2b représentent les PIB par tête en SPA relatifs à la moyenne européenne en 1980 et 1995. Les régions les plus pauvres apparaissent ici avec un PIB par tête inférieur à 75% de la moyenne européenne. Il s'agit principalement des régions du Sud, que les PIB par tête soient en Ecus ou en SPA. Les autres régions défavorisées, dont le PIB par tête est inférieur à la moyenne de l'Union, sont situées à la périphérie et correspondent aux massifs montagneux anciens, aux régions faiblement industrialisées ou aux zones d'industries traditionnelles en crise. Les régions les plus riches apparaissent avec un PIB par tête supérieur à la moyenne européenne. Ces régions sont les zones essentiellement situées à l'est de la France et en Allemagne lorsque les PIB par tête sont exprimés en Ecus alors qu'en SPA, les régions les plus riches concernent surtout les zones centrées sur l'axe Londres-Milan.

Dans les deux cas, il convient enfin de souligner qu'à l'intérieur des pays, des disparités régionales considérables sont également observées. Ainsi, en 1995 et en SPA, 5 des 11 pays de l'échantillon comportaient des régions dont le PIB par tête le plus élevé atteignait environ le double du PIB par tête le moins élevé. Ce fut le cas, par exemple en Belgique (Bruxelles, 178% de la moyenne communautaire ; Hainaut, 85%), en Allemagne (Hamburg, 201% ; Lüneburg, 89%), en France (Ile-de-France, 171% ; Languedoc-Roussillon, 84%) et en Italie (Lombardia, 138% ; Calabria, 62%).

Dans d'autres pays cependant, la différence entre les deux régions ayant le PIB le plus élevé est beaucoup plus faible : en Grèce (Attiki, 76% ; Ipeiros, 44%), aux Pays-Bas (Noord-Holland, 125% ; Drenthe, 94%), au Portugal (Lisboa, 92% ; Alentejo, 59%) et au Royaume-Uni (South East, 117% ; Northern Ireland, 81%).

Ces différentes observations montrent qu'il est difficile de considérer les pays comme des ensembles économiques homogènes. Deux exemples particulièrement frappants sont les économies duales de l'Italie et de l'Allemagne. Il importe donc de considérer les caractéristiques et les évolutions régionales.

L'échantillon étant défini, il nous reste maintenant à présenter les matrices de poids utilisées dans ce cadre particulier des régions européennes.

313 La matrice de poids

Les caractéristiques de la configuration géographique des régions européennes nous conduisent à ne pas utiliser pour l'ensemble de nos analyses empiriques des matrices de contiguïté, mais à leur préférer des matrices de poids de deux types : les matrices des plus proches voisins et les matrices basées sur la distance inverse que nous définissons dans un premier temps (paragraphe 313.1). Nous justifions dans un second temps ce choix en donnant quelques éléments de comparaison entre ces matrices et les matrices de contiguïté (paragraphe 313.2)

313.1 Matrices basées sur la distance

Toutes ces matrices sont basées sur le calcul des distances à l'aide de la distance sphérique entre centroïdes géographiques des régions. Formellement, la distance sphérique en kilomètres entre le centroïde de la région i et le centroïde de la région j est définie de la façon suivante :

$$d_{ij} = 6378 * \arccos\{\cos|Y_i - Y_j| * \cos X_i * \cos X_j + \sin X_i * \sin X_j\} \quad [3.1]$$

où X_i est la latitude du centroïde de la région i , Y_i est la longitude du centroïde de la région i .

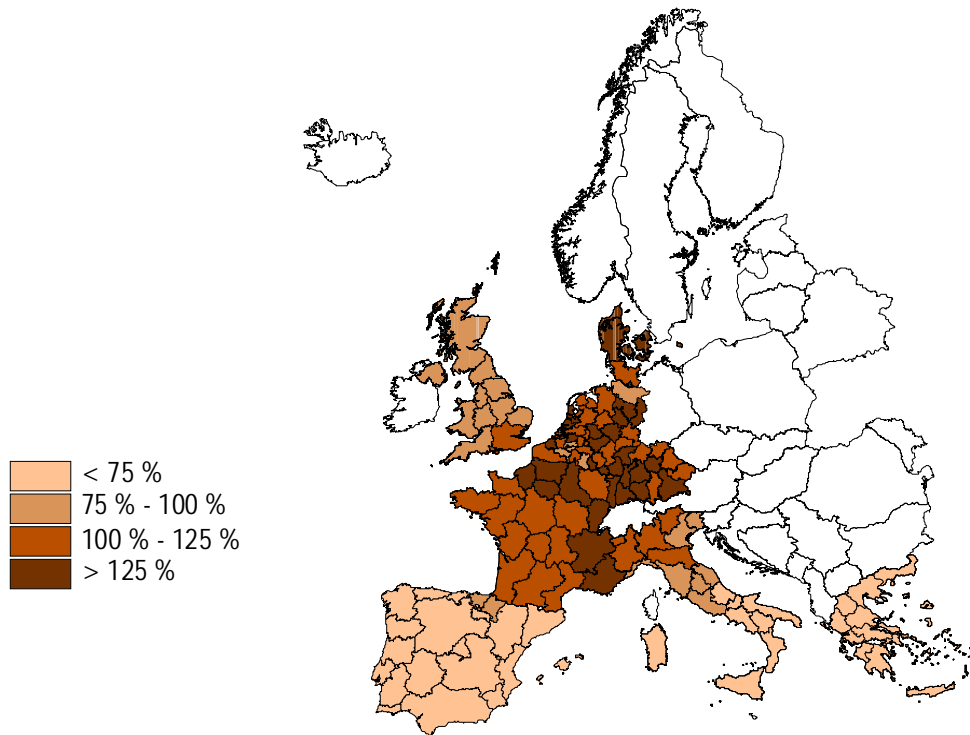


Figure 3.1a : PIB par tête en Ecus relatifs à la moyenne européenne en 1980

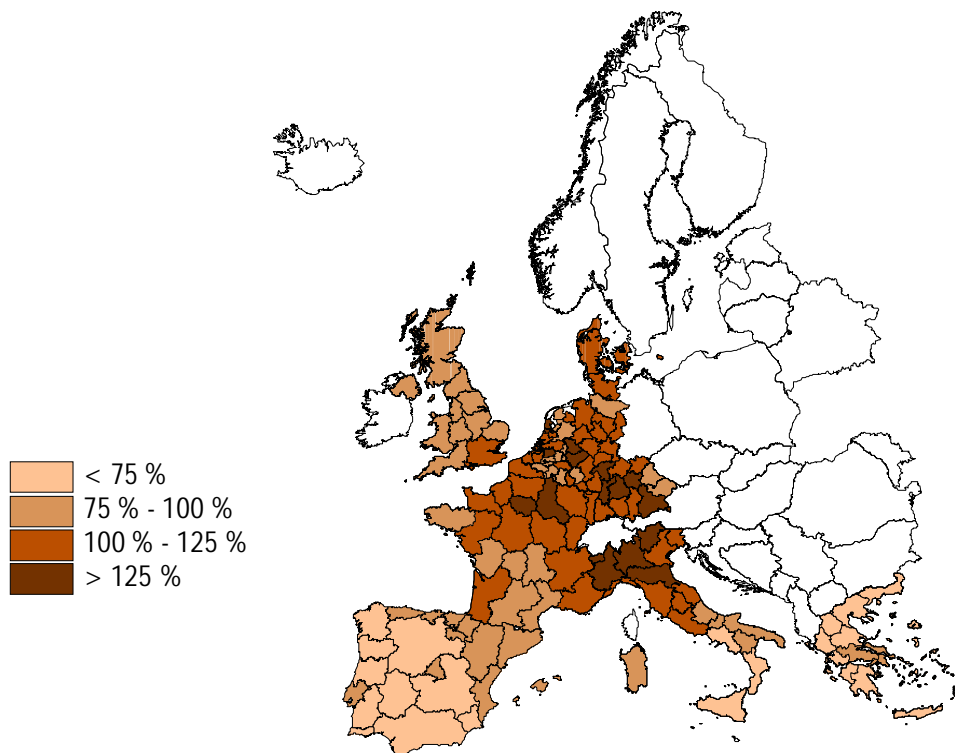


Figure 3.1b : PIB par tête en SPA relatifs à la moyenne européenne en 1980

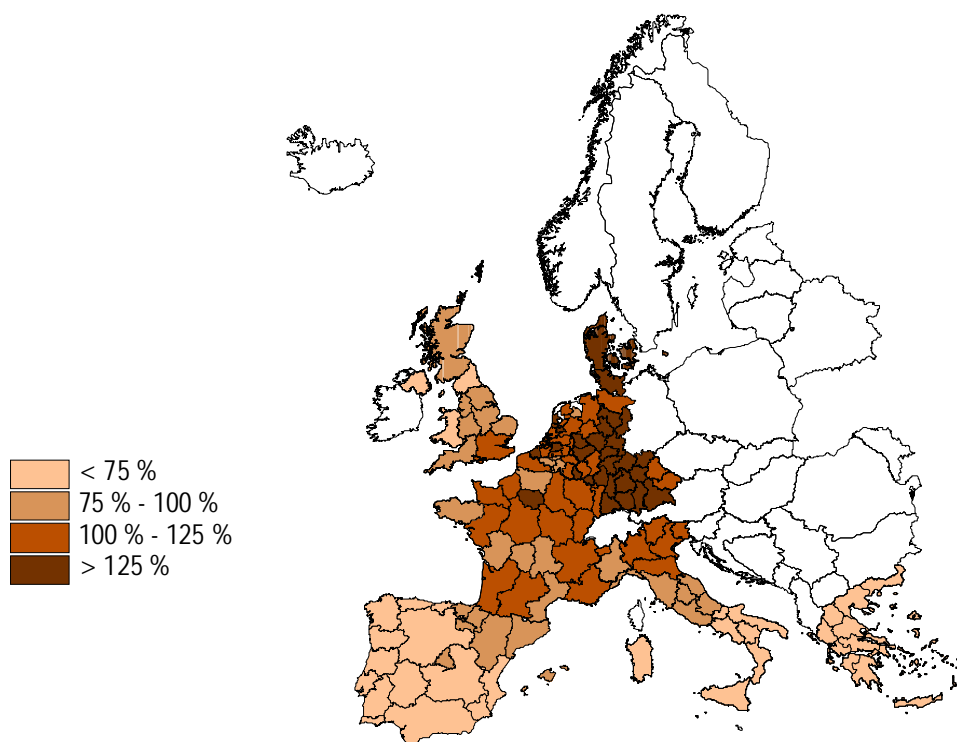


Figure 3.2a : PIB par tête en Ecus relatifs à la moyenne européenne en 1995

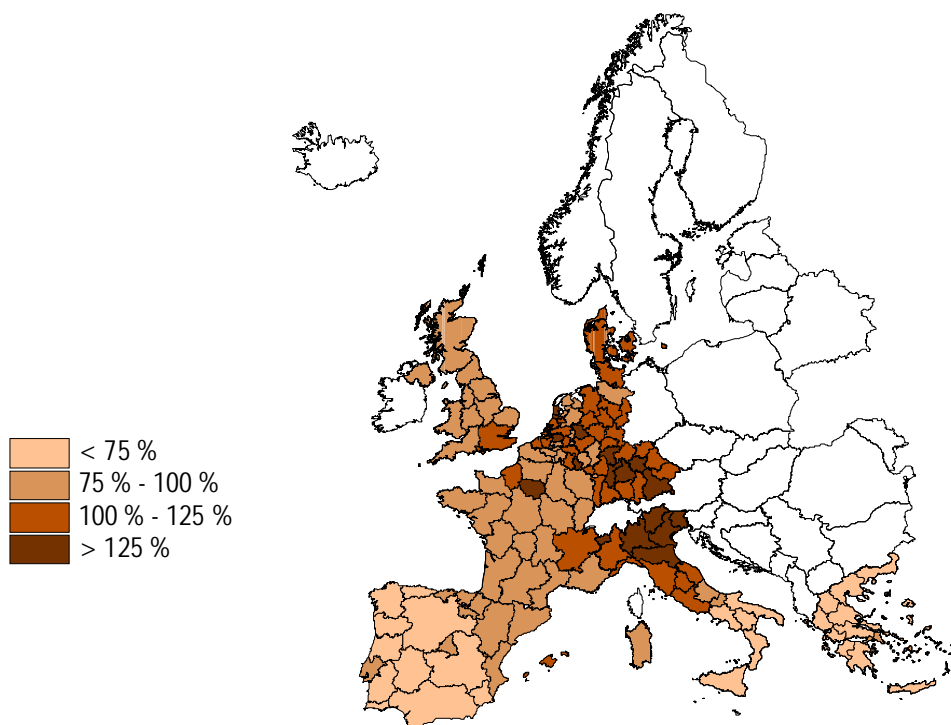


Figure 3.2b : PIB par tête en SPA relatifs à la moyenne européenne en 1995

Les matrices des k -plus proches voisins ont été utilisées par Pace et Barry (1997b), Pinkse et Slade (1998) et Baller *et al.* (2001) dans des contextes différents. La forme générale de la matrice des k -plus proches voisins est définie de la façon suivante :

$$\begin{cases} w_{ij}^*(k) = 0 & \text{si } i = j \quad \forall k \\ w_{ij}^*(k) = 1 & \text{si } d_{ij} \leq d_i(k) \\ w_{ij}^*(k) = 0 & \text{si } d_{ij} > d_i(k) \end{cases} \quad \text{et} \quad w_{ij} = w_{ij}^*(k) / \sum_j w_{ij}^*(k) \quad [3.2]$$

où $w_{ij}^*(k)$ est un élément de la matrice de poids, $w_{ij}(k)$ est un élément de la matrice standardisée en lignes (cf. paragraphe 212.3 du chapitre 2, p. 104), $d_i(k)$ est la valeur seuil définie pour chaque région i . Plus précisément, $d_i(k)$ est la plus petite distance d'ordre k entre les régions i et j , telle que la région i possède exactement k régions voisines. Toutes les analyses ultérieures étant conditionnelles au choix de la matrice de poids, nous utilisons $k = 10, 15$ ou 20 pour vérifier la robustesse de nos résultats.

La forme générale de la matrice des distances inverses s'écrit de la façon suivante :

$$\begin{cases} w_{ij}^*(k) = 0 & \text{si } i = j \quad \forall k \\ w_{ij}^*(k) = 1/d_{ij}^2 & \text{si } d_{ij} \leq D(k) \\ w_{ij}^*(k) = 0 & \text{si } d_{ij} > D(k) \end{cases} \quad \text{et} \quad w_{ij} = w_{ij}^*(k) / \sum_j w_{ij}^*(k) \quad [3.3]$$

où $D(1) = Q_1$, $D(2) = Me$ et $D(3) = Q_3$, Q_1 , Me et Q_3 sont respectivement le premier quartile (516 km), la médiane (953 km) et le troisième quartile (1501 km) de la distribution des distances entre tous les centroïdes des régions de l'échantillon. La robustesse des résultats est évaluée pour $k = 1, 2$ et 3 .

313.2 Distance ou contiguïté entre les régions européennes : quelques éléments de comparaison

Les matrices exposées précédemment sont construites en évaluant la distance entre centroïdes de régions. Dans le cas des régions européennes, elles sont préférées à l'utilisation d'une matrice de contiguïté d'ordre 1, construite à partir des frontières communes entre les régions (comme par exemple dans López-Bazo *et al.*, 1999), pour plusieurs raisons.

Premièrement, avec une matrice de contiguïté d'ordre 1, les observations isolées géographiquement sont implicitement éliminées dans les statistiques d'autocorrélation spatiale du fait de la présence de lignes et/ou de colonnes dans les matrices de poids contenant uniquement des valeurs nulles. Ceci entraîne un changement de taille de l'échantillon qui

devrait être pris en compte pour l'inférence statistique. En revanche, pour $k = 10$ (plus proches voisins) ou $k = 1$ (distance-seuil), les matrices connectent les îles telles que la Sicile, la Sardaigne et les Baléares à l'Europe continentale.

Deuxièmement, pour $k = 10$ ou $k = 1$, ces matrices connectent le Royaume-Uni à l'Europe continentale et la Grèce à l'Italie. En conséquence, cela permet d'éviter la structure bloc-diagonale de la matrice de contiguïté d'ordre 1 lorsqu'elle est ordonnée par pays. Si les matrices de poids sont bloc-diagonales, alors les régions du Royaume-Uni n'interagissent pas avec les régions françaises, belges ou néerlandaises, et les régions grecques n'interagissent pas avec les régions du sud de l'Italie. En revanche, pour les matrices utilisées ici, les connections spatiales entre régions appartenant à ces différents pays sont assurées. Par exemple, pour $k = 10$, 24,28% des 10 plus proches voisins appartiennent à un pays différent du pays d'origine. En outre, ce ratio augmente lorsque le nombre de plus proches voisins dans l'analyse de robustesse est augmenté.

Troisièmement, concernant les matrices des plus proches voisins, choisir un nombre fixe de voisins pour chaque région permet d'éviter des problèmes méthodologiques qui pourraient résulter d'un nombre variable de voisins. Ces problèmes apparaissent avec la matrice de contiguïté simple ou une matrice de distance possédant la même valeur-seuil pour toutes les régions. En revanche, les matrices des plus proches voisins ne sont pas symétriques, ce qui peut donner lieu à des problèmes d'estimation.

Toutes ces caractéristiques sont d'un intérêt particulier lorsqu'on travaille sur les régions européennes, qui sont moins connectées et moins compactes que les états américains. Dans ce chapitre, nous présentons les résultats avec les matrices des k -plus proches voisins dans le texte, alors que les résultats obtenus avec les matrices de distance inverses sont exposés dans l'annexe 3.1 pour les PIB par tête en Ecus et dans l'annexe 3.2 pour les PIB par tête en SPA.

Conclusion de section

Notre échantillon comprend 138 régions européennes sur la période 1980-1995 au niveau NUTS 2 qui est le niveau auquel l'éligibilité à l'objectif 1 des Fonds Structurels est déterminée. Les données de PIB par tête de cet échantillon sont relevées selon le système SEC 79. Nous utilisons à la fois les données de PIB par tête exprimés en Ecus et en SPA. En effet, la correction des parités de pouvoir d'achat au niveau régional permet de prendre en compte les différences de prix qui ne sont pas reflétées par les taux de change mais pose des problèmes de comparabilité temporelle. Nous avons donc choisi de présenter les résultats pour les deux mesures. L'analyse de l'échantillon a alors fait apparaître une forte persistance des disparités économiques au sein des régions européennes sur la période. Nous étudions maintenant le lien existant entre ces disparités économiques et la localisation géographique des régions à l'aide de l'analyse exploratoire des données spatiales et des diverses matrices de poids définies dans le paragraphe 312.

SECTION 32 :

ANALYSE EXPLORATOIRE SPATIALE DES PIB PAR TETE REGIONAUX

La première étape de l'analyse exploratoire consiste à évaluer l'autocorrélation spatiale globale au sein de l'échantillon afin de déterminer si, globalement, il existe une concentration spatiale de régions similaires en termes de PIB par tête en logarithmes (paragraphe 321). Ensuite, nous explicitons les concentrations spatiales de régions riches ou pauvres à l'aide de 3 outils : les statistiques de Getis-Ord (paragraphe 322), le diagramme de Moran (paragraphe 323) et les statistiques LISA (paragraphe 323). Enfin, nous évaluons la robustesse de l'ensemble de nos résultats à l'aide d'une application originale des matrices de transition (paragraphe 324).

321 Une forte autocorrélation spatiale globale

Nous avons calculé la statistique I de Moran à l'aide des matrices $W(10)$, $W(15)$ et $W(20)$. Cette statistique est présentée dans le paragraphe 213.12 du chapitre 2 (p. 108). Le tableau 3.4a fournit l'évolution de la statistique I de Moran appliquée au PIB par tête en logarithmes et en Ecus sur la période 1980-1995 pour les 138 régions européennes de notre échantillon. L'inférence statistique est basée sur l'approche de permutation avec 10 000 permutations (Anselin, 1995).

Il apparaît que les données de PIB par tête sont positivement et spatialement autocorrélées. En effet, les statistiques de Moran sont significatives avec $p = 0,0001$ pour toutes les années. Ce résultat suggère que la distribution des PIB par tête est par nature concentrée sur l'ensemble de la période. En d'autres termes, les régions possédant un PIB par tête relativement élevé (respectivement relativement faible) ont tendance à être localisées près d'autres régions possédant un PIB par tête relativement élevé (respectivement relativement faible) plus souvent que si cette localisation était purement aléatoire. Les valeurs standardisées de la statistique de Moran apparaissent être très élevées, ce qui pourrait indiquer un problème d'échelle spatiale sur notre échantillon : une partie de l'autocorrélation détectée pourrait ainsi provenir du découpage NUTS 2 adopté par Eurostat. Cependant, ce résultat ne remet pas pour autant en cause la présence d'une forte autocorrélation spatiale positive au sein de notre échantillon.

L'évolution de la statistique I de Moran dans le temps indique que les valeurs standardisées restent approximativement égales sur toute la période. Par conséquent, il existe une tendance significative et persistante à la concentration géographique de régions similaires en termes des PIB par tête en logarithmes.

Concernant les PIB par tête en logarithmes et en SPA, les résultats sont similaires même si les valeurs standardisées de la statistique I de Moran sont légèrement inférieures. Les résultats sont donnés dans le tableau 3.4b. Les tableaux 3.4a et 3.4b fournissent également les statistiques de Moran calculées avec les matrices des 15 et 20 plus proches voisins. Nous obtenons les mêmes conclusions sur le signe et la significativité de l'autocorrélation spatiale globale, ce qui souligne la robustesse des résultats vis-à-vis du choix de la matrice de poids.

Année	W(10)			W(15)			W(20)		
	I de Moran	Ecart-type	Stand.	I de Moran	Ecart-type	Stand.	I de Moran	Ecart-type	Stand.
1980	0,774	0,033940	23,024	0,742	0,027106	27,643	0,688	0,022623	30,713
1981	0,760	0,033971	22,574	0,727	0,027127	27,045	0,672	0,022635	30,011
1982	0,746	0,033956	22,161	0,710	0,027084	26,478	0,653	0,022624	29,185
1983	0,779	0,034083	23,060	0,737	0,027186	27,385	0,674	0,022728	29,985
1984	0,757	0,034019	22,446	0,710	0,027128	26,447	0,644	0,022697	28,698
1985	0,766	0,034077	22,692	0,722	0,027184	26,808	0,653	0,022752	29,032
1986	0,785	0,034126	23,213	0,733	0,02722	27,178	0,654	0,022783	29,037
1987	0,789	0,034164	23,289	0,733	0,02725	27,171	0,651	0,022837	28,827
1988	0,773	0,034196	22,802	0,719	0,027259	26,632	0,637	0,022855	28,194
1989	0,750	0,034221	22,113	0,689	0,027206	25,594	0,602	0,022836	26,678
1990	0,762	0,034242	22,461	0,698	0,027248	25,860	0,604	0,022865	26,724
1991	0,754	0,034311	22,174	0,688	0,027292	25,479	0,594	0,022938	26,205
1992	0,770	0,034323	22,651	0,708	0,027316	26,192	0,618	0,022946	27,229
1993	0,790	0,034272	23,259	0,746	0,027306	27,567	0,675	0,022895	29,805
1994	0,799	0,034267	23,514	0,759	0,027304	28,043	0,694	0,022872	30,648
1995	0,802	0,034222	23,653	0,766	0,027268	28,357	0,708	0,022829	31,308

Note : L'espérance de la statistique I de Moran est constante pour toutes les années : $E(I) = -0,007$

Tableau 3.4a : Evolution de la statistique de Moran pour les PIB par tête en logarithmes et en Ecus sur la période 1980-1995

Année	W(10)			W(15)			W(20)		
	I de Moran	Ecart-type	Stand.	I de Moran	Ecart-type	Stand.	I de Moran	Ecart-type	Stand.
1980	0,634	0,034368	18,640	0,602	0,027344	22,640	0,539	0,022788	23,949
1981	0,628	0,034410	18,447	0,591	0,027383	21,844	0,528	0,022797	23,466
1982	0,631	0,034449	18,537	0,597	0,027389	22,055	0,531	0,022828	23,553
1983	0,641	0,034469	18,791	0,602	0,027374	22,249	0,531	0,022815	23,597
1984	0,645	0,034496	18,893	0,604	0,02736	22,330	0,604	0,022832	23,655
1985	0,641	0,034527	18,781	0,605	0,027411	22,344	0,537	0,022887	23,776
1986	0,636	0,034550	18,611	0,604	0,027407	22,280	0,532	0,022913	23,517
1987	0,632	0,034543	18,489	0,598	0,027411	22,071	0,598	0,022969	23,056
1988	0,624	0,034568	18,263	0,592	0,027459	21,810	0,519	0,022992	22,871
1989	0,612	0,034551	17,920	0,580	0,027408	21,421	0,508	0,022956	22,452
1990	0,639	0,034498	18,735	0,605	0,027408	22,342	0,529	0,022932	23,366
1991	0,634	0,034472	18,603	0,599	0,027366	22,130	0,523	0,022939	23,087
1992	0,640	0,034504	18,749	0,608	0,027386	22,472	0,537	0,022947	23,716
1993	0,613	0,034511	17,981	0,589	0,027395	21,757	0,527	0,022956	23,267
1994	0,617	0,034526	18,078	0,592	0,027415	21,835	0,532	0,022962	23,486
1995	0,615	0,034518	18,010	0,587	0,027406	21,675	0,528	0,022958	23,288

Note : L'espérance de la statistique I de Moran est constante pour toutes les années : $E(I) = -0,007$

Tableau 3.4b : Evolution de la statistique de Moran pour les PIB par tête en logarithmes et en SPA sur la période 1980-1995

La statistique I de Moran fournit un résultat unique pour l'ensemble de l'échantillon. Par conséquent, dans le cas d'une autocorrélation spatiale globale positive, elle ne permet pas de discriminer une concentration spatiale de valeurs élevées de PIB par tête d'une concentration spatiale de valeurs faibles de PIB par tête. La détection de telles concentrations spatiales de régions « riches » et de régions « pauvres » est effectuée en calculant les statistiques proposées par Getis et Ord (1992) et Ord et Getis (1995).

322 Concentrations locales de régions pauvres et de régions riches : statistiques Getis-Ord $G_i(k)$

Les statistiques de Getis-Ord ont été présentées dans le paragraphe 213.21 du chapitre 2 (p. 112). Nous utilisons la statistique étendue proposée par Ord et Getis (1995) qui permet l'utilisation de matrices non symétriques telles que les matrices des plus proches voisins.

L'inférence statistique est basée sur une approximation asymptotique normale suggérée par Ord et Getis (1995). Rappelons que l'inférence dans ce cas est compliquée par le fait que les statistiques locales entre deux régions sont corrélées lorsque le voisinage de ces deux régions contient des éléments communs. En d'autres termes, il existe un problème de comparaisons multiples et il est nécessaire de fixer la significativité individuelle de chaque test en utilisant l'approximation de Bonferroni ou celle de Sidák.

Pour déterminer le nombre de comparaisons effectuées dans l'échantillon, noté m , certains auteurs (comme Pereira *et al.*, 1998) choisissent d'utiliser le nombre total de régions dans l'échantillon, soit : $m = N$. Cependant, utiliser le nombre total de régions est trop conservateur pour déterminer la significativité des statistiques $G_i(k)$ (Anselin, 1995 ; Ord et Getis, 1995). Dans le cas des matrices des k -plus proches voisins, le nombre de comparaisons ne peut pas excéder k car deux régions données ne peuvent pas avoir plus de k voisins communs. Ainsi, nous suggérons d'utiliser $m = k$ dans ce cas particulier. Nous présentons donc les résultats obtenus avec le niveau de significativité usuel de 5%, qui est trop libéral, et le pseudo-niveau de significativité de Sidák avec $m = k$. Ces deux niveaux de significativité peuvent ainsi être considérés comme les deux bornes extrêmes pour l'inférence statistique.

Pour les matrices basées sur les distances inverses (pour lesquelles les résultats sont reportés dans les annexes 3.2 et 3.3), cette approximation n'est pas possible. Nous avons donc choisi de fixer le nombre de comparaisons multiples au nombre moyen de régions voisines que possède chaque région : 34 pour $W(1)$, 69 pour $W(2)$ et 103 pour $W(3)$.

Les résultats de cette procédure sont résumés pour $k = 10$ dans le tableau de l'annexe 3.4 pour les PIB par tête en Ecus. Les colonnes 3 et 4 reportent le nombre d'années durant

lesquelles $G_i(10)$ est significativement positive ou négative. Les années correspondantes sont reportées dans les colonnes 5 et 6 ⁵¹.

En utilisant le niveau de significativité de 5%, nous détectons principalement des concentrations spatiales locales persistantes de valeurs élevées autour des régions du nord de la France, des régions belges, néerlandaises et allemandes. Notons cependant que seules les régions allemandes restent significatives au pseudo-niveau de 5% de Sidák. Les concentrations locales de valeurs faibles sont également très persistantes et concernent les régions les plus pauvres des régions européennes : la plupart des régions espagnoles, certaines régions de l'Italie du sud ainsi que toutes les régions grecques et portugaises ⁵². Ces régions pauvres sont entrées dans l'Union Européenne en 1986 et appartiennent aux « pays de la cohésion ». Elles ont bénéficié depuis 1989 de l'aide régionale au titre de l'objectif 1 mais sur toute la période, le PIB par tête de ces régions reste inférieur à la moyenne.

Les résultats pour les PIB par tête en SPA sont résumés dans le tableau de l'annexe 3.5 ⁵³. Les concentrations de valeurs faibles sont inchangées. Concernant les concentrations de valeurs élevées, deux différences principales apparaissent. Premièrement, il y a moins de régions françaises, belges, néerlandaises et allemandes significatives à 5%. On détecte ainsi des concentrations autour des régions de l'Allemagne du sud, 3 régions françaises entourant ces régions allemandes. Deuxièmement, les régions du nord de l'Italie deviennent significatives. Ces concentrations sont persistantes sur toute la période et semblent représenter le centre des régions européennes développant des activités à haute valeur ajoutée, où l'industrie est concentrée et dynamique et où l'activité commerciale est intense, en contraste aux vieilles régions industrielles. Cependant, il est possible de noter que certaines de ces dernières régions deviennent significatives à nouveau à la fin de la période.

Globalement, ces résultats indiquent une forte persistance de l'inégalité spatiale entre les régions européennes dans le temps suivant la polarisation Nord-Sud, dans la mesure où seul le voisinage de ces régions est considéré. Ce type de polarisation peut être rapproché du schéma centre-périphérie développé en économie géographique. Nous assimilons alors le centre aux régions ayant les valeurs les plus élevées de PIB par tête (le Nord) et la périphérie aux régions ayant les valeurs de PIB par tête les plus faibles (le Sud). Nous détectons ainsi une forme particulière d'hétérogénéité spatiale entre les régions européennes.

Des résultats détaillés pour 1980 et 1995 sont présentés dans les figures 3.3a et 3.4a et dans le tableau 3.14a (p. 255) pour les PIB par tête en Ecus, et dans les figures 3.3b et 3.4b et dans le tableau 3.14b (p. 256) pour les PIB par tête en SPA.

⁵¹ Notons que 54,12% de ces indicateurs sont significatifs au niveau de 5% (1 195 contre un total de 2 208) alors que seuls 29,30% des indicateurs sont significatifs au pseudo-niveau de significativité de Sidák (647 contre un total de 2 208).

⁵² Toutes les régions grecques et portugaises sont même significatives en utilisant le pseudo-niveau de significativité à 5% de Sidák.

⁵³ Notons que 46,87% de ces indicateurs sont significatifs au niveau de 5% (1 035 contre un total de 2 208) alors que seuls 26,13% des indicateurs sont significatifs au pseudo-niveau de significativité de Sidák (577 contre un total de 2 208).

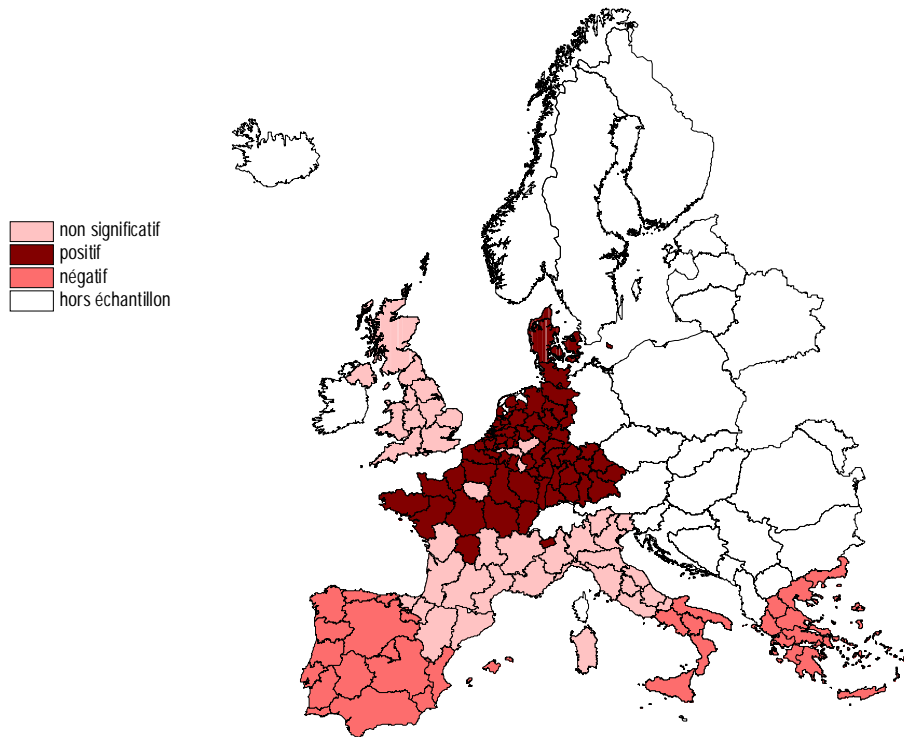


Figure 3.3a : Statistiques de Getis-Ord significatives ;
PIB par tête en logarithmes et en Ecus pour 1980

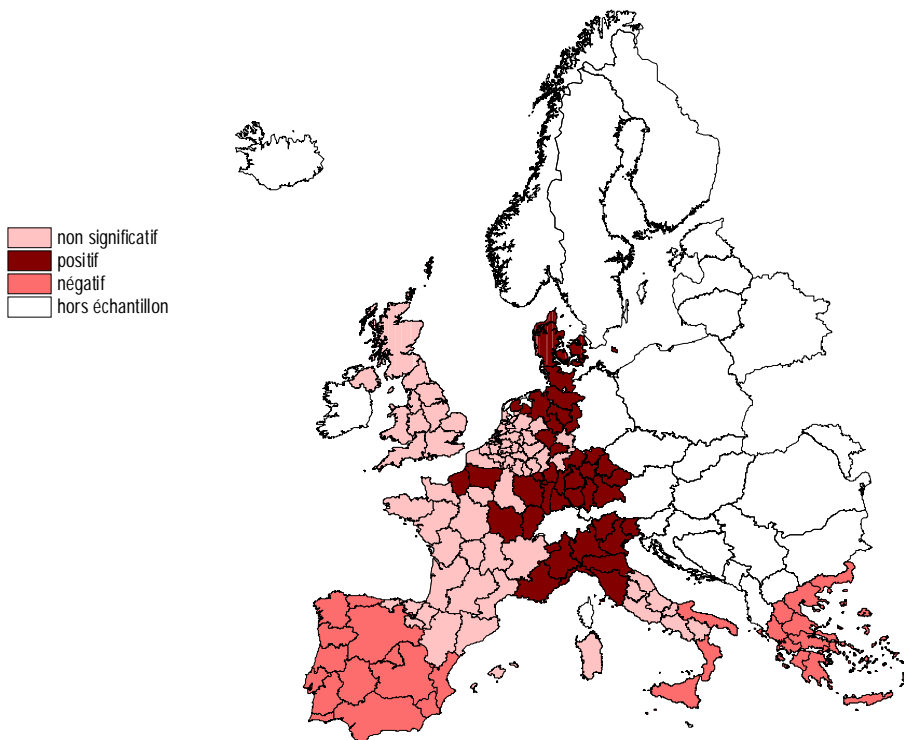


Figure 3.3b : Statistiques de Getis-Ord significatives ;
PIB par tête en logarithmes et en SPA pour 1980

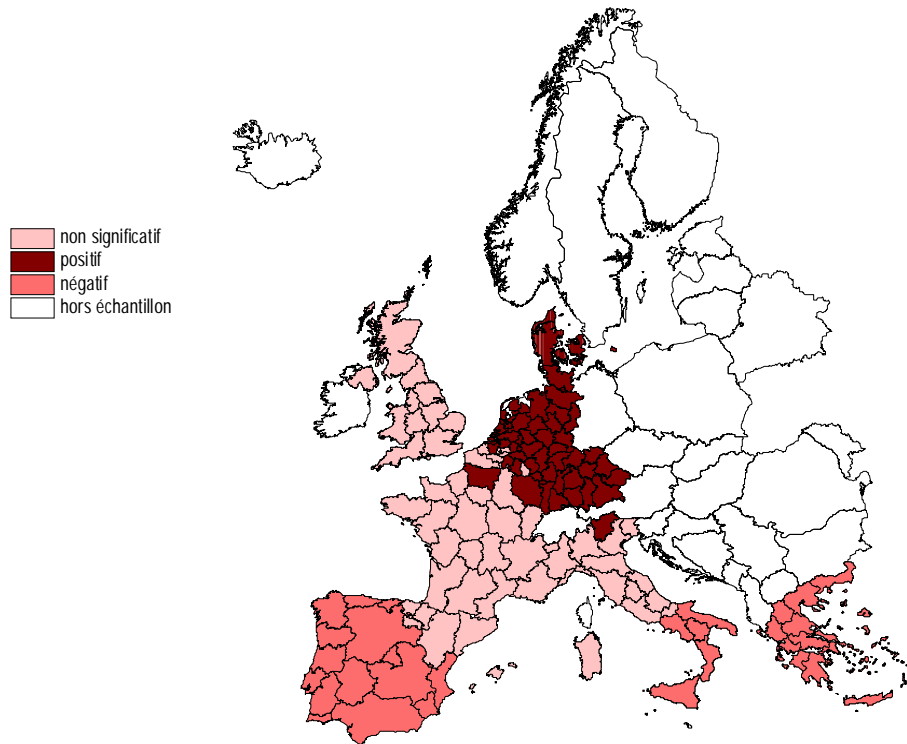


Figure 3.4a : Statistiques de Getis-Ord significatives ;
PIB par tête en logarithmes et en Ecus pour 1995

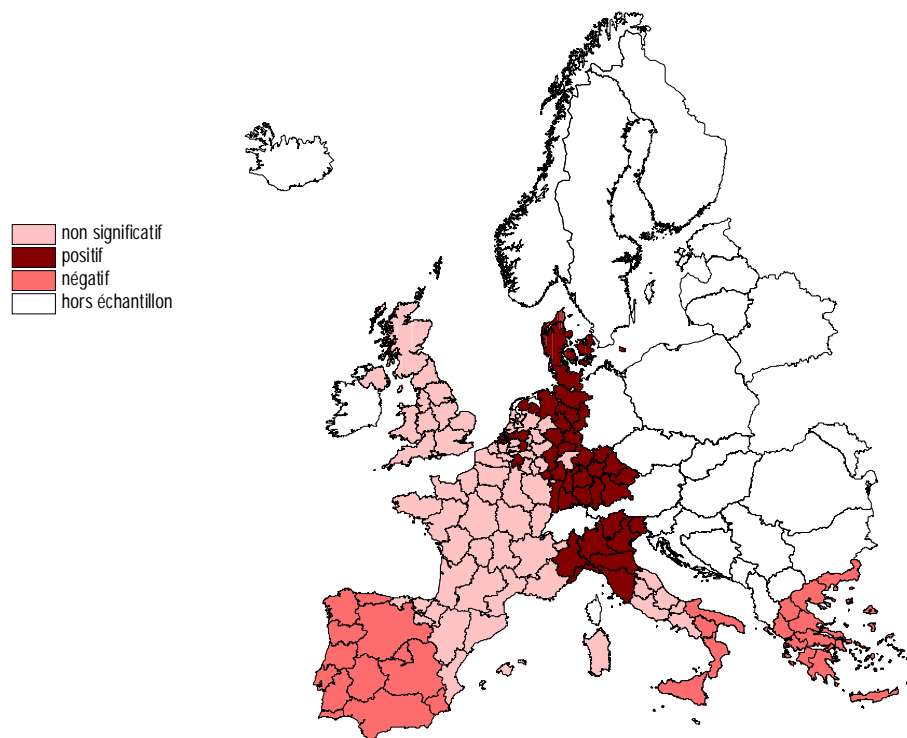


Figure 3.4b : Statistiques de Getis-Ord significatives ;
PIB par tête en logarithmes et en SPA pour 1995

Le tableau 3.5a indique les 10 valeurs les plus élevées et les 10 valeurs les plus faibles des statistiques de Getis-Ord pour les PIB par tête exprimés en Ecus. En 1980 et en 1995, les concentrations de valeurs très élevées sont exclusivement situées en Allemagne. En 1980, les concentrations de valeurs très faibles concernent principalement les régions grecques et portugaises alors qu'en 1995 toutes les régions portugaises sont remplacées par d'autres régions grecques. Le tableau 3.5b regroupe les 10 valeurs les plus extrêmes pour les PIB par tête mesurés en SPA. En 1980 et 1995, certaines régions allemandes sont remplacées par des régions du nord de l'Italie. Les 10 valeurs les plus faibles concernent principalement les mêmes régions du sud.

Toutes ces différences sont dues au fait que l'utilisation des Ecus renforce les disparités économiques de telle sorte que les régions pauvres en SPA apparaissent encore plus pauvres en Ecus. Les concentrations de valeurs faibles ne sont donc pas affectées. En revanche, certaines régions riches dans les pays pauvres, comme les régions du nord de l'Italie, semblent être pauvres en Ecus mais émergent lorsqu'on effectue la correction par la parité des pouvoirs d'achat.

Habituellement, les statistiques de Getis-Ord sont utilisées pour déterminer l'échelle spatiale appropriée en augmentant la distance critique d (dans notre cas, cela reviendrait à augmenter le nombre de voisins k). Une possibilité est de choisir la valeur de d (ou de k) qui maximise la somme en valeur absolue des statistiques $G_i(d)$ pour toutes les observations (Getis, 1995). Pourtant, dans notre cas, la somme en valeur absolue ne décroît jamais lorsqu'on augmente le nombre de voisins. Ceci indique que les effets spatiaux ne cessent pas lorsque le nombre de voisins est augmenté. Par conséquent, nous pensons qu'il est plus informatif de considérer la robustesse de ces résultats empiriques en utilisant différentes valeurs de k (voir *infra*, paragraphe 325, p. 241).

1980		1995	
Getis-Ord 10 plus grandes valeurs	Getis-Ord 10 plus petites valeurs	Getis-Ord 10 plus grandes valeurs	Getis-Ord 10 plus petites valeurs
Tübingen	Lisboa (1)	Oberpfalz	Kentriki Makedonia (1)
Freiburg	Extremadura (1)	Tübingen	Sterea Ellada (1)
Karlsruhe	Alentejo (1)	Oberfranken	Thessalia (1)
Oberpfalz	Algarve (1)	Mittelfranken	Anatoliki Makedonia (1)
Lüneburg	Centro (1)	Karlsruhe	Attiki (1)
Schleswig-Holstein	Andalucia (1)	Stuttgart	Peloponnisos (1)
Hannover	Norte (1)	Unterfranken	Dytiki Ellada (1)
Oberfranken	Sterea Ellada (1)	Braunschweig	Kriti (1)
Braunschweig	Attiki (1)	Schwaben	Dytiki Makedonia (1)
Weser-Ems	Kentriki Makedonia (1)	Lüneburg	Ionia Nisia (1)

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 3.5a : Valeurs extrêmes de la statistique Getis-Ord ;
PIB par tête en logarithmes et en Ecus pour 1980 et 1995

1980		1995	
Getis-Ord 10 plus grandes valeurs	Getis-Ord 10 plus petites valeurs	Getis-Ord 10 plus grandes valeurs	Getis-Ord 10 plus petites valeurs
Tübingen	Stereia Ellada (1)	Tübingen	Kentriki Makedonia (1)
Piemonte	Lisboa (1)	Oberpfalz	Stereia Ellada (1)
Veneto	Atiki (1)	Trentino	Thessalia (1)
Liguria	Kentriki Makedonia (1)	Oberfranken	Anatoliki Makedonia (1)
Lombardia	Thessalia (1)	Mittelfranken	Atiki (1)
Freiburg	Peloponnisos (1)	Lombardia	Peloponnisos (1)
Trentino	Dytiki Ellada (1)	Veneto	Dytiki Ellada (1)
Emilia-Romagna	Alentejo (1)	Friuli	Kriti (1)
Karlsruhe	Anatoliki Makedonia (1)	Niederbayern	Ionia Nisia (1)
Friuli	Algarve (1)	Schwaben	Dytiki Makedonia (1)

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 3.5b : Valeurs extrêmes de la statistique Getis-Ord ; PIB par tête en logarithmes et en SPA pour 1980 et 1995

Les statistiques de Getis-Ord ont permis de mettre en évidence une polarisation des régions européennes en deux groupes suivant le schéma Nord-Sud. Une autre façon de détecter les regroupements spatiaux, mais également d'analyser l'instabilité locale sous la forme de localisations atypiques, d'observations extrêmes et de régimes spatiaux, est d'utiliser simultanément le diagramme de Moran et les statistiques LISA (Anselin, 1995, 1996). Nous présentons d'abord les résultats obtenus à l'aide du diagramme de Moran.

323 Concentrations de régions similaires et détection de régions atypiques : le diagramme de Moran

Tout d'abord, nous analysons les caractéristiques de polarisation ou de stratification des régions européennes obtenues à l'aide du diagramme de Moran au début et à la fin de la période (paragraphe 323.1). Ensuite, nous vérifions la validité de ces résultats en étudiant la stabilité de ces diagrammes sur la période à l'aide des mesures de cohésion et de flux (paragraphe 323.2). Enfin, nous détectons les observations extrêmes à l'aide de divers diagnostics de régression (paragraphe 323.3).

323.1 Les diagrammes de Moran en 1980 et 1995

Dans le diagramme de Moran, la valeur standardisée du PIB par tête en logarithmes figure en abscisse et son décalage spatial (également standardisé) figure en ordonnée. Les quatre différents quadrants du diagramme correspondent alors aux quatre différents types d'association spatiale locale existant entre une région et ses voisines : HH (Haut-Haut) ; BH

(Bas-Haut) ; BB (Bas-Bas) et HB (Haut-Bas). Les quadrants HH et BB (respectivement BH et HB) représentent une association spatiale positive (respectivement négative) ⁵⁴.

Les figures 3.5a et 3.6a représentent les diagrammes de Moran pour l'année initiale et pour l'année finale de notre échantillon. Ils sont construits à partir de la matrice des 10 plus proches voisins pour les PIB par tête en logarithmes et en Ecus. L'utilisation de variables standardisées permet de comparer les diagrammes de Moran dans le temps.

Il apparaît que la plupart des régions européennes sont caractérisées par une association spatiale positive. Ainsi, en 1980, 97,8% des régions européennes relèvent d'une association de valeurs similaires (65,2% des régions se trouvent dans le quadrant HH et 32,6% se trouvent dans le quadrant BB). En 1995, 94,9% des régions européennes relèvent de cette association positive (56,5% se trouvent dans le quadrant HH et 32,4% se trouvent dans le quadrant BB). La différence principale entre 1980 et 1995 concerne les régions anglaises qui se trouvaient dans le quadrant HH en 1980 et se retrouvent dans le quadrant BB en 1995. Si l'on compare ces résultats avec ceux obtenus sur les données de PIB par tête en SPA (figures 3.5b et 3.6b), on remarque que la prédominance de l'association spatiale positive est moins prononcée. En effet, en 1980, seules 84,8% des régions européennes relèvent d'une association de valeurs similaires (53,6% se trouvent dans le quadrant HH et 31,2% se trouvent dans le quadrant BB) et en 1995, seules 81,8% des régions européennes sont caractérisées par cette association positive (47% se trouvent dans le quadrant HH et 34,1% se trouvent dans le quadrant BB). Il est intéressant de noter que différentes régions anglaises sont les seules régions du « Nord de l'Europe » à appartenir au quadrant BB en 1980 et/ou 1995. Les résultats soulignent ainsi les difficultés des vieilles régions industrielles au Royaume-Uni, comme North, North-West ou Yorkshire and Humberside.

Les diagrammes de Moran permettent également de détecter les « régions atypiques », c'est-à-dire les régions qui dévient du schéma global d'association spatiale. Il s'agit des régions qui se trouvent dans le quadrant BH ou dans le quadrant HB.

Lorsque les PIB par tête sont exprimés en Ecus, il existe très peu de régions atypiques. Ainsi, en 1980, seules 3 régions relèvent d'une association de valeurs dissemblables : Wales et Northern Ireland (Royaume-Uni) dans le quadrant BH ; Aquitaine (France) dans le quadrant HB. En 1995, il y a 7 régions atypiques : Hainaut et Namur (Belgique), Languedoc-Roussillon (France) et East Anglia (Royaume-Uni) dans le quadrant BH ; Aquitaine, Midi-Pyrénées (France) et Lazio (Italie) dans le quadrant HB.

Lorsque les PIB par tête sont exprimés en SPA, il existe davantage de régions atypiques. Ainsi, en 1980, 21 régions sont caractérisées par une association de valeurs dissemblables (18 dans le quadrant BH et 3 dans le quadrant HB). Certaines régions belges, allemandes et anglaises apparaissent être moins riches que leurs voisines et sont du type BH. Seules 3 régions HB sont détectées : Aquitaine (France), North-West et Scotland (Royaume-Uni). Cette image est différente en 1995 : 25 régions atypiques sont détectées (18 dans le

⁵⁴ Les résultats détaillés pour 1980 et 1995 se trouvent dans le tableau 3.14a (p. 255) pour les PIB par tête en Ecus et dans le tableau 3.14b (p. 256) pour les PIB par tête en SPA.

quadrant BH et 7 dans le quadrant HB). La différence majeure entre ces deux dates concerne 6 régions françaises (Picardie, Bourgogne, Nord-Pas-de-Calais, Auvergne, Bretagne et Poitou-Charentes). En 1980, ces dernières étaient riches comme leurs voisines alors qu'en 1995, elles doivent être considérées comme des régions pauvres entourées de régions riches. Inversement, 4 régions espagnoles (Navarra, Madrid, Cataluna et Islas Baleares) se retrouvent dans le quadrant HB en 1995 alors qu'elles se trouvaient dans le quadrant BB en 1980. Ceci indique un haut niveau d'activité économique comparé à leurs voisines à la fin de la période ⁵⁵.

Concernant l'existence d'une hétérogénéité spatiale entre les régions européennes, deux conclusions différentes se dégagent selon la variable utilisée.

Lorsque les PIB par tête sont mesurés en Ecus, les résultats favorisent l'existence de 2 régimes uniquement, le premier correspondant aux régions se trouvant dans le quadrant HH et le second regroupant les régions appartenant au quadrant BB. On retrouve alors le schéma de polarisation simple Nord-Sud obtenu précédemment avec les statistiques Getis-Ord.

En revanche, lorsque les PIB par tête sont exprimés en SPA, l'ensemble des résultats semblent indiquer la présence d'une hétérogénéité sous la forme de 3 régimes spatiaux distincts. Le premier correspond au schéma HH incluant principalement les régions du nord, le second correspond au schéma BB incluant surtout les régions pauvres éligibles à l'objectif 1, et le troisième correspond au schéma BH représentant une association spatiale négative. Notons que le nombre de régions appartenant au quadrant HB est trop faible pour constituer un quatrième régime. Cette dernière implication étend la polarisation simple Nord-Sud obtenue précédemment vers un schéma plus complexe de « stratification », pour reprendre les termes de Quah (1996a) dans un contexte légèrement différent ⁵⁶.

Les résultats décrits dans ce paragraphe sont basés sur l'analyse de l'année initiale 1980 et l'année finale 1995. Ils peuvent être étendus à toute la période, ce qui permet d'étudier l'évolution temporelle de la classification donnée par les diagrammes de Moran. Ceci est effectué à l'aide des statistiques de cohésion et de flux suggérées par Rey (2001).

323.2 Evolution temporelle des diagrammes de Moran

Ce qui nous intéresse ici est de savoir si une région et ses voisines occupent au cours de la période le même quadrant ou un quadrant différent dans le diagramme de Moran. Pour étudier de telles évolutions, Rey (2001) suggère de calculer différentes mesures basées sur une typologie en 4 groupes des transitions dans le temps d'une région et de ses voisines.

⁵⁵ Les diagrammes de Moran construits avec les autres matrices des plus proches voisins donnent des résultats très similaires.

⁵⁶ Quah (1996a) utilise le terme de polarisation lorsque la distribution des PIB par tête est bimodale (2 clubs de convergence) et le terme stratification lorsque cette distribution est multimodale (multiples clubs de convergence).

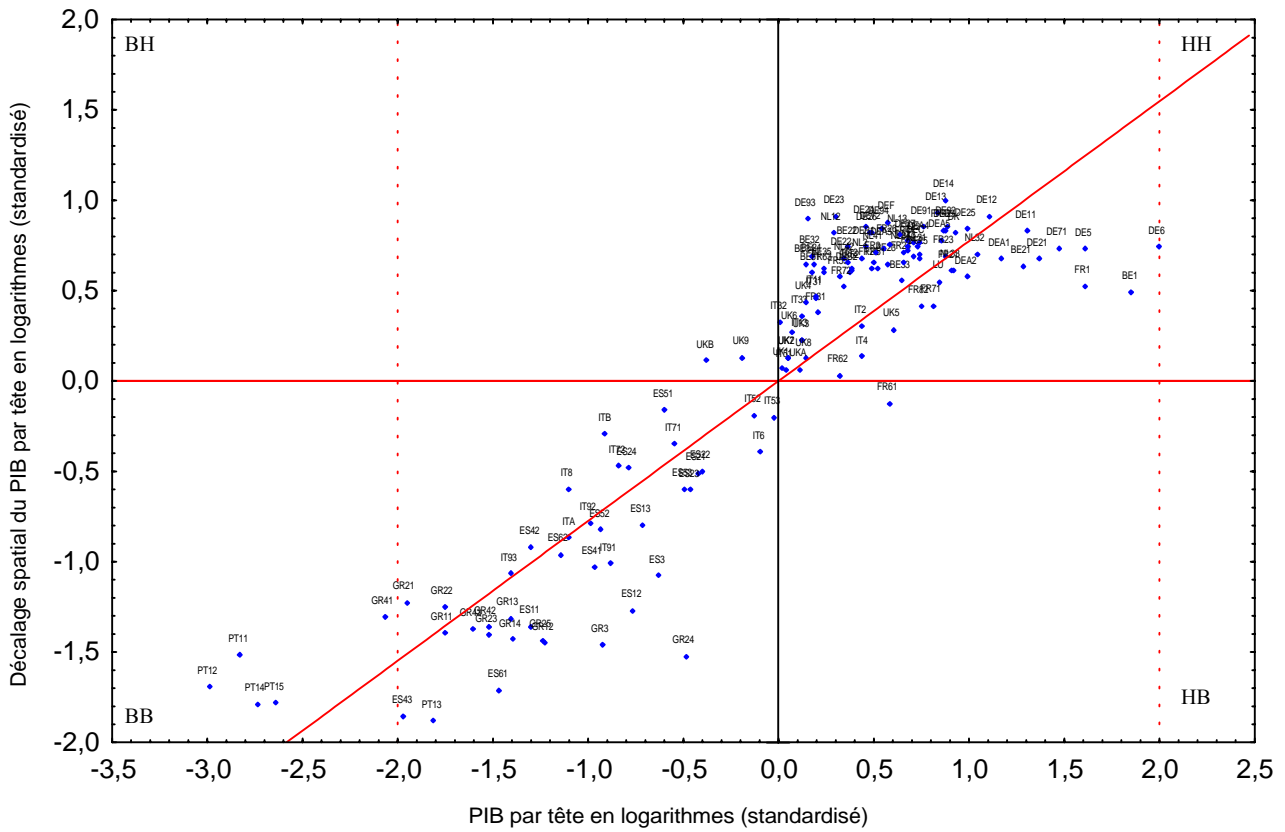


Figure 3.5a : Diagramme de Moran ; PIB par tête en logarithmes et en Ecus pour 1980

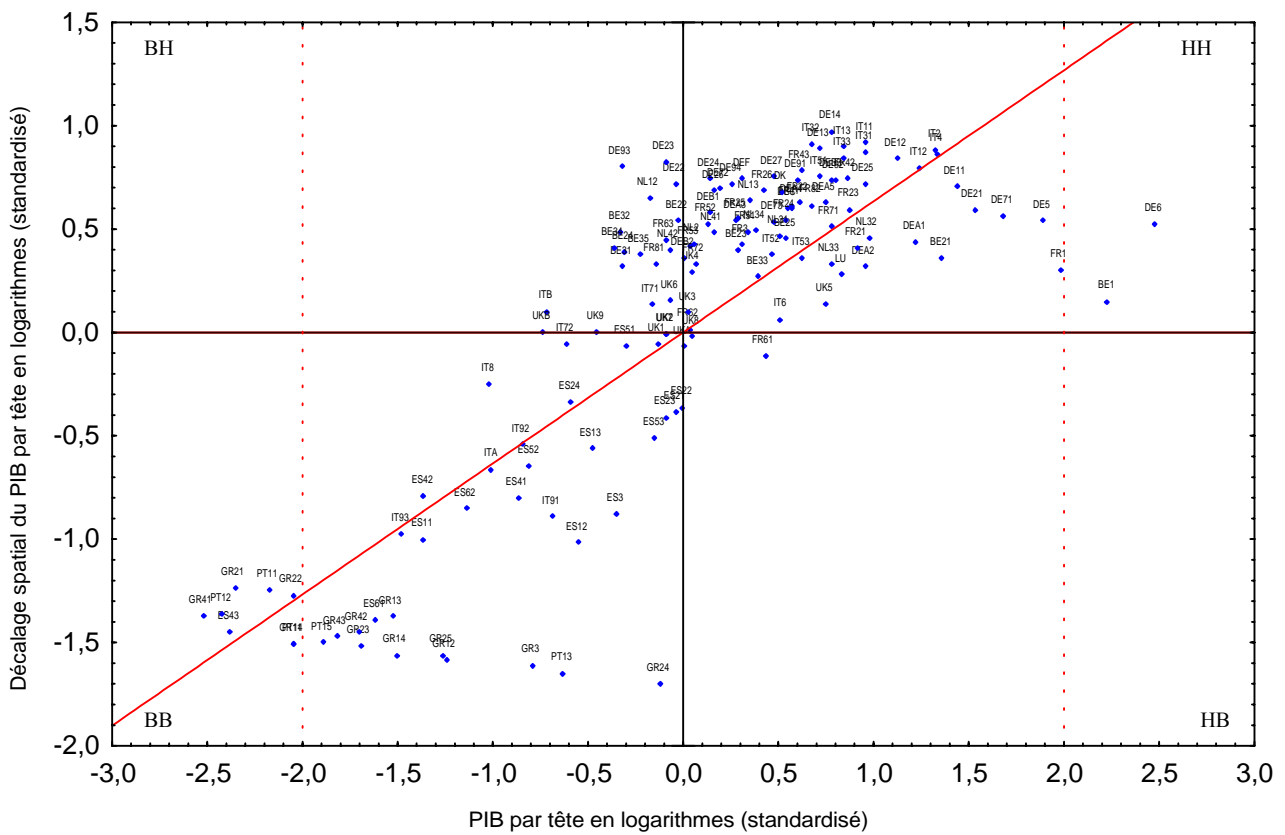


Figure 3.5b : Diagramme de Moran ; PIB par tête en logarithmes et en SPA pour 1980

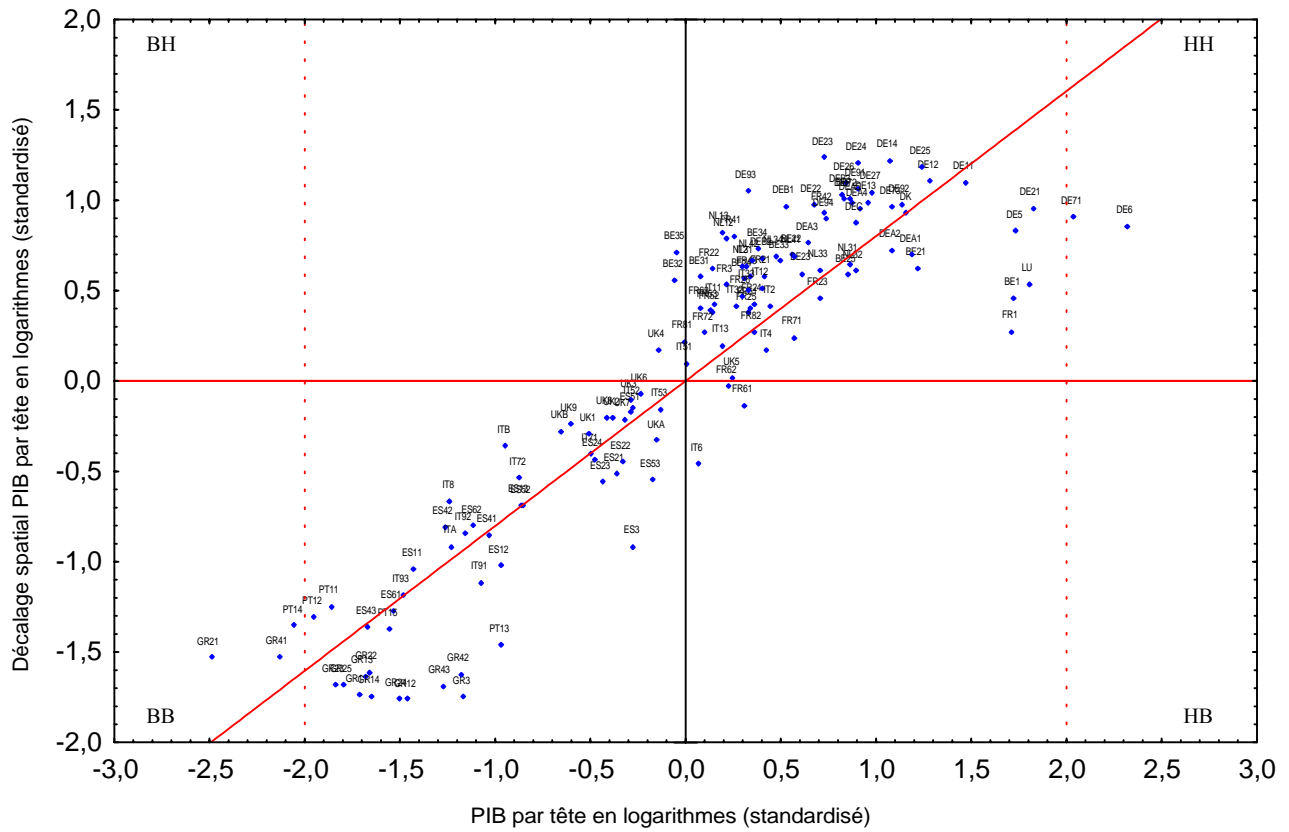


Figure 3.6a : Diagramme de Moran ; PIB par tête en logarithmes et en Ecus pour 1995

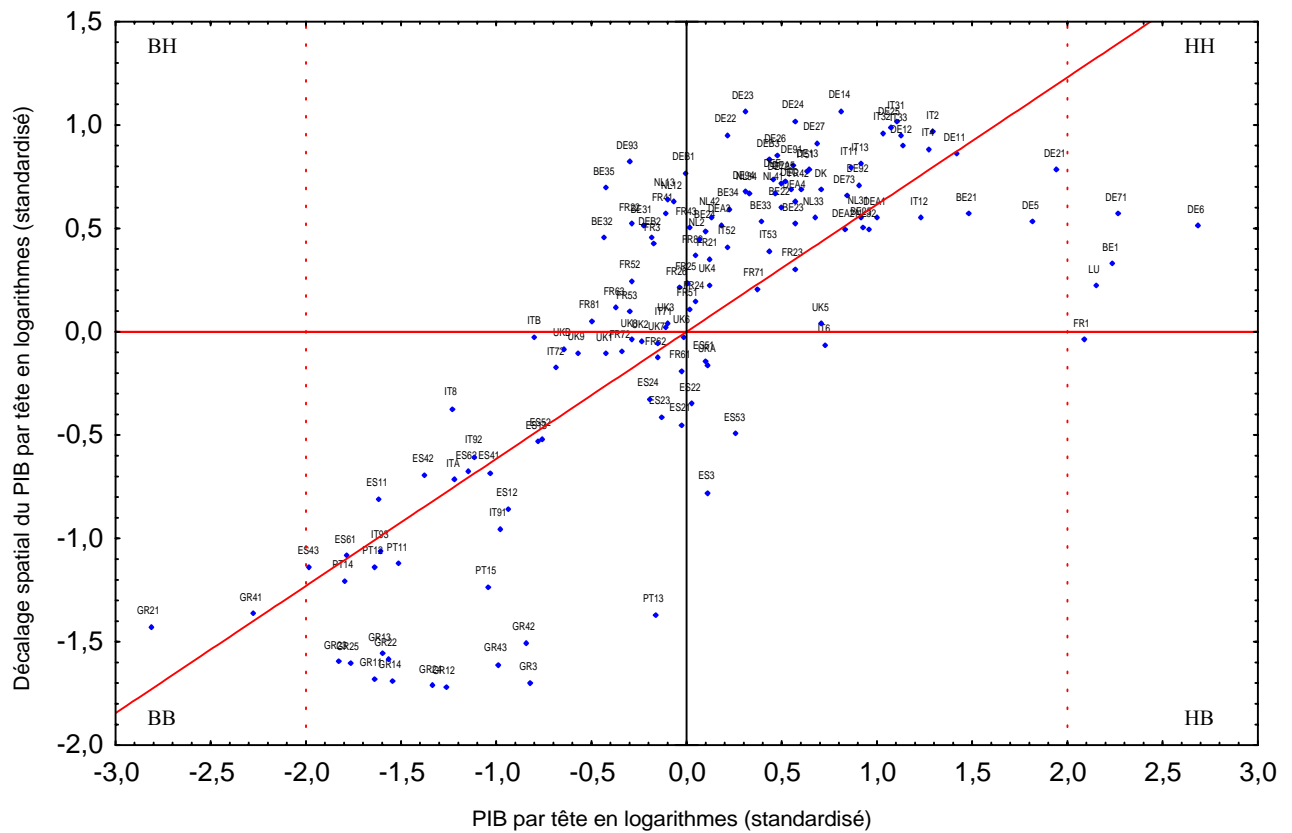


Figure 3.6b : Diagramme de Moran ; PIB par tête en logarithmes et en SPA pour 1995

- *Le premier groupe* (transitions de type I) inclut les transitions dans lesquelles seul un mouvement relatif d'une région est observé, comme par exemple, une région HH dans la première période qui devient une région BH dans la période suivante. Les autres cas sont HB-BB, BH-HH et BB-HB.
- *Le second groupe* (transitions de type II) contient les transitions dans lesquelles seules les régions voisines changent de classe : HH-HB, HB-HH, BH-BB et BB-BH.
- *Le troisième groupe* (transitions de type III) comprend les transitions dans lesquelles à la fois les régions et leurs voisines changent de classe : HH-BB, BB-HH (transition de type IIIA), HB-BH, BH-HB (transition de type IIIB).
- *Le quatrième groupe* (transitions de type IV) inclut les transitions dans lesquelles les régions et leurs voisines restent dans la même classe : HH-HH, BB-BB, HB-HB et BH-BH.

Sur la période, chaque région subit 15 transitions. Rey définit *la cohésion* comme la somme des fréquences des transitions de type IIIA et IV. En effet, il s'agit des cas où une région et ses voisines évoluent de la même façon. Par ailleurs, *le flux* est défini comme la fréquence des transitions de type I et II sur les 15 transitions de la période. Une grande stabilité dans les types de transition est donc reflétée par de faibles valeurs du flux.

Les résultats pour ces différentes mesures sont reportés dans le tableau 3.6a pour les PIB par tête en Ecus et en considérant des intervalles de temps de 1, 5, 10 et 15 ans. On remarque tout d'abord que la mobilité augmente avec l'intervalle de temps. Par exemple, pour des intervalles d'un an, la probabilité qu'une région et ses voisines restent dans le même quadrant est de 95,6%. Sur les 15 ans, cette probabilité diminue pour atteindre 89,1%. La mobilité reste cependant très faible puisque la mesure du flux n'augmente que très légèrement avec l'intervalle de temps. Les transitions les plus communes sont les transitions de type II (pour les intervalles d'1 et 5 ans) et les transitions de type I (pour les intervalles de temps plus élevés de 10 et 15 ans). Aucune transition de type IIIB n'a été observée sur les différents intervalles de temps.

Ces différents résultats dénotent une grande stabilité entre les régions européennes et leurs voisines qui ne se démarquent pas, et un faible taux de mobilité, qui n'augmente que lentement avec l'intervalle de transition. La polarisation détectée dans le paragraphe précédent reste ainsi valable sur l'ensemble de la période.

Pour les PIB par tête en SPA (tableau 3.6b), les résultats reflètent la plus grande instabilité déjà détectée dans le paragraphe précédent. La mobilité, mesurée par le flux, est bien plus importante : elle est de 5,8% pour des intervalles d'un an et atteint 20,3% pour la période de 15 ans (contre seulement 5,8% lorsque les PIB par tête sont exprimés en Ecus). Cependant, compte tenu des problèmes de comparabilité temporelle existant avec la mesure en SPA, ces résultats doivent être maniés avec précaution.

Intervalle		I	II	IIIA	IIIB	IV	Cohésion	Flux
1 an	Total	0,018	0,022	0,005	0,000	0,956	0,960	0,040
	Hausse	0,003	0,014	0,001	0,000	-		
	Baisse	0,015	0,008	0,003	0,000	-		
5ans	Total	0,036	0,043	0,022	0,000	0,899	0,921	0,079
	Hausse	0,005	0,020	0,005	0,000	-		
	Baisse	0,031	0,022	0,017	0,000	-		
10 ans	Total	0,047	0,041	0,059	0,000	0,853	0,912	0,088
	Hausse	0,007	0,022	0,011	0,000	-		
	Baisse	0,040	0,019	0,048	0,000	-		
15 ans	Total	0,036	0,022	0,051	0,000	0,891	0,942	0,058
	Hausse	0,007	0,000	0,000	0,000	-		
	Baisse	0,029	0,022	0,051	0,000	-		

Tableau 3.6a : Cohésion et flux dans les diagrammes de Moran ; PIB par tête en logarithmes et en Ecus sur la période 1980-1995

Intervalle		I	II	IIIA	IIIB	IV	Cohésion	Flux
1 an	Total	0,027	0,031	0,001	0,000	0,941	0,942	0,058
	Hausse	0,006	0,021	0,000	0,000	-		
	Baisse	0,021	0,010	0,001	0,000	-		
5ans	Total	0,054	0,055	0,013	0,001	0,877	0,890	0,109
	Hausse	0,011	0,032	0,004	0,001	-		
	Baisse	0,043	0,024	0,009	0,000	-		
10 ans	Total	0,087	0,069	0,011	0,005	0,829	0,839	0,156
	Hausse	0,018	0,034	0,002	0,004	-		
	Baisse	0,069	0,035	0,008	0,001	-		
15 ans	Total	0,116	0,087	0,014	0,000	0,783	0,797	0,203
	Hausse	0,029	0,043	0,000	0,000	-		
	Baisse	0,087	0,043	0,014	0,000	-		

Tableau 3.6b : Cohésion et flux dans les diagrammes de Moran ; PIB par tête en logarithmes et en SPA sur la période 1980-1995

323.3 Analyse des observations extrêmes

On peut détecter facilement les observations extrêmes sur le diagramme de Moran en repérant les régions se trouvant au-delà de 2 unités de l'origine du diagramme. Par ailleurs, divers diagnostics de régression permettent d'évaluer dans quelle mesure la statistique d'autocorrélation spatiale globale de Moran est influencée par la présence d'observations extrêmes dans l'échantillon.

Sur le diagramme de Moran, on visualise aisément les observations extrêmes, qui sont définies comme les points se situant au-delà de 2 unités de l'origine du diagramme. Sur les graphiques 3.5a, 3.5b, 3.6a et 3.6b, ces points sont ceux se situant au-delà des lignes parallèles à l'axe des ordonnées représentées en pointillés. Par exemple, pour 1995 en Ecus (figure 3.6a), deux régions possèdent des PIB par tête supérieurs à deux écarts-types au-delà de la moyenne : Hamburg et Darmstadt (Allemagne). Nous repérons ainsi deux régions-états dont l'activité de production est susceptible d'être surestimée à cause de l'importance des flux de migrants journaliers. Nous détectons également 3 régions possédant des PIB par tête inférieurs à deux fois l'écart-type en deçà de la moyenne : Alentejo (Portugal), Kentriki Makedonia et Voreio Aigaio (Grèce). Toutes trois sont des régions pauvres éligibles à l'objectif 1. Notons enfin qu'il n'y a jamais d'observations extrêmes sur l'axe des ordonnées. En d'autres termes, aucune région ne possède de régions voisines qui se démarquent significativement de la moyenne.

Il est encore possible de détecter les points extrêmes, qui exercent une forte influence sur la statistique I de Moran, grâce aux diagnostics de régression standards. En effet, nous avons démontré dans le paragraphe 213.22 du chapitre 2 (p. 114) que la statistique globale I de Moran est formellement équivalente à la pente de la régression linéaire de Wz sur z en utilisant une matrice de poids standardisée. Ces droites de régression sont représentées sur les graphiques 3.5a, 3.5b, 3.6a et 3.6b.

Trois diagnostics sont effectués à partir des résidus studentisés, des valeurs de levier et des distances de Cook (Belsley *et al.*, 1980 ; Donald et Maddala, 1993). Les deux premières mesures permettent de détecter les valeurs extrêmes alors que la distance de Cook est une mesure d'influence.

La valeur de levier représente la distance entre les valeurs X de l'observation i et les moyennes de toutes les valeurs X . Une valeur de levier élevée indique que l'observation i est distante de l'observation moyenne de l'échantillon. Formellement, la valeur de levier fait référence aux éléments situés sur la diagonale de la « matrice chapeau » définie par : $X(X'X)^{-1}X'$ où X est la matrice des variables explicatives de la régression. Si on note h_{ii} le levier pour l'observation i , une valeur de levier est considérée comme élevée si elle excède $2p/N$, où $p = k + 1$ (k est le nombre de variables explicatives dans la régression, soit $k = 1$) et N est le nombre d'observations (Haoglin et Welsch, 1978).

Les résidus studentisés sont une mesure du caractère extrême d'une observation sur l'ensemble des prédicteurs sur la variable dépendante. Formellement, le résidu studentisé associé à l'observation i se définit de la façon suivante :

$$r_i = \frac{\hat{u}_i}{\hat{\sigma}(1-h_{ii})^{1/2}} \quad [3.4]$$

où \hat{u}_i représente l'erreur estimée de la régression associée à l'observation i , $\hat{\sigma}$ est l'écart-type estimé de la régression et h_{ii} représente le levier de l'observation i .

La distance de Cook est une mesure de l'influence d'une observation particulière sur l'équation de la régression. Elle représente la différence entre les coefficients estimés et les valeurs qui auraient été obtenues si cette observation avait été éliminée de l'analyse. Pour une observation i , la distance de Cook est définie de la façon suivante (Cook, 1977 ; Weisberg, 1985) :

$$D_i = \frac{r_i^2 \left(\frac{h_{ii}}{1-h_{ii}} \right)}{p} \quad [3.5]$$

où $p = 2$ et h_{ii} représente le levier de l'observation i . Une observation est considérée comme influente si elle excède la valeur seuil à 5% de la loi de Fisher à p et $N - p$ degrés de liberté, soit 3,062 pour $p = 2$ et $N = 138$.

Les tableaux 3.7a et 3.7b représentent les observations extrêmes détectées selon ces différents diagnostics de régression pour 1980 et 1995, pour les PIB par tête en Ecus et en SPA respectivement. Nous avons ainsi reporté les résidus studentisés supérieurs à 2 en valeur absolue ainsi que les observations associées à une valeur de levier supérieure à $4/N$ (avec $N = 138$). Pour les deux variables, les résultats sont très similaires et concernent essentiellement des régions allemandes très riches (Hamburg, Darmstadt), les régions capitales (Ile-de-France, Bruxelles) ainsi que les régions pauvres du Portugal et de la Grèce.

En revanche, aucune région ne peut être considérée comme influente suivant le critère de la distance de Cook pour toutes les années et pour les deux variables considérées. Ce dernier résultat suggère que, même si certaines régions possèdent de forts leviers et résidus studentisés, il n'existe pas de régions qui sont particulièrement influentes dans l'échantillon. Par conséquent, la statistique I de Moran ne peut pas être considérée comme étant fortement influencée par une région particulière et l'autocorrélation spatiale globale doit être considérée comme une caractéristique générale de notre échantillon.

Les diagrammes de Moran nous ont permis de révéler des caractéristiques intéressantes en termes de PIB par tête des régions européennes. Sur la période étudiée, les disparités économiques persistent et plusieurs groupes de régions doivent être distingués. Avec les PIB par tête en Ecus, on aboutit à un résultat de polarisation en deux régimes alors qu'un processus plus complexe de stratification semble émerger avec les PIB par tête en SPA. Pourtant, puisque les diagrammes de Moran ne fournissent pas de résultats de significativité, ces différents schémas doivent être confirmés par l'utilisation des statistiques LISA.

	1980		1995	
	Région	Résidus studentisés	Région	Résidus studentisés
Résidu-type excédant 2 en valeur absolue	Stereia Ellada (1)	-3,445158	Ile-de-France	-3,139385
	Bruxelles	-2,893242	Hamburg	-2,886250
	Hamburg	-2,500151	Bruxelles	-2,654439
	Attiki (1)	-2,298205	Luxembourg (Lux)	-2,612451
	Ile-de-France	-2,225954	Attiki (1)	-2,337432
	Asturias (1)	-2,099516	Darmstadt	-2,130149
	Lüneburg	2,073019	Madrid	-2,005542
	Région	Levier	Région	Levier
Levier excédant 4/N	Centro (1)	0,072428	Ipeiros (1)	0,052553
	Norte (1)	0,065610	Hamburg	0,046399
	Alentejo (1)	0,062048	Voreio Aigaio (1)	0,040424
	Algarve (1)	0,058095	Alentejo (1)	0,038027
	Voreio Aigaio (1)	0,038353	Darmstadt	0,037616
	Hamburg	0,036314	Centro (1)	0,034994
	Extremadura (1)	0,035489	Norte (1)	0,032597
	Ipeiros (1)	0,035076	Dyptiki Ellada (1)	0,031850
	Bruxelles	0,032278	Oberbayern	0,031539
	Lisboa (1)	0,031164	Luxembourg (Lux)	0,030974
	Ionias Nisia (1)	0,029664	Peloponnissos (1)	0,030740
	Anatoliki Makedonia (1)	0,029641	Bremen	0,029290

Tableau 3.7a : Observations extrêmes du diagramme de Moran ; PIB par tête en logarithmes et en Ecus pour 1980 et 1995

	1980		1995	
	Région	Résidus studentisés	Région	Résidus studentisés
Résidu-type excédant 2 en valeur absolue	Lüneburg	2,128682	Kentriki Makedonia (1)	-2,035619
	Ile-de-France	-2,264290	Notio Aigaio (1)	-2,121667
	Hamburg	-2,490462	Kriti (1)	-2,146351
	Attiki (1)	-2,583399	Bruxelles	-2,274116
	Lisboa (1)	-2,897142	Luxembourg (Lux)	-2,383646
	Bruxelles	-2,971167	Hamburg	-2,486370
	Stereia Ellada (1)	-3,727925	Attiki (1)	-2,538986
	Région	Levier	Région	Levier
Levier excédant 4/N	Voreio Aigaio (1)	0,053695	Ipeiros (1)	0,065033
	Hamburg	0,052013	Hamburg	0,059825
	Centro (1)	0,050041	Voreio Aigaio (1)	0,044988
	Extremadura (1)	0,048691	Darmstadt	0,044826
	Ipeiros (1)	0,047648	Bruxelles	0,043646
	Bruxelles	0,043457	Luxembourg (Lux)	0,040955
	Norte (1)	0,041845	Ile-de-France	0,038955
	Ionias Nisia (1)	0,037876	Extremadura (1)	0,035889
	Anatoliki Makedonia (1)	0,037838	Oberbayern	0,034872
	Alentejo (1)	0,037740	Dyptiki Ellada (1)	0,031672
	Ile-de-France	0,035943	Bremen	0,031306
	Bremen	0,033375	Alentejo (1)	0,030798
	Algarve (1)	0,033346	Andalucia (1)	0,030518
	Kriti (1)	0,031384	Peloponnissos (1)	0,030022

Tableau 3.7b : Observations extrêmes du diagramme de Moran ; PIB par tête en logarithmes et en SPA pour 1980 et 1995

324 La significativité des concentrations locales : les statistiques LISA

Pour chaque année de la période 1980-1995, nous calculons les statistiques LISA et nous concentrons nos interprétations sur les phénomènes de regroupements spatiaux pour lesquels ces statistiques sont significatives.

Puisqu'il y a autocorrélation spatiale globale, l'inférence statistique doit être basée sur l'approche de permutation conditionnelle avec 10 000 permutations. Dans ce cas, les probabilités critiques obtenues pour les différentes statistiques locales sont en réalité des pseudo-niveaux de significativité. En outre, comme la distribution des statistiques LISA n'est probablement pas une distribution normale, nous utilisons l'approximation de Bonferroni à 5% au lieu de l'approximation de Sidák. Comme précédemment pour les statistiques de Getis-Ord, le nombre de comparaisons multiples est fixé à $m = k$, où k est le nombre des plus proches voisins dans la matrice de poids. Pour les matrices de poids basées sur les distances inverses, le nombre de comparaisons multiples est fixé au nombre moyen de régions voisines que possède chaque région.

Les résultats sont fournis dans l'annexe 3.4 pour les PIB par tête en Ecus et dans l'annexe 3.5 pour les PIB par tête en SPA. Le nombre d'années, sur toute la période, pour lesquelles les statistiques LISA sont significatives se trouve dans la septième colonne de ces tableaux⁵⁷. Le nombre d'années, pour lesquelles une région se trouve significativement dans un certain quadrant du diagramme de Moran, est reporté dans les colonnes suivantes et les années correspondantes sont dans les colonnes 12 et 13 de chaque tableau. Enfin, les cartes de significativité de Moran associées au niveau de 5% pour les années initiales et finales sont représentées dans les figures 3.7a et 3.8a pour les PIB par tête en Ecus et dans les figures 3.7b et 3.8b pour les PIB par tête en SPA⁵⁸.

Différents résultats se dégagent de l'analyse de ces cartes et de ces tableaux.

Premièrement, pour les PIB par tête en Ecus, le schéma local d'association spatiale reflète la tendance globale à l'autocorrélation spatiale positive. En effet, 98,84% des statistiques LISA significatives, en utilisant le pseudo-niveau de significativité à 5%, tombent dans le quadrant HH ou dans le quadrant BB du diagramme de Moran. Cependant, la répartition entre les associations de type HH et celles de type BB est inégale car 62,27% des régions tombent dans le quadrant HH.

⁵⁷ Pour les PIB par tête en Ecus, on note que 66,26% de ces indicateurs sont significatifs au pseudo-niveau de significativité de 5% (1 463 contre un total de 2 208) alors que seuls 41,35% des indicateurs sont significatifs au pseudo-niveau de significativité de Bonferroni (913 contre un total de 2 208). Pour les PIB par tête en SPA, on note que 58,38% de ces indicateurs sont significatifs au pseudo-niveau de significativité de 5% (1 289 contre un total de 2 208) alors que seuls 34,15% des indicateurs sont significatifs au pseudo-niveau de significativité de Bonferroni (754 contre un total de 2 208).

⁵⁸ Des résultats détaillés pour 1980 et 1995 sont reportés dans les tableaux 3.14a (p. 255) et 3.14b (p. 256).

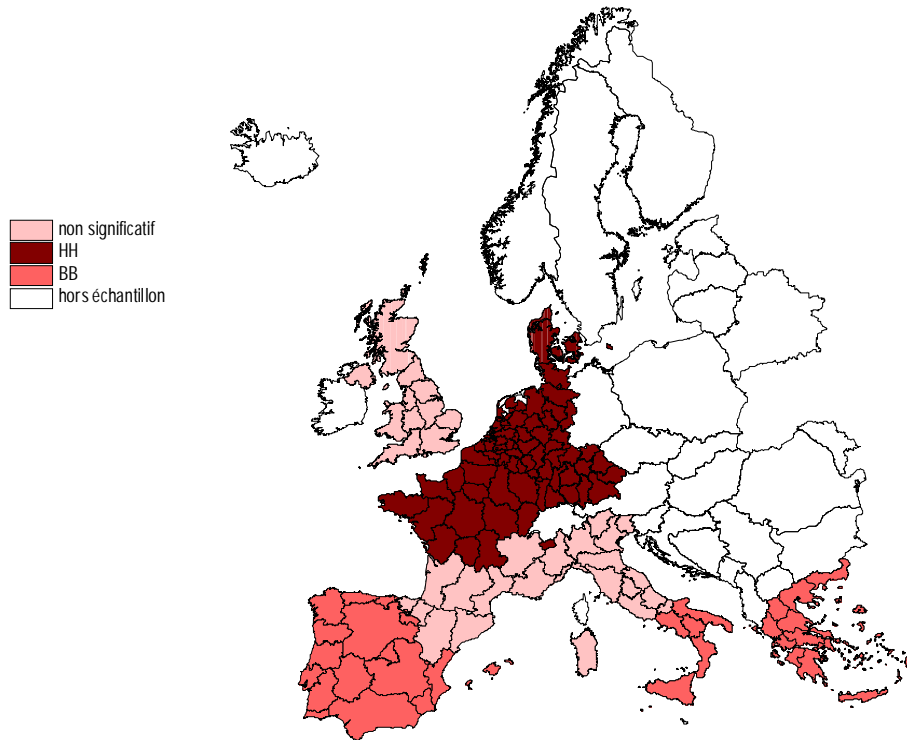


Figure 3.7a : Statistiques LISA significatives ;
PIB par tête en logarithmes et en Ecus pour 1980

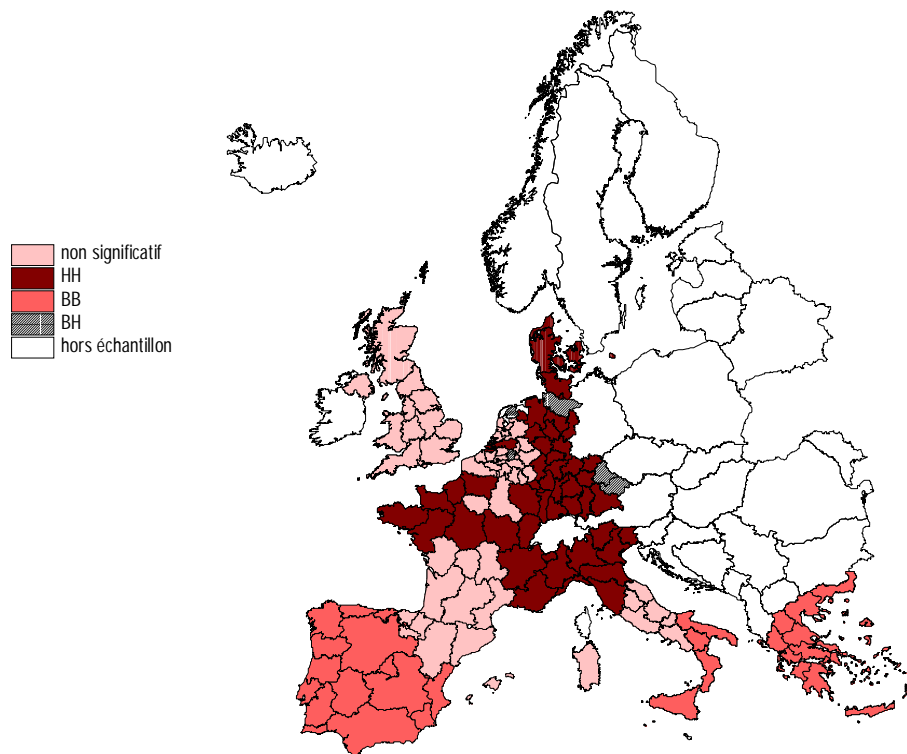


Figure 3.7b : Statistiques LISA significatives ;
PIB par tête en logarithmes et en SPA pour 1980

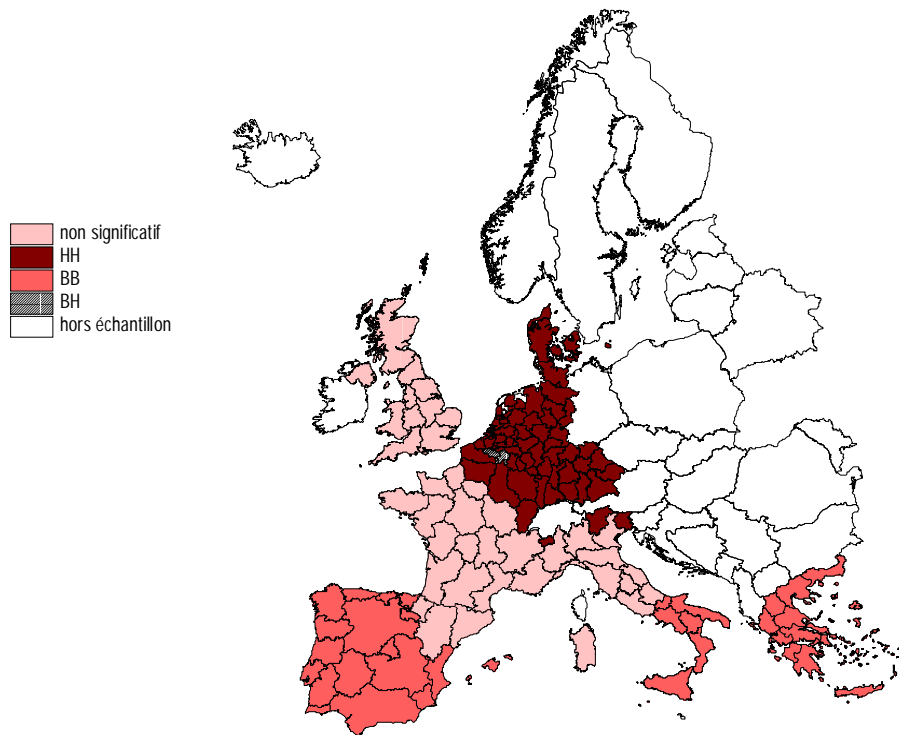


Figure 3.8a : Statistiques LISA significatives ;
PIB par tête en logarithmes et en Ecus pour 1995

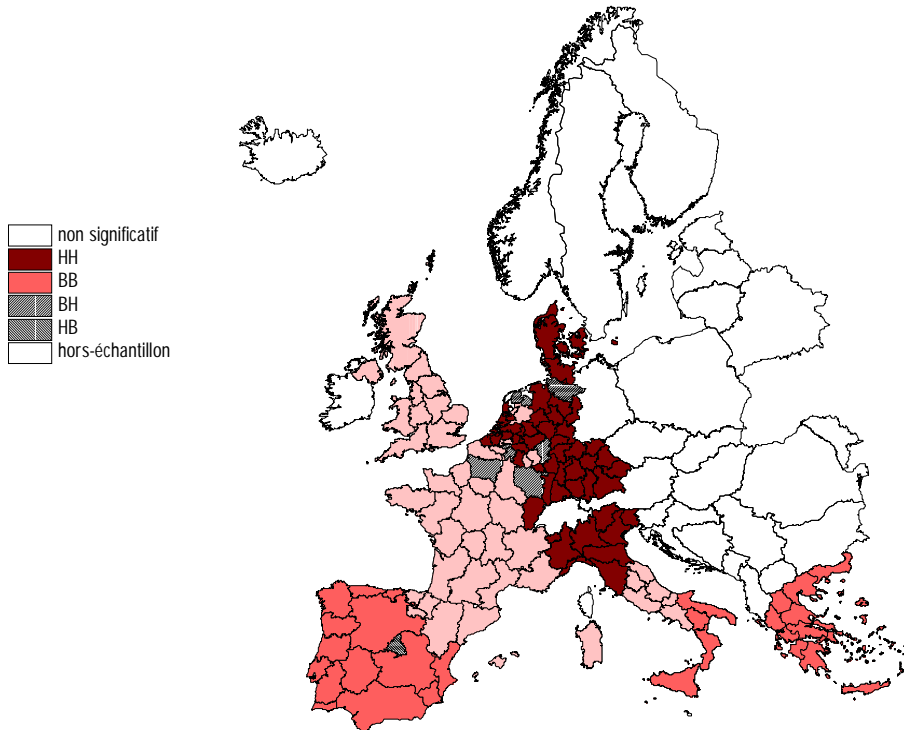


Figure 3.8b : Statistiques LISA significatives ;
PIB par tête en logarithmes et en SPA pour 1995

Ainsi, on détecte principalement des régions (ou des groupes de régions) possédant un PIB par tête élevé entourées de régions possédant un PIB par tête également élevé ⁵⁹.

Ces résultats sont très similaires pour les PIB par tête en SPA : 94,34% des statistiques LISA significatives tombent dans le quadrant HH ou dans le quadrant BB (62,27% tombent dans le quadrant HH) ⁶⁰.

Deuxièmement, pour les PIB par tête en Ecus, les déviations à la tendance globale sont extrêmement limitées et dominées par les associations spatiales négatives de type BH (1,03% des statistiques LISA significatives). Seule la région Madrid (Espagne) est significativement HB (« îlot de richesse ») pour 1991 et 1992. Les concentrations de type BH (« mouton noir ») sont le Brabant Wallon (Belgique) pour 3 ans, Hainaut (Belgique) pour 2 ans, Namur (Belgique) pour 3 ans, Friesland (Pays-Bas) pour 6 ans et Drenthe (Pays-Bas) pour seulement une année ⁶¹.

Pour les PIB par tête en SPA, les déviations à la tendance globale sont également dominées par les associations spatiales négatives de type BH (4,9% des LISA significatifs). Seules 2 régions HB sont détectées : Madrid (Espagne) entre 1987 et 1995 et Islas Baleares (Espagne) en 1982. Les régions BH sont surtout situées en Belgique (3), France (3), Allemagne (4) et Pays-Bas (3). Les régions françaises et les régions de Belgique du sud semblent donc constituer une poche de non-stationarité à la fin des années 90. Avec le pseudo-niveau de Bonferroni à 5%, seules les régions allemandes Niederbayern, Oberpfalz et Lüneburg restent significatives. Les deux premières régions étaient significatives au début de la période car leur économie était principalement basée sur l'agriculture. Elles ont connu depuis un fort développement industriel et tertiaire. Le cas de la région Lüneburg est plus spécifique puisqu'elle est entourée de 2 régions-états, Bremen et Hamburg. Elle est donc affectée par le problème des flux de migrants journaliers.

Troisièmement, pour les PIB par tête en Ecus, 4 concentrations spatiales persistent dans le temps. La première est une concentration significative de type BB entre toutes les régions portugaises et presque toutes les régions espagnoles. Deux autres concentrations de type BB sont situées entre les régions italiennes éligibles à l'objectif 1 (Campania, Puglia, Basilicata, Calabria, Sicilia) et entre toutes les régions grecques (les régions grecques et portugaises sont mêmes significatives au pseudo-niveau de significativité de Bonferroni). La dernière concentration, de type HH, concerne principalement les régions allemandes, françaises, belges et néerlandaises.

Pour les PIB par tête en SPA, les régions de l'Italie du nord deviennent significatives. En outre, presque toutes les régions françaises qui étaient significatives en 1980 n'appartiennent plus à la concentration en 1995 (sauf les régions Alsace et Franche-Comté).

⁵⁹ En utilisant le pseudo-niveau de significativité de Bonferroni à 5%, 53,89% des régions tombent dans le quadrant HH alors que 46,11% des régions tombent dans le quadrant BB.

⁶⁰ En utilisant le pseudo-niveau de significativité de Bonferroni à 5%, 44,30% des régions tombent dans le quadrant HH alors que 53,32% des régions tombent dans le quadrant BB. Ainsi, la répartition entre les quadrants HH et BB est inversée lorsqu'on considère les PIB par tête en SPA.

⁶¹ Aucune localisation atypique n'est trouvée en utilisant le pseudo-niveau de significativité de Bonferroni à 5%.

Tous ces résultats montrent une forte persistance des disparités spatiales entre les régions européennes dans le temps : les régions riches entourées par des régions riches bénéficient toujours de leur environnement positif alors que les régions pauvres entourées de régions pauvres restent négativement affectées.

En conclusion, pour les PIB par tête en Ecus, les résultats confirment l'existence de 2 régimes uniquement, le premier correspondant au schéma HH et le second au schéma BB, les deux régimes représentant l'association spatiale positive. Pour les PIB par tête en SPA, seules quelques régions sont significativement dans le quadrant BH. En conséquence, le résultat de stratification obtenu précédemment en utilisant le diagramme de Moran doit être remis en cause : la détection d'un régime de type BH n'est pas aussi évidente que pour les régimes HH et BB. En effet, en 1980 il n'y a que 4 régions significatives de type BH contre les 49 régions significatives de type HH et de type BB.

Finalement, le schéma de polarisation Nord-Sud apparaît être la caractéristique principale de notre échantillon concernant l'hétérogénéité spatiale. Ces résultats sont compatibles avec ceux précédemment obtenus avec les statistiques Getis-Ord.

A première vue, il peut sembler surprenant qu'aucune des régions anglaises ne soit significative dans nos résultats. En particulier, la région South East contenant Londres, n'est pas significativement HH, même s'il s'agit d'une des régions les plus riches en Europe. A notre avis, ce résultat est dû à l'échelle spatiale adoptée pour le Royaume-Uni (NUTS 1). Les régions à ce niveau sont sans doute trop grandes pour permettre de capter des disparités spatiales significatives entre les régions anglaises et les régions continentales du nord.

Les observations extrêmes des statistiques LISA (c'est-à-dire celles qui sont supérieures à deux fois l'écart-type de la distribution des statistiques LISA) sont reportées dans le tableau 3.8 pour les PIB par tête en logarithmes, en Ecus et en SPA pour 1980 et 1995. En 1980 pour les PIB par tête en Ecus, toutes les régions portugaises ainsi que la région Extremadura (Espagne) présentent des concentrations de valeurs très similaires. En 1980, 4 des 5 régions portugaises ainsi que les régions espagnoles Extremadura et quelques régions grecques présentent des concentrations de valeurs très similaires. La situation en 1995 est très différente puisque les régions portugaises (à l'exception d'Alentejo) sont remplacées par d'autres régions grecques, ce qui souligne la bonne performance en termes de développement économique de ces régions après l'intégration à l'Union Européenne. Pour les PIB par tête en SPA, 3 régions portugaises sont extrêmes en 1980. L'image est similaire en 1995.

Tous les calculs présentés dans les paragraphes 321 à 324 ont été effectués sur la base de la matrice des 10 plus proches voisins. Il convient maintenant d'évaluer la robustesse de ces résultats pour les autres matrices de poids.

Ecus		SPA	
1980	1995	1980	1995
Extremadura (1)	Anatoliki Makedonia (1)	Extremadura (1)	Extremadura (1)
Norte (1)	Kentriki Makedonia (1)	Anatoliki Makedonia (1)	Anatoliki Makedonia (1)
Centro (1)	Dytiki Makedonia (1)	Ipeiros (1)	Dyptiki Makedonia (1)
Lisboa (1)	Thessalia (1)	Ionia Nisia (1)	Thessalia (1)
Alentejo (1)	Ipeiros (1)	Dytiki Ellada (1)	Ipeiros (1)
Algarve (1)	Ionia Nisia (1)	Voreio Aigaio (1)	Ionia Nisia (1)
	Dytiki Ellada (1)	Notio Aigaio (1)	Dyptiki Ellada (1)
	Stereia Ellada (1)	Kriti (1)	Stereia Ellada (1)
	Peloponnisos (1)	Norte (1)	Peloponnisos (1)
	Voreio Aigaio (1)	Centro (1)	Voreio Aigaio (1)
	Alentejo (1)	Alentejo (1)	
		Algarve (1)	

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 3.8 : Valeurs extrêmes des statistiques LISA ;
PIB par tête en logarithmes, en Ecus et SPA pour 1980 et 1995

325 Des résultats robustes au choix de la matrice de poids

Il est souvent considéré comme arbitraire de travailler avec une matrice de poids plutôt qu'avec une autre. La réponse à cette critique consiste alors à répliquer les résultats avec de multiples matrices de poids. Cependant, cette réponse est intéressante si on peut disposer d'une méthode permettant d'étudier la robustesse des résultats au choix de la matrice de poids. Dans notre cas, nous considérons que les caractéristiques des concentrations locales sont robustes si une région se trouvant dans un état particulier avec la matrice des 10 plus proches voisins reste dans le même état pour les autres matrices de poids (15 ou 20 voisins). Rappelons que les différents états possibles pour les statistiques $G_i(k)$ sont : « non significatif », « positif » ou « négatif ». Pour les statistiques LISA, les différents états sont « non significatif », « HH », « BB », « HB » ou « BH ».

Ces matrices de transition constituent un outil approprié pour étudier cet aspect. En effet, elles peuvent être utilisées pour décrire les transitions entre deux distributions différentes pour un point donné du temps. Nous avons donc défini une analyse originale de robustesse à partir des matrices de transition suivantes : tableaux 3.9a et 3.9b pour les statistiques $G_i(k)$, tableaux 3.10a et 3.10b pour les statistiques LISA.

Dans chaque tableau, la première matrice reporte les transitions des régions entre les différents états possibles lorsque l'on passe de 10 à 15 voisins, tandis que la deuxième matrice reporte ces transitions lorsque l'on passe de 10 à 20 voisins. Si les résultats sont robustes à la matrice de poids, alors ces deux matrices seront diagonales : pour une année donnée, une région est dans le même état quel que soit le nombre de voisins choisi.

Pour les statistiques $G_i(k)$ et les données en Ecus (tableau 3.9a) avec le pseudo-niveau de significativité de Sidák ⁶², il y a 1 561 occurrences d'une région associée à une statistique $G_i(k)$ non-significative pour une année donnée et 10 voisins. 77,5% de ces régions sont associées à une statistique $G_i(k)$ toujours non significative la même année avec 15 voisins. En revanche, 16,3% (respectivement 6,1%) de ces régions sont associées à une statistique $G_i(k)$ significativement positive (respectivement négative) la même année pour 15 voisins.

10 à 15 voisins				
	Nombre total	non significatif	positif	négatif
non significatif	1561	0,775	0,163	0,061
positif	228	0,000	1,000	0,000
négatif	419	0,000	0,000	1,000
10 à 20 voisins				
	Nombre total	non significatif	positif	négatif
non significatif	1561	0,585	0,324	0,091
positif	228	0,000	1,000	0,000
négatif	419	0,000	0,000	1,000

Tableau 3.9a : Analyse de robustesse pour les statistiques Getis-Ord ; PIB par tête en logarithmes et en Ecus sur la période 1980-1995

10 à 15 voisins				
	Nombre total	non significatif	positif	négatif
non significatif	1631	0,861	0,078	0,061
positif	189	0,079	0,921	0,000
négatif	388	0,005	0,000	0,995
10 à 20 voisins				
	Nombre total	non significatif	positif	négatif
non significatif	1631	0,776	0,145	0,080
positif	189	0,000	1,000	0,000
négatif	388	0,000	0,000	1,000

Tableau 3.9b : Analyse de robustesse pour les statistiques Getis-Ord ; PIB par tête en logarithmes et en SPA sur la période 1980-1995

Deux résultats intéressants sont par ailleurs à souligner (tableau 3.9a pour les données en Ecus et 3.9b pour les données en SPA).

Premièrement, plus le nombre de voisins est grand, plus le nombre de régions associées à une statistique $G_i(k)$ significative est important. Ceci est reflété par la première ligne de chaque matrice. Le pourcentage de régions, associées à une statistique $G_i(k)$ non-significative pour 10 voisins et qui sont toujours associées à une statistique $G_i(k)$ non-

⁶² Tous les résultats de robustesse sont similaires en utilisant le niveau de significative de 5% avec l'approximation asymptotique normale.

significative pour 15 voisins la même année, est de 77,5%. Cette probabilité diminue pour atteindre 58,5% pour 20 voisins. Lorsqu'une région est associée à une statistique $G_i(k)$ devenant significative avec un plus grand nombre de voisins, alors la statistique $G_i(k)$ associée devient principalement positive (16,3% avec 10 voisins et 32,4% avec 15 voisins). C'est le cas pour quasiment toutes les régions allemandes, 3 régions françaises proches de l'Allemagne et 2 régions dans le nord de l'Italie. En revanche, la statistique $G_i(k)$ devient significativement négative pour toutes les régions espagnoles lorsqu'on augmente le nombre de voisins.

Deuxièmement, les sous-matrices correspondant aux deux états $G_i(k)$ (positif ou négatif) sont diagonales pour les PIB par tête en Ecus. Ceci indique que les statistiques $G_i(k)$ significatives avec 10 voisins gardent le même signe lorsque le nombre de voisins augmente. Pour les PIB par tête en SPA (tableau 3.9b), seules les régions Emilia-Romagna (Italie) et Asturias (Espagne), associées à une statistique $G_i(k)$ significative pour la matrice des 10-plus proches voisins, deviennent non significatives pour la matrice des 15-plus proches voisins (8,4%).

L'analyse de robustesse pour les statistiques LISA est fournie dans le tableau 3.10a pour les PIB par tête en Ecus et dans le tableau 3.10b pour les PIB par tête en SPA. Les matrices de transition sont calculées en utilisant le pseudo-niveau de significativité à 5% de Bonferroni⁶³. A nouveau, plus le nombre de voisins dans la matrice de poids est élevé, plus il y a de régions qui sont significativement dans le premier quadrant (17% pour 15 voisins et 28,3% pour 15 voisins). Les régions concernées sont les mêmes que celles associées à une statistique $G_i(k)$ devenant positive. Les sous-matrices correspondant aux 4 états des statistiques LISA (HH, BH, BB, et HB) sont quasi-diagonales. Par conséquent, une région associée à une statistique LISA significative reste dans le même état la même année quel que soit le nombre de voisins choisi.

10 à 15 voisins						
	Nombre total	non significatif	HH	BH	BB	HB
non significatif	1295	0,754	0,170	0,072	0,002	0,002
HH	492	0,002	0,998	0,000	0,000	0,000
BH	421	0,000	0,000	1,000	0,000	0,000
BB	0	0,000	0,000	0,000	0,000	0,000
BL	0	0,000	0,000	0,000	0,000	0,000
10 à 20 voisins						
	Nombre total	non significatif	HH	BH	BB	HB
non significatif	1295	0,586	0,283	0,110	0,002	0,019
HH	492	0,000	1,000	0,000	0,000	0,000
BH	421	0,000	0,000	1,000	0,000	0,000
BB	0	0,000	0,000	0,000	0,000	0,000
BL	0	0,000	0,000	0,000	0,000	0,000

Tableau 3.10a : Analyse de robustesse pour les statistiques LISA ; PIB par tête en logarithmes et en Ecus sur la période 1980-1995

⁶³ Tous les résultats de robustesse restent similaires en utilisant le pseudo-niveau de significativité de 5%.

10 à 15 voisins						
	Nombre total	non significatif	HH	BH	BB	HB
non significatif	1454	0,847	0,091	0,003	0,055	0,005
HH	334	0,081	0,919	0,000	0,000	0,000
BH	16	0,313	0,000	0,688	0,000	0,000
BB	402	0,000	0,000	0,000	1,000	0,000
HB	2	0,000	0,000	0,000	0,000	1,000
10 à 20 voisins						
	Nombre total	non significatif	HH	BH	BB	HB
non significatif	1454	0,750	0,134	0,028	0,079	0,009
HH	334	0,042	0,958	0,000	0,000	0,000
BH	16	0,063	0,000	0,938	0,000	0,000
BB	402	0,000	0,000	0,000	1,000	0,000
HB	2	0,000	0,000	0,000	0,000	1,000

Tableau 3.10b : Analyse de robustesse pour les statistiques LISA ; PIB par tête en logarithmes et en SPA sur la période 1980-1995

Tous ces éléments indiquent la grande robustesse des résultats au choix de la matrice de poids et au nombre de voisins.

Conclusion de section

Les différentes analyses exploratoires réalisées dans cette section soulignent, sur l'ensemble de la période 1980-1995, un schéma persistant de polarisation en termes de PIB par tête entre les régions du Nord et les régions du Sud de l'Europe. Ces résultats, robustes au choix de la matrice de poids, suggèrent la présence de deux régimes spatiaux dans notre échantillon : au Nord, les régions à hauts niveaux de PIB par tête entourées de régions à hauts niveaux de PIB par tête, tandis qu'au Sud, les régions à faibles niveaux de PIB par tête sont entourées de régions à faibles niveaux de PIB par tête. Pourtant, ces caractéristiques restent relatives et ne nous permettent pas d'induire des résultats sur le dynamisme économique de ces régions. En effet, les indicateurs locaux d'association spatiale sont relatifs à la moyenne européenne. Par conséquent, si une région possède un PIB par tête inférieur à la moyenne en début et en fin de période, cela ne signifie pas pour autant qu'elle est restée aussi pauvre. Pour caractériser les liens entre localisation et dynamisme économique, il convient donc d'étendre l'analyse exploratoire des données spatiales sur les taux de croissance annuels moyens entre 1980 et 1995.

SECTION 33

ANALYSE EXPLORATOIRE SPATIALE DES TAUX DE CROISSANCE ANNUELS MOYENS

Les résultats précédents ont mis en évidence la persistance de fortes disparités spatiales en termes de niveaux de PIB par tête sur la période 1980-1995. Cependant, l'analyse des taux de croissance annuels moyens est susceptible de nuancer cette affirmation. En effet, si le taux de croissance des régions pauvres est plus élevé que celui des régions riches, les inégalités économiques entre le Nord et le Sud s'atténueront avec le temps. Ainsi, nous avons déjà souligné la bonne performance économique de certaines régions pauvres, telles que les régions portugaises. Par exemple, ces dernières étaient associées à des observations extrêmes pour les statistiques LISA en 1980, ce qui signifie qu'elles étaient caractérisées par de très faibles niveaux de PIB par tête et qu'elles étaient entourées de régions également caractérisées par de très faibles PIB par tête. En revanche, ces régions ne faisaient plus partie en 1995 des observations extrêmes des statistiques LISA. Ce résultat suggère qu'elles ont fait partie, durant la période, de concentrations géographiques de régions ayant connu de forts taux de croissance.

Pour corroborer ces enseignements, nous étudions dans cette section les performances économiques des régions européennes en appliquant les techniques de l'analyse exploratoire des données spatiales aux taux de croissance annuels moyens 1980 et 1995. Tout d'abord, nous évaluons l'autocorrélation spatiale globale et locale pour les taux de croissance annuels afin de déterminer les concentrations locales de taux de croissance (paragraphe 331). Ensuite, nous étudions la relation existant entre le taux de croissance annuel moyen et le niveau initial du PIB par tête afin d'analyser dans quelle mesure les régions les plus pauvres en début de période ont connu des taux de croissance plus élevés (paragraphe 332).

331 Les concentrations spatiales de taux de croissance

Le taux de croissance annuel moyen d'une région i sur la période 1980-1995 est approximé de la façon suivante :

$$g_i = \frac{1}{15} [\ln y_{i,1995} - \ln y_{i,1980}] \quad [3.6]$$

où $y_{i,1980}$ et $y_{i,1995}$ représentent le PIB par tête de la région i respectivement en 1980 et en 1995.

A l'aide de cette approximation, nous reprenons la démarche appliquée précédemment pour les niveaux de PIB par tête. Nous évaluons d'abord l'autocorrélation spatiale globale (paragraphe 331.1). Ensuite, nous déterminons les concentrations locales de régions à forts ou à faibles taux de croissance à l'aide des statistiques de Getis-Ord (paragraphe 331.2), du diagramme de Moran (paragraphe 331.3) et des statistiques LISA (paragraphe 331.4).

331.1 Une forte autocorrélation spatiale globale positive

L'autocorrélation spatiale globale des taux de croissance annuels moyens entre 1980 et 1995 est analysée à l'aide de la statistique I de Moran. Les résultats révèlent une autocorrélation spatiale globale positive. En effet, la valeur standardisée de la statistique I est égale à 12,709 pour les PIB par tête en Ecus. Cette valeur est associée à une probabilité critique de 0,0001. Lorsque les PIB par tête sont exprimés en SPA, la valeur standardisée de la statistique I est égale à 5,955, associée avec une probabilité critique de 0,0002.

On note que ces valeurs standardisées sont bien inférieures à celles calculées pour les niveaux de PIB par tête en logarithmes. Ces résultats indiquent néanmoins que les régions associées à un taux de croissance annuel moyen relativement élevé (respectivement relativement faible) sont localisées près d'autres régions associées à des taux de croissance annuels moyens relativement élevés (respectivement relativement faibles) plus souvent que si cette localisation était purement aléatoire. Nous explicitons ces concentrations de régions à forts et à faible taux de croissance à l'aide des statistiques locales d'autocorrélation spatiale.

331.2 Concentrations locales et statistiques Getis-Ord

Comme précédemment, l'inférence statistique est basée sur une approximation asymptotique normale et le nombre de comparaisons multiples est fixé à 10 pour la matrice des 10 plus proches voisins. Pour la matrice basée sur les distances inverses (pour laquelle les résultats sont reportés dans les annexes 3.2 et 3.3), le nombre de comparaisons multiples est fixé au nombre moyen que possède chaque région, soit 34 pour $W(1)$.

Les résultats fournis par la procédure d'évaluation des statistiques Getis-Ord appliquées au taux de croissance annuel moyen calculé avec les PIB par tête en Ecus et la matrice des 10 plus proches voisins sont reportés dans le tableau 3.14a (p. 255) et dans la figure 3.9a.

Il apparaît que 34,9% des statistiques Getis-Ord sont significatives au niveau de 5%⁶⁴. Les concentrations spatiales locales de valeurs élevées sont principalement situées au Portugal, en Espagne du sud et de l'ouest, ainsi que dans les régions d'Allemagne du sud.

⁶⁴ 13,8% de ces statistiques sont statistiquement positives et 21% sont statistiquement négatives.

Ainsi, ces dernières sont les seules régions du « Nord » de l'Europe à faire partie d'une concentration locale dans laquelle les taux de croissance sont significativement élevés. Inversement, les régions anglaises, les régions françaises et quelques régions grecques du nord présentent des concentrations locales de valeurs faibles. Ainsi, ces régions grecques du Sud de l'Europe se démarquent des régions espagnoles et portugaises car elles appartiennent à une concentration locale dans laquelle les taux de croissance sont significativement faibles. Cette concentration peut alors s'interpréter comme une « trappe de sous-développement ».

Les valeurs extrêmes se trouvent dans le tableau 3.11. Elles confirment la bonne performance des régions du Portugal et de l'Espagne qui sont associées aux 10 plus grandes valeurs des statistiques Getis-Ord. En revanche, les plus faibles valeurs sont associées aux régions anglaises.

Les mêmes calculs appliqués au taux de croissance annuel moyen des PIB par tête en SPA sont reportés dans le tableau 3.14b (p. 256) et dans la figure 3.9b. Seulement 18,8% des statistiques Getis-Ord sont significatives au niveau de 5%⁶⁵. Les concentrations spatiales locales de valeurs élevées sont principalement situées au Portugal et au sud de l'Espagne. Les régions Namur et Liège (Belgique) et la région Oberfranken (Allemagne) sont également au cœur de concentrations spatiales de valeurs élevées. Seules les régions françaises font partie de concentrations locales de valeurs faibles alors que les régions anglaises ne sont plus significatives. Les valeurs extrêmes se trouvent également dans le tableau 3.11 et font ressortir la mauvaise performance des régions françaises en termes de taux de croissance.

Ecus		SPA	
10 plus grandes valeurs	10 plus petites valeurs	10 plus grandes valeurs	10 plus petites valeurs
Extremadura (1)	Northern Ireland (1)	Extremadura (1)	Bourgogne
Alentejo (1)	Scotland	Alentejo (1)	Auvergne
Lisboa (1)	Wales	Andalucia (1)	Centre
Algarve (1)	North	Lisboa (1)	Pays de la Loire
Andalucia (1)	South West	Algarve (1)	Limousin
Centro (1)	West Midlands	Centro (1)	Ile-de-France
Galicia (1)	Yorkshire and Humberside	Namur	Basse Normandie
Norte (1)	East Midlands	Galicia (1)	Poitou-Charentes
Asturias (1)	North West	Liège	Rhône-Alpes
Oberfranken	Bourgogne	Oberfranken	Franche-Comté

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 3.11 : Valeurs extrêmes de la statistique Getis-Ord ;
taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus et en SPA

⁶⁵ 7,2% de ces statistiques sont statistiquement positives et 11,6% sont statistiquement négatives.

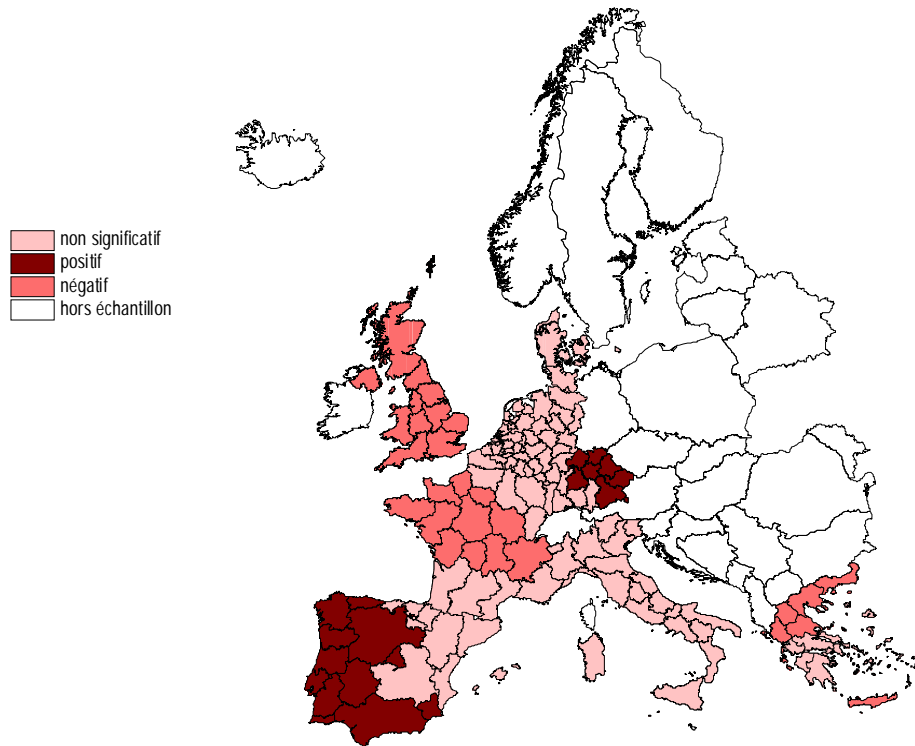


Figure 3.9a : Statistiques de Getis-Ord significatives ;
taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus

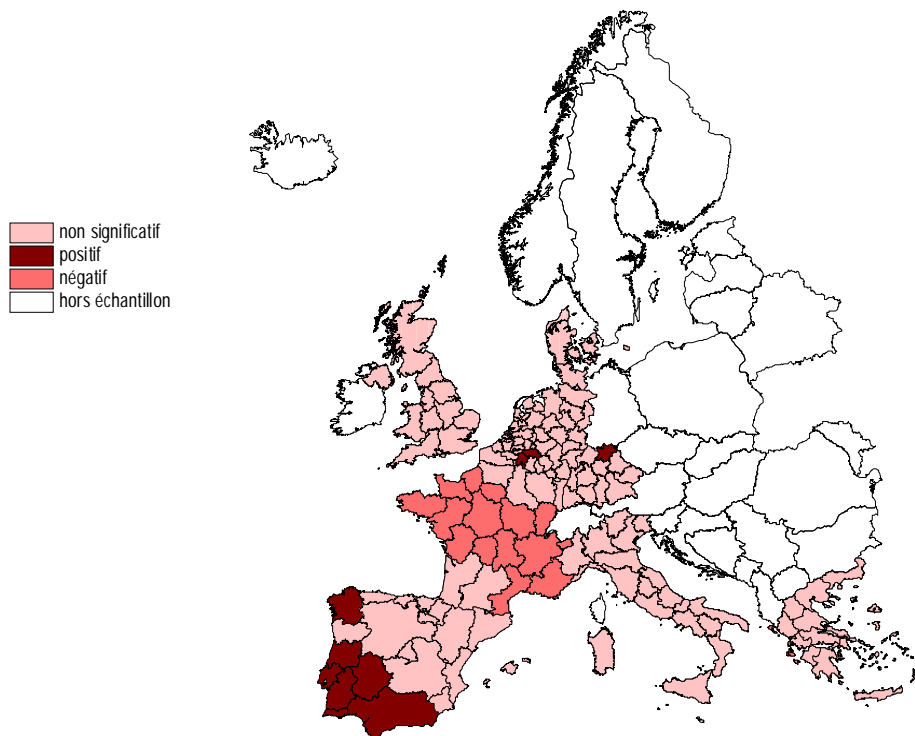


Figure 3.9b : Statistiques de Getis-Ord significatives ;
taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en SPA

331.3 Concentrations locales, régions atypiques et diagramme de Moran

Lorsqu'on compare le diagramme de Moran pour les taux de croissance en Ecus (figure 3.10a) aux diagrammes de Moran pour les niveaux en 1980 et 1995 (figures 3.5a et 3.6a), il apparaît une plus forte instabilité spatiale dans le sens où plus de régions appartiennent au quadrant HB ou au quadrant BH. En effet, seules 73,19% des régions européennes relèvent d'une association de valeurs similaires (33,33% se trouvent dans le quadrant HH et 39,86% se trouvent dans le quadrant BB). En revanche, 26,81% de ces régions sont caractérisées par une association locale de valeurs dissemblables (11,59% se trouvent dans le quadrant BH et 15,22% se trouvent dans le quadrant HB). Il apparaît également que toutes les régions portugaises ainsi que le Luxembourg possèdent des taux de croissance supérieurs à deux fois l'écart-type au-delà de la moyenne. Notons cependant que le Luxembourg se trouve dans le quadrant HB, ce qui indique que le taux de croissance de ce pays a été beaucoup plus élevé que la moyenne alors que les régions voisines ont connu des taux de croissance inférieurs à la moyenne. Seule la région Sterea Ellada a connu un taux de croissance inférieur à deux écart-types en deçà de la moyenne.

Finalement, les observations extrêmes détectées selon les divers diagnostics de régression, basés sur le diagramme de Moran, sont reportées dans le tableau 3.12. Seuls les résidus studentisés et les valeurs de levier sont représentés car il n'y avait pas d'observations influentes selon le critère de la distance de Cook. En conséquence, l'autocorrélation spatiale globale du taux de croissance n'est pas due à l'influence de quelques régions mais apparaît être une caractéristique générale de l'ensemble de l'échantillon.

Concernant les PIB par tête exprimés en SPA (figure 3.10b), on constate qu'il y a, à nouveau, plus d'instabilité dans le diagramme de Moran pour les taux de croissance que dans les diagrammes pour les niveaux en 1980 et en 1995. Ainsi, seules 59,42% des régions européennes relèvent d'une association de valeurs similaires (20,09% se trouvent dans le quadrant HH et 33,33% se trouvent dans le quadrant BB) alors que 40,58% de ces régions sont caractérisées par une association spatiale de valeurs dissemblables (21,01% se trouvent dans le quadrant BH et 19,57% se trouvent dans le quadrant HB). Trois régions portugaises (Norte, Centro et Algarve), deux régions grecques (Notio Aigaio et Kriti) et Luxembourg possèdent des taux de croissance supérieurs à deux écart-types au-delà de la moyenne. Comme précédemment, Luxembourg est une observation extrême se trouvant dans le quadrant HB. La région grecque Kriti se trouve également dans ce cas. En revanche, 2 régions françaises, se trouvant dans le quadrant BB (les vieilles régions industrielles Champagne-Ardenne et Picardie) et la région grecque Sterea Ellada, se trouvant dans le quadrant BH, ont connu des taux de croissance inférieurs à deux écart-types en deçà de la moyenne. Dans le premier cas, ce résultat indique que les régions Champagne-Ardenne et Picardie ont connu des taux de croissance très inférieurs à la moyenne alors que leurs voisines ont connu des taux de croissance inférieurs à la moyenne. Dans le second cas, Sterea Ellada apparaît être une région isolée puisqu'elle a connu un taux de croissance très inférieur à la moyenne alors que

ses voisines ont connu des taux de croissance supérieurs à la moyenne. Rappelons également que certaines de ces régions sont des observations extrêmes dans le quadrant opposé en 1980 pour les niveaux de PIB par tête. Nous revenons à cette relation inverse entre le niveau de PIB par tête en 1980 et le taux de croissance moyen dans le paragraphe 332. Finalement, les observations extrêmes détectées selon les diagnostics de régression basés sur le diagramme de Moran sont reportées dans le tableau 3.12. Comme précédemment, il n'y a pas de régions influentes au sens de la distance de Cook.

	Ecus		SPA	
	Région	Résidus studentisés	Région	Résidus studentisés
Résidu-type excédant 2 en valeur absolue	Andalucia (1)	3,497511	Andalucia (1)	2,535605
	Extremadura (1)	2,822284	Stereia Ellada (1)	2,125763
	Galicia (1)	2,745314	Namur	2,119678
	Luxembourg (Lux)	-2,666020	Algarve (1)	2,031897
	Asturias (1)	2,591420	Centre	-2,009620
	Kriti (1)	-2,436728	Luxembourg	-2,020224
	Ionia Nisia (1)	-2,195220	Pays de la Loire	-2,060235
	Notio Agaio (1)	-2,142425	Ionia Nisia (1)	-2,174145
			Ile-de-France	-2,174375
			Auvergne	-2,216134
			Bourgogne	-2,560234
	Région	Levier	Région	Levier
Levier excédant 4/N	Algarve (1)	0,105763	Luxembourg (Lux)	0,089036
	Centro (1)	0,102492	Stereia Ellada (1)	0,081831
	Norte (1)	0,089878	Notio Aigaio (1)	0,054936
	Stereia Ellada (1)	0,065942	Algarve (1)	0,054338
	Lisboa (1)	0,064531	Kriti (1)	0,052348
	Luxembourg (Lux)	0,055558	Picardie	0,052143
	Alentejo (1)	0,054656	Centro (1)	0,051240
	Picardie	0,030390	Champagne-Ardenne	0,044446
			Norte (1)	0,039800
			Lorraine	0,033541
		PACA	0,030651	

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 3.12 : Valeurs extrêmes du diagramme de Moran ; taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus et en SPA

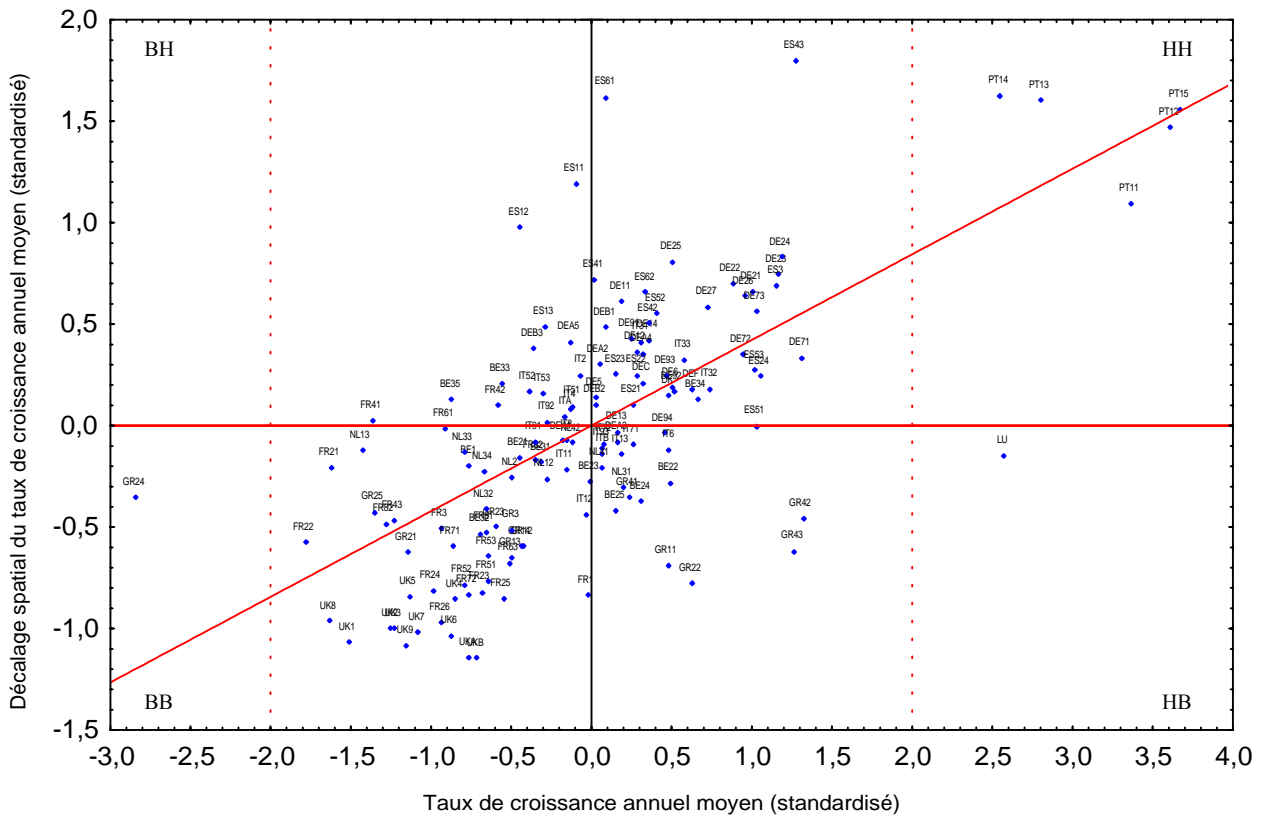


Figure 3.10a : Diagramme de Moran pour le taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus

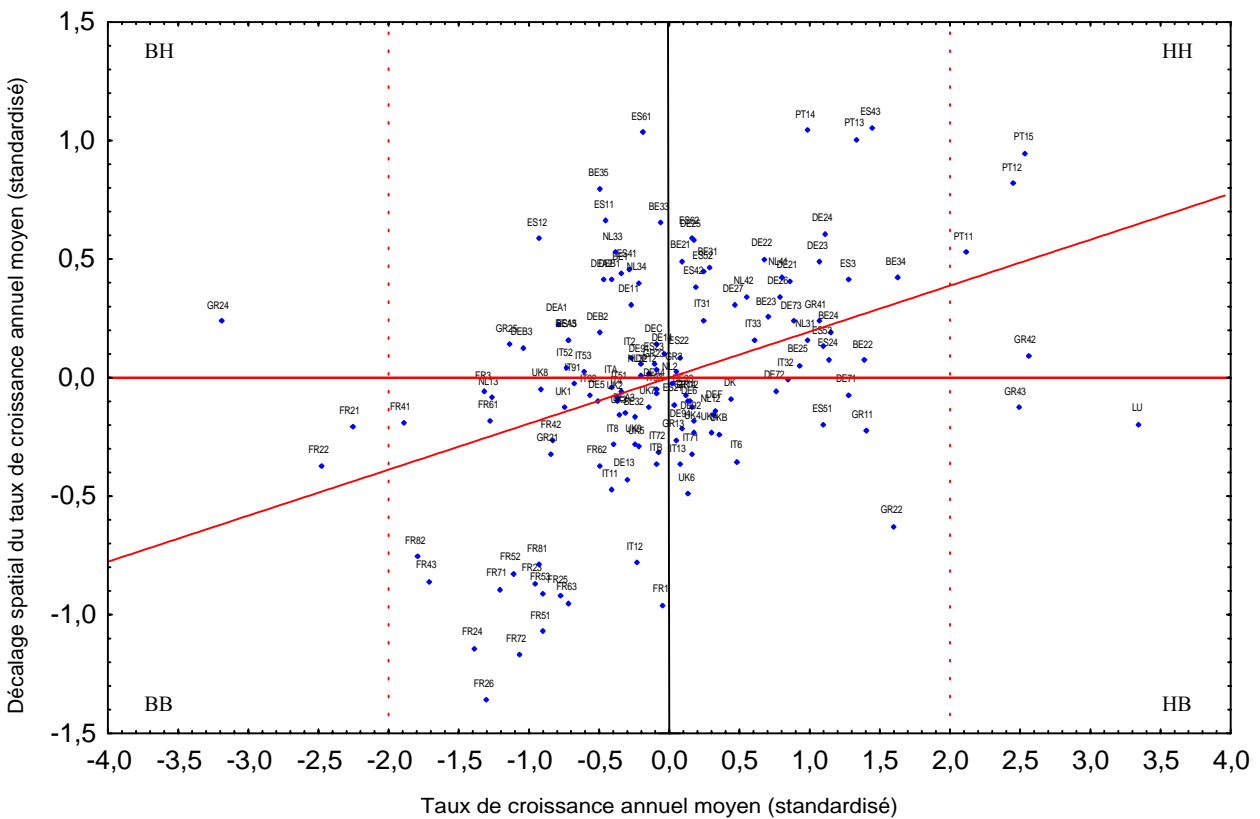


Figure 3.10b : Diagramme de Moran pour le taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en SPA

331.4 Concentrations locales, régions atypiques et statistiques LISA

Les résultats fournis par la procédure d'évaluation des statistiques LISA appliquée au taux de croissance annuel moyen sont reportés dans le tableau 3.14a et dans la figure 3.11a pour les PIB par tête en Ecus et dans le tableau 3.14b et dans la figure 3.11b pour les PIB par tête en SPA.

Concernant les taux de croissance en Ecus, il apparaît que les concentrations de type BB (16 régions françaises, toutes les régions anglaises, 6 régions grecques) et les concentrations de type HH (9 régions de l'Allemagne du sud, 6 régions espagnoles et toutes les régions portugaises) dominent toujours les schémas d'association spatiale. Plus précisément, 42,75% (respectivement 21,01%) des statistiques LISA calculées sont significatives au pseudo-niveau de 5% (respectivement au pseudo-niveau de Bonferroni à 5%). Ces résultats rejoignent ceux trouvés avec les statistiques Getis-Ord : le dynamisme des régions espagnoles, portugaises et d'Allemagne du sud contraste avec le comportement des régions françaises, anglaises et grecques du nord. Cependant, l'analyse des statistiques LISA permet en outre de détecter des régions atypiques, c'est-à-dire qui se trouvent significativement dans le quadrant HB ou dans le quadrant BH. Ainsi, les régions Galicia et Asturias (Espagne) sont significativement du type BH, alors que les régions Anatoliki Makedonia, Ionia Nisia et Kriti (Grèce) sont significativement du type HB. Bien qu'étant une observation extrême dans le diagramme de Moran, la région Sterea Ellada n'est pas une région significativement de type BH.

Les valeurs extrêmes des statistiques LISA sont représentées dans le tableau 3.13. Il apparaît que seules les régions portugaises sont dans ce cas. Ainsi, ces régions ont connu des taux de croissance très supérieurs à la moyenne en étant entourées de régions connaissant elles-mêmes des taux de croissance très supérieurs à la moyenne. Ces résultats soulignent ainsi le dynamisme de ces régions au cours de la période 1980-1995.

Concernant les taux de croissance en SPA, l'inspection des statistiques LISA significatives indique que les schémas d'associations spatiales locales restent également dominés par des concentrations de type BB (14 régions françaises) ou HH (2 régions allemandes, 2 régions espagnoles et toutes les régions portugaises). Plus précisément, 22,46% (respectivement 15,22%) des statistiques LISA calculées sont significatives au pseudo-niveau de 5% (respectivement au pseudo-niveau de Bonferroni à 5%). Certaines régions ressortent significativement dans les quadrants HB ou BH. Ainsi, les régions espagnoles Galicia, Asturias et Andalucia apparaissent être les régions avec de faibles taux de croissance entourées de régions avec de forts taux de croissance. Elles apparaissent alors moins dynamiques que leurs voisines portugaises. La seule région avec un fort taux de croissance entourée de régions avec de faibles taux de croissance est la région touristique Ionia Nisia.

Les valeurs extrêmes des statistiques LISA sont également représentées dans le tableau 3.13. Si on compare aux résultats précédents en Ecus, seules trois régions portugaises

restent des observations extrêmes alors que trois régions françaises apparaissent. La correction des pouvoirs d'achat a donc pour effet, d'une part, d'amoindrir la performance économique des régions portugaises et d'autre part, de renforcer les performances moyennes des régions françaises en termes de taux de croissance.

Ecus	SPA
Norte (1)	Extremadura (1)
Centro (1)	Centre
Lisboa (1)	Bourgogne
Alentejo (1)	Franche-Comté
Algarve (1)	Auvergne
	PACA
	Ionía Nisia (1)
	Centro (1)
	Lisboa (1)
	Algarve (1)

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 3.13 : Valeurs extrêmes des statistiques LISA ;
taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus et en SPA

Tous ces résultats sont robustes au choix du nombre de voisins dans la matrice de poids. En revanche, ils doivent être interprétés avec prudence pour deux raisons. D'une part, comme nous l'avons déjà souligné précédemment, les résultats obtenus en utilisant les données SPA posent des problèmes de comparabilité temporelle. D'autre part, que les PIB par tête soient exprimés en Ecus ou en SPA, les résultats sont dépendants de la date initiale et de la date finale. En effet, si ces dates sont à différents points du cycle d'affaire, le calcul des taux de croissance annuels moyens est potentiellement biaisé (Dowrick et Nguyen, 1989). Certains auteurs corrigent les distorsions dues aux cycles d'affaires en estimant les taux de croissance annuels moyens de chaque région à l'aide d'une régression incluant une tendance temporelle pour les PIB par tête en logarithmes de chaque région de l'échantillon (Fingleton *et al.*, 1997). Cette méthode peut cependant poser d'autres problèmes, concernant notamment la présence de racines unitaires et la fiabilité de l'inférence statistique pour les statistiques d'autocorrélation spatiale globales et locales basées sur ces taux de croissance estimés. Par conséquent, nous ne considérons pas ce problème dans le cadre de cette thèse.

Les résultats montrent la bonne performance économique de certaines régions en termes de taux de croissance alors qu'elles étaient parmi les plus pauvres en 1980. Nous étudions plus avant cette relation entre niveau de PIB initial et taux de croissance.

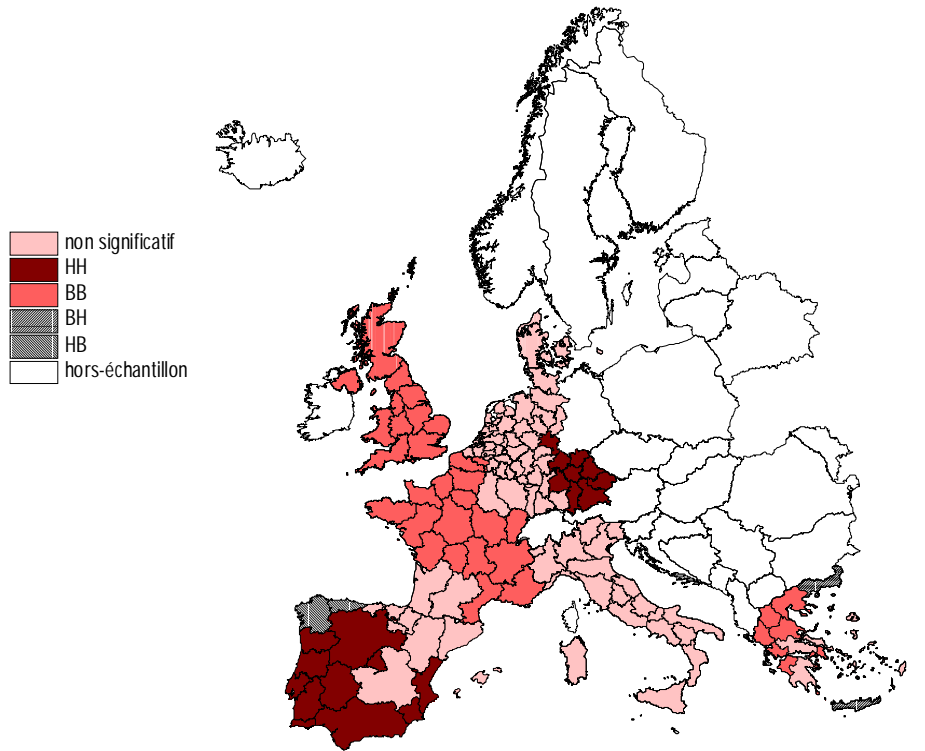


Figure 3.11a : Statistiques LISA significatives ;
taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus

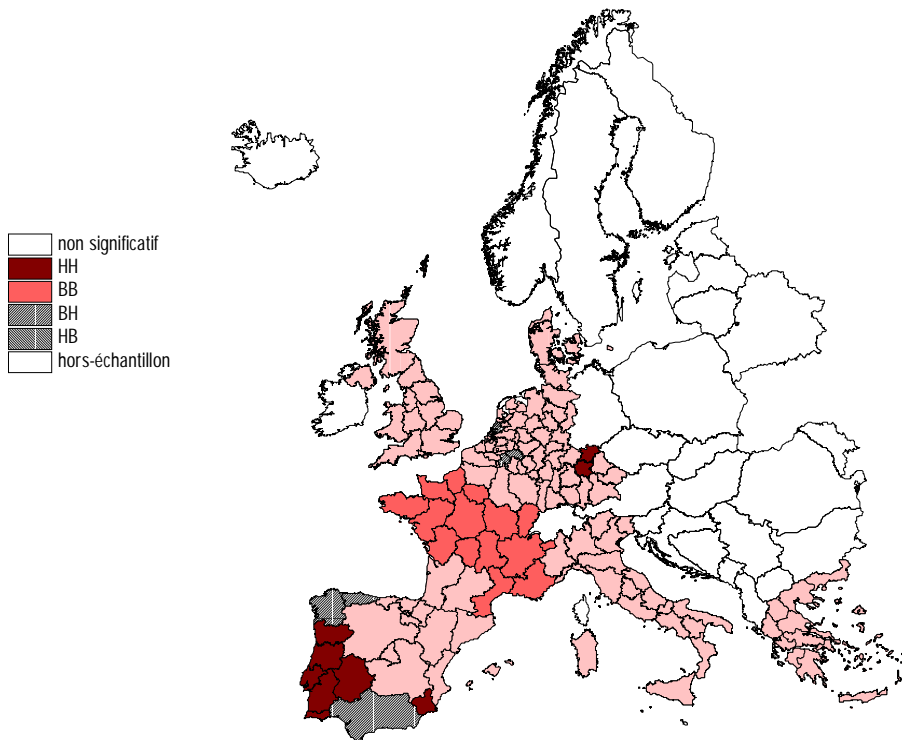


Figure 3.11b : Statistiques LISA significatives ;
taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en SPA

Code	Région	G_i	I_i	G_i	I_i	G_i	I_i	Code	Région	G_i	I_i	G_i	I_i	G_i	I_i
		1980	1980	1995	1995	croiss.	croiss.			1980	1980	1995	1995	croiss.	Croiss.
BELGIQUE								Fr51	Pays de la Loire	+*	HH*	+	HH	-	BB
Be1	Bruxelles	+	HH*	+	HH	-	BB	Fr52	Bretagne	+*	HH*	+	HH	-*	BB**
Be21	Anvers	+*	HH*	+*	HH*	-	BB	Fr53	Poitou-Charentes	+	HH*	+	HH	-*	BB*
Be22	Limburg (B)	+*	HH**	+*	HH*	-	HB	Fr61	Aquitaine	-	HB	-	HB	-	BB
Be23	Oost Vlaanderen	+*	HH*	+*	HH*	-	BB	Fr62	Midi-Pyrénées	+	HH	-	HB	-	BB
Be24	Vlaams Brabant	+*	HH*	+	HH*	-	HB	Fr63	Limousin	+*	HH*	+	HH	-*	BB*
Be25	West Vlaanderen	+*	HH*	+	HH*	-	HB	Fr71	Rhône-Alpes	+	HH	+	HH	-*	BB*
Be31	Brabant wallon	+	HH*	+	HH*	-	BB	Fr72	Auvergne	+	HH*	+	HH	-*	BB**
Be32	Hainaut	+*	HH**	+	BH*	-	BB*	Fr81	Languedoc-Roussillon	+	HH	+	BH	-*	BB*
Be33	Liège	+	HH*	+*	HH*	+	BH	Fr82	PACA	+	HH	+	HH	-	BB*
Be34	Luxembourg (B)	+*	HH*	+*	HH*	+	HH	GRECE							
Be35	Namur	+*	HH*	+*	BH*	+	BH	Gr11	Anatoliki Makedonia (1)	-**	BB**	-**	BB**	-*	HB*
ALLEMAGNE								Gr12	Kentriki Makedonia (1)	-**	BB**	-**	BB**	-*	BB*
De11	Stuttgart	+*	HH**	+**	HH**	+*	HH*	Gr13	Dytiki Makedonia (1)	-**	BB**	-**	BB**	-*	BB*
De12	Karlsruhe	+**	HH**	+**	HH**	+	HH	Gr14	Thessalia (1)	-**	BB**	-**	BB**	-*	BB*
De13	Freiburg	+**	HH**	+**	HH**	-	HB	Gr21	Ipeiros (1)	-**	BB**	-**	BB**	-*	BB*
De14	Tübingen	+**	HH**	+**	HH**	+	HH	Gr22	Ionia Nisia (1)	-**	BB**	-**	BB**	-*	HB**
De21	Oberbayern	+*	HH**	+**	HH**	+	HH*	Gr23	Dytiki Ellada (1)	-**	BB**	-**	BB**	-	HB**
De22	Niederbayern	+*	HH*	+**	HH**	+*	HH*	Gr24	Sterea Ellada (1)	-**	BB**	-**	BB**	-	BB
De23	Oberpfalz	+**	HH**	+**	HH**	+*	HH*	Gr25	Peloponnisos (1)	-**	BB**	-**	BB**	-	BB
De24	Oberfranken	+**	HH**	+**	HH**	+*	HH*	Gr3	Attiki (1)	-**	BB**	-**	BB**	-	BB*
De25	Mittelfranken	+*	HH**	+**	HH**	+	HH*	Gr41	Voreio Aigaio (1)	-**	BB**	-**	BB**	-	HB
De26	Unterfranken	+*	HH**	+**	HH**	+*	HH*	Gr42	Notio Aigaio (1)	-**	BB**	-**	BB**	-	HB
De27	Schwaben	+*	HH**	+**	HH**	+	HH*	Gr43	Kriti (1)	-**	BB**	-**	BB**	-*	HB*
De5	Bremen	+*	HH**	+*	HH**	+	HH	ITALIE							
De6	Hamburg	+*	HH**	+**	HH**	+	HH	It11	Piemonte	+	HH	+	HH	-	BB
De71	Darmstadt	+*	HH**	+**	HH**	+	HH	It12	Valle d'Aosta	+*	HH*	+	HH*	-	BB
De72	Giessen	+*	HH**	+**	HH**	+	HH	It13	Liguria	+	HH	+	HH	-	HB
De73	Kassel	+*	HH**	+**	HH**	+	HH*	It2	Lombardia	+	HH	+	HH	+	BH
De91	Braunschweig	+**	HH**	+**	HH**	+	HH	It31	Trentino – Alto Adige	+	HH	+	HH*	+	HH
De92	Hannover	+**	HH**	+**	HH**	+	HH	It32	Veneto	+	HH	+	HH	+	HH
De93	Lüneburg	+**	HH**	+**	HH**	+	HH	It33	Friuli – Venezia Giulia	+	HH	+	HH*	+	HH
De94	Weser-Ems	+*	HH**	+**	HH**	-	HB	It4	Emilia – Romagna	+	HH	+	HH	+	BH
Dea1	Düsseldorf	+*	HH*	+*	HH*	-	BB	It51	Toscana	+	HH	+	HH	+	BH
Dea2	Köln	+	HH*	+*	HH*	+	HH	It52	Umbria	-	BB	-	BB	+	BH
Dea3	Münster	+*	HH**	+*	HH**	-	HB	It53	Marche	-	BB	-	BB	+	BH
Dea4	Detmold	+*	HH**	+**	HH**	+	HH	It6	Lazio	-	BB	-	BB	-	HB
Dea5	Arnsberg	+*	HH**	+**	HH**	+	BH	It71	Abruzzo (1)	-	BB	-	BB	-	HB
Deb1	Koblenz	+*	HH**	+**	HH**	+	HH	It72	Molise (1)	-	BB	-	BB*	-	HB
Deb2	Trier	+*	HH*	+*	HH*	+	HH	It8	Campania (1)	-*	BB*	-*	BB*	-	BB
Deb3	Rhein Hessen-Pfalz	+*	HH**	+**	HH**	+	BH	It91	Puglia (1)	-**	BB**	-**	BB**	-	BB
Dec	Saarland	+*	HH**	+**	HH**	+	HH	It92	Basilicata (1)	-*	BB*	-*	BB**	+	BH
Def	Schleswig-Holstein	+**	HH**	+**	HH**	+	HH	It93	Calabria (1)	-**	BB**	-**	BB**	-	HB
Dk	DANEMARK	+*	HH**	+**	HH**	+	HH	Ita	Sicilia (1)	-**	BB**	-**	BB**	+	BH
ESPAGNE								Itb	Sardegna (1)	-	BB	-	BB	-	HB
Es11	Galicia (1)	-**	BB**	-**	BB**	+**	BH**	Lu	LUXEMBOURG	+	HH*	+	HH*	-	HB
Es12	Asturias (1)	-**	BB**	-**	BB**	+**	BH**	PAYS-BAS							
Es13	Cantabria	-*	BB*	-*	BB*	+	BH	NI12	Friesland	+*	HH**	+*	HH**	-	BB
Es21	Pais Vasco	-	BB	-	BB*	+	HH	NI13	Drenthe	+*	HH**	+*	HH**	-	BB
Es22	Navarra	-	BB	-	BB	+	HH	NI2	Oost Nederland	+*	HH*	+*	HH*	-	BB
Es23	La Rioja	-*	BB*	-	BB*	+	HH	NI31	Utrecht	+*	HH**	+*	HH**	-	HB
Es24	Aragon	-	BB	-	BB	+	HH	NI32	Noord-Holland	+*	HH**	+*	HH**	-	BB
Es3	Madrid	-**	BB**	-**	BB**	+*	HH*	NI33	Zuid-Holland	+*	HH*	+*	HH*	-	BB
Es41	Castilla-Leon (1)	-**	BB**	-**	BB**	+*	HH*	NI34	Zeeland	+*	HH**	+*	HH**	-	BB
Es42	Castilla-la Mancha (1)	-**	BB**	-**	BB**	+	HH	NI41	Noord-Brabant	+*	HH*	+*	HH*	-	HB
Es43	Extremadura (1)	-**	BB**	-**	BB**	+**	HH**	NI42	Limburg (NL)	+*	HH*	+*	HH*	-	BB
Es51	Cataluna	-	BB	-	BB	+	HB	PORTUGAL							
Es52	Valenciana (1)	-*	BB*	-*	BB*	+	HH*	Pt11	Norte (1)	-**	BB**	-**	BB**	+**	HH**
Es53	Islas Baleares	-**	BB*	-**	BB*	+	HH	Pt12	Centro (1)	-**	BB**	-**	BB**	+**	HH**
Es61	Andalucia (1)	-**	BB**	-**	BB**	+**	HH**	Pt13	Lisboa e vale do Tejo (1)	-**	BB**	-**	BB**	+**	HH**
Es62	Murcia (1)	-**	BB**	-**	BB**	+*	HH*	Pt14	Alentejo (1)	-**	BB**	-**	BB**	+**	HH**
FRANCE								Pt15	Algarve (1)	-**	BB**	-**	BB**	+**	HH**
Fr1	Ile-de-France	+	HH*	+	HH	-*	BB**	ROYAUME-UNI							
Fr21	Champagne-Ardenne	+*	HH*	+	HH*	-	BB	Uk1	North	+	HH	-	BB	-**	BB**
Fr22	Picardie	+*	HH**	+*	HH*	-	BB*	Uk2	Yorkshire and Humberside	+	HH	-	BB	-**	BB**
Fr23	Haute-Normandie	+*	HH**	+	HH	-*	BB**	Uk3	East Midlands	+	HH	-	BB	-**	BB**
Fr24	Centre	+*	HH*	+	HH	-*	BB**	Uk4	East Anglia	+	HH	+	BH	-**	BB**
Fr25	Basse-Normandie	+*	HH*	+	HH	-*	BB**	Uk5	South East	+	HH	+	HH	-**	BB**
Fr26	Bourgogne	+*	HH**	+	HH	-**	BB**	Uk6	South West	+	HH	-	BB	-**	BB**
Fr3	Nord-Pas-De-Calais	+*	HH*	+	HH*	-	BB*	Uk7	West Midlands	+	HH	-	BB	-**	BB**
Fr41	Lorraine	+*	HH**	+*	HH**	+	BH	Uk8	North West	+	HH	-	BB	-**	BB**
Fr42	Alsace	+*	HH**	+**	HH**	+	BH	Uk9	Wales	+	BH	-	BB	-**	BB**
Fr43	Franche-Comté	+*	HH**	+	HH*	-	BB*	Uka	Scotland	+	HH	-	BB	-**	BB**
								Ukb	Northern Ireland (1)	+	BH	-	BB	-**	BB**

Notes : (*) significatif à 5%, basée sur une approximation normale pour Getis-Ord et sur 10 000 permutations pour LISA
 (**) significatif à 5% Sidak pour Getis-Ord et à 5% Bonferroni pour LISA

Tableau 3.14a : Statistiques Getis-Ord et LISA pour la matrice $W(10)$;

PIB par tête en logarithmes et en Ecus pour 1980, 1995 et pour le taux de croissance annuel moyen

Code	Région	G _i	I _i	G _i	I _i	G _i	I _i	Code	Région	G _i	I _i	G _i	I _i	G _i	I _i
		1980	1980	1995	1995					croiss.	croiss.	1980	1980		
	BELGIQUE							Fr51	Pays de la Loire	-**	HH*	+	HH	-**	BB**
Be1	Bruxelles	+	HH	+	HH	+	BH	Fr52	Bretagne	+	HH*	+	BH	-*	BB**
Be21	Anvers	+	HH	+	HH*	+	HH	Fr53	Poitou-Charentes	+	HH	+	BH	-**	BB**
Be22	Limburg (B)	+	BH*	+	HH*	+	HH	Fr61	Aquitaine	-	HB	-	BB	-	BB
Be23	Oost Vlaanderen	+	HH	+	HH*	+	HH	Fr62	Midi-Pyrénées	+	HH	-	BB	-	BB
Be24	Vlaams Brabant	+	BH	+	HH	+	HH	Fr63	Limousin	+	BH	+	BH	-**	BB**
Be25	West Vlaanderen	+	HH	+	HH*	+	HH	Fr71	Rhône-Alpes	+	HH*	+	HH	-**	BB**
Be31	Brabant wallon	+	BH	+	BH*	+	HH	Fr72	Auvergne	+	HH	-	BB	-**	BB**
Be32	Hainaut	+	BH	+	BH	-	BB	Fr78	Languedoc-Roussillon	+	BH	+	BH	-*	BB**
Be33	Liège	+	HH	+	HH*	+	BH*	Fr82	PACA	+	HH*	+	HH	-*	BB**
Be34	Luxembourg (B)	+	BH	+	HH*	+	HH		GRECE						
Be35	Namur	+	BH	+	BH*	+	BH*	Gr11	Anatoliki Makedonia (1)	-**	BB**	-**	BB**	-	HB
	ALLEMAGNE							Gr12	Kentriki Makedonia (1)	-**	BB**	-**	BB**	-	HB
De11	Stuttgart	+	HH*	+	HH**	+	BH	Gr13	Dytiki Makedonia (1)	-**	BB**	-**	BB**	-	HB
De12	Karlsruhe	+	HH**	+	HH**	+	BH	Gr14	Thessalia (1)	-**	BB**	-**	BB**	-	HB
De13	Freiburg	+	HH**	+	HH**	-	BB	Gr21	Ipeiros (1)	-**	BB**	-**	BB**	-	BB
De14	Tübingen	+	HH**	+	HH**	+	BH	Gr22	Ionia Nisia (1)	-**	BB**	-**	BB**	-*	HB*
De21	Oberbayern	+	HH*	+	HH**	+	HH	Gr23	Dytiki Ellada (1)	-**	BB**	-**	BB**	+	BH
De22	Niederbayern	+	BH*	+	HH**	+	HH	Gr24	Stereia Ellada (1)	-**	BB**	-**	BB**	+	BH
De23	Oberpfalz	+	BH**	+	HH**	+	HH	Gr25	Peloponnisos (1)	-**	BB**	-**	BB**	+	BH
De24	Oberfranken	+	HH*	+	HH**	+	HH*	Gr3	Attiki (1)	-**	BB**	-**	BB**	+	HH
De25	Mittelfranken	+	HH**	+	HH**	+	HH*	Gr41	Voreio Aigaio (1)	-**	BB**	-**	BB**	+	HH
De26	Unterfranken	+	HH*	+	HH**	+	HH	Gr42	Notio Aigaio (1)	-**	BB**	-**	BB**	+	HH
De27	Schwaben	+	HH**	+	HH**	+	HH	Gr43	Kriti (1)	-**	BB**	-**	BB**	-	HB
De5	Bremen	+	HH*	+	HH*	-	BB		ITALIE						
De6	Hamburg	+	HH*	+	HH*	-	HB	It11	Piemonte	+	HH**	+	HH**	-	BB
De71	Darmstadt	+	HH*	+	HH*	-	HB	It12	Valle d'Aosta	+	HH**	+	HH*	-*	BB**
De72	Giessen	+	HH*	+	HH*	-	HB	It13	Liguria	+	HH**	+	HH**	-	HB
De73	Kassel	+	HH*	+	HH*	+	HH	It2	Lombardia	+	HH**	+	HH**	+	BH
De91	Braunschweig	+	HH**	+	HH**	+	BH	It31	Trentino – Alto Adige	+	HH**	+	HH**	+	HH
De92	Hannover	+	HH**	+	HH*	-	HB	It32	Veneto	+	HH**	+	HH**	+	HB
De93	Lüneburg	+	BH**	+	BH**	-	HB	It33	Friuli – Venezia Giulia	+	HH**	+	HH**	+	HH
De94	Weser-Ems	+	HH*	+	HH*	-	HB	It4	Emilia – Romagna	+	HH**	+	HH**	-	BB
Dea1	Düsseldorf	+	HH	+	HH*	+	BH	It51	Toscana	+	HH**	+	HH**	-	BB
Dea2	Köln	+	HH	+	HH*	+	BH	It52	Umbria	+	HH	+	HH	+	BH
Dea3	Münster	+	HH*	+	HH*	-	BB	It53	Marche	+	HH	+	HH	+	BH
Dea4	Detmold	+	HH*	+	HH*	-	BB	It6	Lazio	+	HH	-	HB	-	HB
Dea5	Arnsberg	+	HH*	+	HH*	+	BH	It71	Abruzzo (1)	+	BH	+	BH	-	HB
Deb1	Koblenz	+	HH*	+	BH*	+	BH	It72	Molise (1)	-	BB	-	BB	-	BB
Deb2	Trier	+	HH	+	BH	+	BH	It8	Campania (1)	-	BB	-	BB	-	BB
Deb3	Rheinhessen-Pfalz	+	HH**	+	HH**	+	BH	It91	Puglia (1)	-**	BB**	-**	BB**	-	BB
Dec	Saarland	+	HH*	+	HH*	+	BH	It92	Basilicata (1)	-**	BB*	-*	BB*	-	BB
Def	Schleswig-Holstein	+	HH**	+	HH*	-	HB	It93	Calabria (1)	-**	BB**	-**	BB**	-	BB
Dk	DANEMARK	+	HH*	+	HH*	-	HB	Ita	Sicilia (1)	-*	BB*	-*	BB*	-	BB
	ESPAGNE							Itb	Sardegna (1)	+	BH	-	BB	-	BB
Es11	Galicie (1)	-**	BB**	-*	BB**	+	BH*	Lu	LUXEMBOURG	+	HH	+	HH	-	HB
Es12	Asturies (1)	-**	BB**	-**	BB**	+	BH*		PAYS-BAS						
Es13	Cantabria	-	BB*	-	BB*	+	BH	Ni12	Friesland	+	BH*	+	BH*	-	HB
Es21	Pais Vasco	-	BB	-	BB	-	HB	Ni13	Drenthe	+	HH*	+	BH*	-	BB
Es22	Navarra	-	BB	-	HB	+	HH	Ni2	Oost Nederland	+	HH	+	HH	-	HB
Es23	La Rioja	-	BB	-	BB	+	BH	Ni31	Utrecht	+	HH	+	HH*	+	HH
Es24	Aragon	-	BB	-	BB	+	HH	Ni32	Noord-Holland	+	HH	+	HH*	+	BH
Es3	Madrid	-**	BB**	-*	HB*	+	HH	Ni33	Zuid-Holland	+	HH	+	HH*	+	BH*
Es41	Castilla-Leon (1)	-*	BB*	-*	BB*	+	BH	Ni34	Zeeland	+	HH*	+	HH*	+	BH
Es42	Castilla-la Mancha (1)	-*	BB*	-*	BB*	+	HH	Ni41	Noord-Brabant	+	HH*	+	HH*	+	HH
Es43	Extremadura (1)	-**	BB**	-**	BB**	+	HH**	Ni42	Limburg (NL)	+	BH	+	HH*	+	HH
Es51	Cataluna	-	BB	-	HB	-	HB		PORTUGAL						
Es52	Valenciana (1)	-*	BB*	-	BB*	+	HH	Pt11	Norte (1)	-**	BB**	-**	BB**	+	HH*
Es53	Islas Baleares	-	BB	-	HB	+	HH	Pt12	Centro (1)	-**	BB**	-**	BB**	+	HH**
Es61	Andalucia (1)	-**	BB**	-**	BB**	+	BH**	Pt13	Lisboa e vale do Tejo (1)	-**	BB**	-**	BB**	+	HH**
Es62	Murcia (1)	-**	BB**	-*	BB*	+	HH*	Pt14	Alentejo (1)	-**	BB**	-**	BB**	+	HH**
	FRANCE							Pt15	Algarve (1)	-**	BB**	-**	BB**	+	HH**
Fr1	Ile-de-France	+	HH	-	HB	-**	BB**		ROYAUME-UNI						
Fr21	Champagne-Ardenne	+	HH	+	HH	-	BB	Uk1	North	-	BB	-	BB	-	BB
Fr22	Picardie	+	HH*	+	BH*	-	BB	Uk2	Yorkshire and Humberside	-	BB	-	BB	-	BB
Fr23	Haute-Normandie	+	HH*	+	HH	-**	BB**	Uk3	East Midlands	+	HH	+	BH	-	BB
Fr24	Centre	+	HH*	+	HH	-**	BB**	Uk4	East Anglia	+	HH	+	HH	-	HB
Fr25	Basse-Normandie	+	HH*	+	HH	-**	BB**	Uk5	South East	+	HH	+	HH	-	BB
Fr26	Bourgogne	+	HH*	+	BH	-**	BB**	Uk6	South West	+	BH	-	BB	-	HB
Fr3	Nord-Pas-De-Calais	+	HH	+	BH	-	BB	Uk7	West Midlands	-	BB	-	BB	-	BB
Fr41	Lorraine	+	HH*	+	BH*	-	BB	Uk8	North West	-	HB	-	BB	-	BB
Fr42	Alsace	+	HH**	+	HH*	-	BB	Uk9	Wales	+	BH	-	BB	-	BB
Fr43	Franche-Comté	+	HH**	+	HH*	-**	BB**	Uka	Scotland	-	HB	-	HB	-	HB
								Ukb	Northern Ireland (1)	-	BH	-	BB	-	HB

Notes : (*) significatif à 5%, basée sur une approximation normale pour Getis-Ord et sur 10 000 permutations pour LISA
 (**) significatif à 5% Sidák pour Getis-Ord et à 5% Bonferroni pour LISA

Tableau 3.14b : Statistiques Getis-Ord et LISA pour la matrice $W(10)$;

PIB par tête en logarithmes et en SPA pour 1980, 1995 et pour le taux de croissance annuel moyen

332 Analyse exploratoire de la relation entre taux de croissance et PIB par tête initial

L'étude de la relation inverse entre le PIB par tête en logarithmes en 1980 et le taux de croissance annuel moyen permet d'étudier dans quelle mesure le dynamisme des régions sur la période 1980-1995 dépend de leur point de départ (tableaux 3.14a et 3.14b).

Concernant les PIB par tête en Ecus, il apparaît ainsi que seules 42,75% des régions qui étaient dans un certain quadrant du diagramme de Moran pour les PIB par tête en 1980 (HH, BB, HB ou BH) se retrouvent dans le quadrant opposé pour leur taux de croissance annuel moyen (BB, HH, BH ou HB). Pour les PIB par tête en SPA, ce chiffre est encore plus réduit et est égal à 36,33%. Lorsqu'on considère les statistiques de Getis-Ord, cette relation inverse est légèrement plus importante. Ainsi, pour les PIB par tête en Ecus, 57,25% des statistiques Getis-Ord, qui étaient significativement positives (respectivement négatives), étaient positives (respectivement négatives) pour les taux de croissance annuels moyens. Ce chiffre se monte à 55,8% pour les PIB par tête en SPA.

Cette caractéristique globale masque des comportements très différents parmi les régions européennes. Ainsi, toutes les régions du Portugal et certaines régions espagnoles possédaient en 1980 de faibles PIB par tête et étaient entourées de régions possédant également de faibles PIB par tête (concentration de type BB). Pourtant, leur taux de croissance entre 1980 et 1995 est, comme pour leurs voisines, plus élevé que la moyenne européenne (concentration de type HH). Les indicateurs d'autocorrélation spatiale locale permettent ainsi de faire ressortir le caractère dynamique de ces régions, dont les performances économiques au sein du groupe des régions du Sud ont souvent été soulignées, comme par exemple dans les différents rapports de la Commission Européenne (1991, 1994, 1996, 1999). En revanche, presque toutes les régions françaises et anglaises sont caractérisées par une configuration du type HH pour les PIB par tête en 1980 et par une configuration du type BB pour les taux de croissance annuels moyens.

D'autres tendances entre les schémas d'association spatiale peuvent être soulignées. D'une part, au sein du groupe des régions du Sud, quelques régions pauvres d'Espagne, d'Italie et de Grèce ne parviennent pas à décoller, tout comme leurs voisines (configurations de type BB pour les PIB par tête initiaux et pour les taux de croissance) ou en dépit du dynamisme de leurs voisines (configuration de type BB pour les PIB par tête initiaux et de type BH pour les taux de croissance). C'est le cas par exemple des régions espagnoles Galicia, Asturias, Cantabria et des régions Umbria et Marche en Italie. Par conséquent, ces régions montrent de forts signes de retard de développement. En outre, la situation de certaines régions grecques a continué à se détériorer sur la période par rapport à celui du reste de la Communauté. En effet, nombre d'entre elles étaient pauvres entourées de régions pauvres en début de période et ont connu des taux de croissance significativement inférieurs à la moyenne, tout comme leurs voisines (Kentriki Makedonia, Dytiki Makedonia, Attiki,...).

Dans le groupe des régions Nord, les dynamismes relatifs des régions vis-à-vis de leurs voisines apparaissent plus fluctuants et aucune configuration de types BH ou HB sur les taux de croissance ne semble dominante. Soulignons cependant la très bonne performance de certaines régions allemandes qui étaient caractérisées par une configuration de type HH à la fois pour les PIB par tête en 1980 et pour leur taux de croissance annuels moyens (Oberfranken, Mittelfranken...).

Conclusion de section

Cette section a mis en évidence les disparités en termes de taux de croissance entre les régions européennes et a montré que la croissance économique reste géographiquement concentrée. Ainsi, certaines régions bénéficient d'un environnement économique favorable et d'un cercle vertueux de croissance. Ces régions se répartissent en deux groupes, d'une part les régions périphériques de l'Union Européenne situées au Portugal et en Espagne et d'autre part, les régions de l'Allemagne du Sud. En revanche, les régions françaises et anglaises ont connu des taux de croissance inférieurs à la moyenne de l'échantillon alors qu'elles avaient démarré avec des niveaux élevés de PIB par tête. Ces deux éléments pourraient suggérer un phénomène de rattrapage de ces régions par des régions plus pauvres. Globalement, il apparaît néanmoins que les disparités économiques et spatiales n'ont que très faiblement diminué au cours de la période, malgré les performances nettement meilleures, par rapport à la moyenne communautaire, réalisées par certaines régions appartenant, dans notre échantillon, aux pays de la cohésion (Espagne, Portugal et Grèce). La très bonne performance des régions allemandes laisse en outre présager une divergence de ces régions les plus riches avec le reste de l'Europe Communautaire.

Conclusion du chapitre 3

L'étude de la distribution spatiale des PIB par tête en Europe sur la période 1980-1995 à l'aide de l'analyse exploratoire des données spatiales (ESDA) souligne l'importance de la localisation géographique dans les processus de croissance régionale. Malgré la baisse des coûts de transport, le progrès des technologies de la communication et la restructuration économique de la dernière décennie, les disparités territoriales traditionnelles persistent : les activités économiques ne sont pas réparties également dans l'espace et la croissance reste géographiquement concentrée.

L'ESDA que nous avons menée apparaît être un outil puissant pour révéler finement les caractéristiques du développement économique de chaque région en relation avec celui de son environnement géographique. Elle doit être considérée comme une première étape descriptive avant de suggérer des facteurs expliquant les schémas spatiaux détectés et estimer des modèles économétriques plus sophistiqués. Nous avons ainsi détecté la présence d'autocorrélation spatiale et d'hétérogénéité spatiale dans la distribution des PIB par tête régionaux bien que nous n'ayons pas tenté dans ce chapitre de détailler les facteurs explicatifs de la performance économique propres à chaque région. Nous nous concentrons plutôt sur les implications de ces résultats sur l'analyse économétrique.

Du point de vue de l'économétrie appliquée, les résultats de l'ESDA ont deux implications importantes pour l'estimation des modèles de β -convergence standards (Barro et Sala-I-Martin, 1991, 1995 ; Armstrong, 1995a). Habituellement, ces modèles sont estimés à l'aide des MCO. Cependant, les estimateurs obtenus grâce à cette méthode sont inefficients, l'inférence statistique n'est pas fiable lorsque les hypothèses statistiques d'absence de corrélation et d'homoscédasticité ne sont pas respectées.

Premièrement, l'ESDA révèle une autocorrélation spatiale globale positive, qui est persistante sur toute la période : les régions possédant des PIB par tête élevés (respectivement faibles) sont et restent localisées près d'autres régions possédant des PIB par tête élevés (respectivement faibles). Ceci est également vrai pour les taux de croissance des PIB par tête sur la période 1980-1995. En conséquence, l'autocorrélation spatiale des termes d'erreurs du modèle de β -convergence devrait être systématiquement testée. Si elle est détectée, une spécification spatiale appropriée (variable endogène décalée ou autocorrélation des erreurs) devrait être estimée à l'aide des outils économétriques adéquats pour permettre une inférence statistique fiable.

Deuxièmement, les statistiques de Getis et Ord indiquent la concentration de valeurs faibles ou élevées, persistantes sur toute la période, ce qui confirme la polarisation Nord-Sud des régions européennes. Les diagrammes de Moran et les statistiques LISA renforcent cette conclusion et révèlent en outre l'existence de régions atypiques, principalement des régions pauvres entourées de régions riches (type BH). Ces résultats soulignent une hétérogénéité

spatiale conjointe à l'autocorrélation spatiale globale positive et peuvent indiquer la co-existence de deux régimes spatiaux distincts, qui peuvent alors être interprétés comme des clubs de convergence spatiaux. Ainsi, le processus de convergence, s'il existe, pourrait être différent selon les régimes. En outre, les effets spatiaux peuvent également être différents dans chacun de ces régimes. Du point de vue de l'économétrie appliquée, ce résultat suggère que l'existence potentielle de régimes spatiaux distincts doit être considérée dans l'estimation des modèles de β -convergence. En d'autres termes, l'hétérogénéité spatiale doit être testée sous la forme d'une instabilité structurelle (coefficients variables selon le régime) et/ou d'une hétéroscédasticité des erreurs. La méthodologie empirique suggérée par Durlauf et Johnson (1995) doit donc être étendue dans le but de prendre en compte les effets spatiaux.

Ces différents aspects sont étudiés dans le chapitre suivant.

Annexe 3.1

Les régions de l'échantillon

Code	Région	Code	Région	Code	Région
	BELGIQUE				ITALIE
Be1	Bruxelles	Es22	Navarra	It11	Piemonte
Be21	Anvers	Es23	La Rioja	It12	Valle d'Aosta
Be22	Limburg (B)	Es24	Aragon	It13	Liguria
Be23	Oost Vlaanderen	Es3	Madrid	It2	Lombardia
Be24	Vlaams Brabant	Es41	Castilla-Leon (1)	It31	Trentino – Alto Adige
Be25	West Vlaanderen	Es42	Castilla-la Mancha (1)	It32	Veneto
Be31	Brabant Wallon	Es43	Extremadura (1)	It33	Friuli – Venezia Giulia
Be32	Hainaut	Es51	Cataluna	It4	Emilia – Romagna
Be33	Liège	Es52	Valenciana (1)	It51	Toscana
Be34	Luxembourg (B)	Es53	Islas Baleares	It52	Umbria
Be35	Namur	Es61	Andalucia (1)	It53	Marche
	ALLEMAGNE	Es62	Murcia (1)	It6	Lazio
De11	Stuttgart		FRANCE	It71	Abruzzo (1)
De12	Karlsruhe	Fr1	Ile-de-France	It72	Molise (1)
De13	Freiburg	Fr21	Champagne-Ardenne	It8	Campania (1)
De14	Tübingen	Fr22	Picardie	It91	Puglia (1)
De21	Oberbayern	Fr23	Haute-Normandie	It92	Basilicata (1)
De22	Niederbayern	Fr24	Centre	It93	Calabria (1)
De23	Oberpfalz	Fr25	Basse-Normandie	Ita	Sicilia (1)
De24	Oberfranken	Fr26	Bourgogne	Itb	Sardegna (1)
De25	Mittelfranken	Fr3	Nord-Pas-De-Calais	Lu	LUXEMBOURG
De26	Unterfranken	Fr41	Lorraine		PAYS-BAS
De27	Schwaben	Fr42	Alsace	NI12	Friesland
De5	Bremen	Fr43	Franche-Comté	NI13	Drenthe
De6	Hamburg	Fr51	Pays de la Loire	NI2	Oost Nederland
De71	Darmstadt	Fr52	Bretagne	NI31	Utrecht
De72	Giessen	Fr53	Poitou-Charentes	NI32	Noord-Holland
De73	Kassel	Fr61	Aquitaine	NI33	Zuid-Holland
De91	Braunschweig	Fr62	Midi-Pyrénées	NI34	Zeeland
De92	Hannover	Fr63	Limousin	NI41	Noord-Brabant
De93	Lüneburg	Fr71	Rhône-Alpes	NI42	Limburg (NL)
De94	Weser-Ems	Fr72	Auvergne		PORTUGAL
Dea1	Düsseldorf	Fr81	Languedoc-Roussillon	Pt11	Norte (1)
Dea2	Köln	Fr82	PACA	Pt12	Centro (1)
Dea3	Münster		GRECE	Pt13	Lisboa e vale do Tejo (1)
Dea4	Detmold	Gr11	Anatoliki Makedonia (1)	Pt14	Alentejo (1)
Dea5	Arnsberg	Gr12	Kentriki Makedonia (1)	Pt15	Algarve (1)
Deb1	Koblenz	Gr13	Dytiki Makedonia (1)		ROYAUME-UNI
Deb2	Trier	Gr14	Thessalia (1)	Uk1	North
Deb3	Rheinhessen-Pfalz	Gr21	Ipeiros (1)	Uk2	Yorkshire and Humberside
Dec	Saarland	Gr22	Ionia Nisia (1)	Uk3	East Midlands
Def	Schleswig-Holstein	Gr23	Dytiki Ellada (1)	Uk4	East Anglia
Dk	DANEMARK	Gr24	Stereia Ellada (1)	Uk5	South East
	ESPAGNE	Gr25	Peloponnisos (1)	Uk6	South West
Es11	Galicie (1)	Gr3	Attiki (1)	Uk7	West Midlands
Es12	Asturias (1)	Gr41	Voreio Aigaio (1)	Uk8	North West
Es13	Cantabria	Gr42	Notio Aigaio (1)	Uk9	Wales
Es21	Pais Vasco	Gr43	Kriti (1)	Uka	Scotland
				Ukb	Northern Ireland (1)

Tableau 1 : Liste des régions et des codes

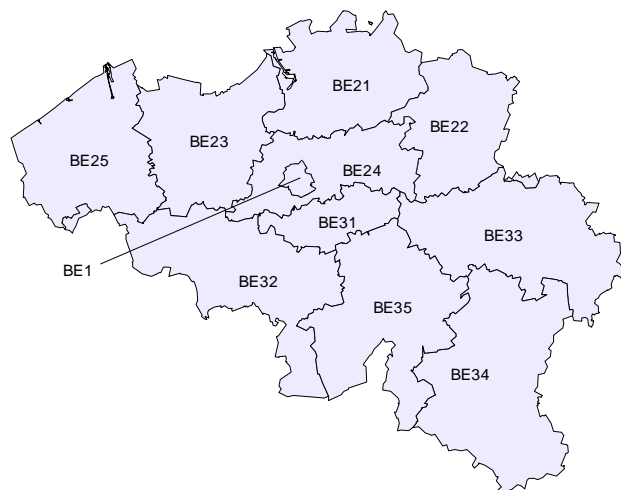


Figure 1 : Les régions de Belgique

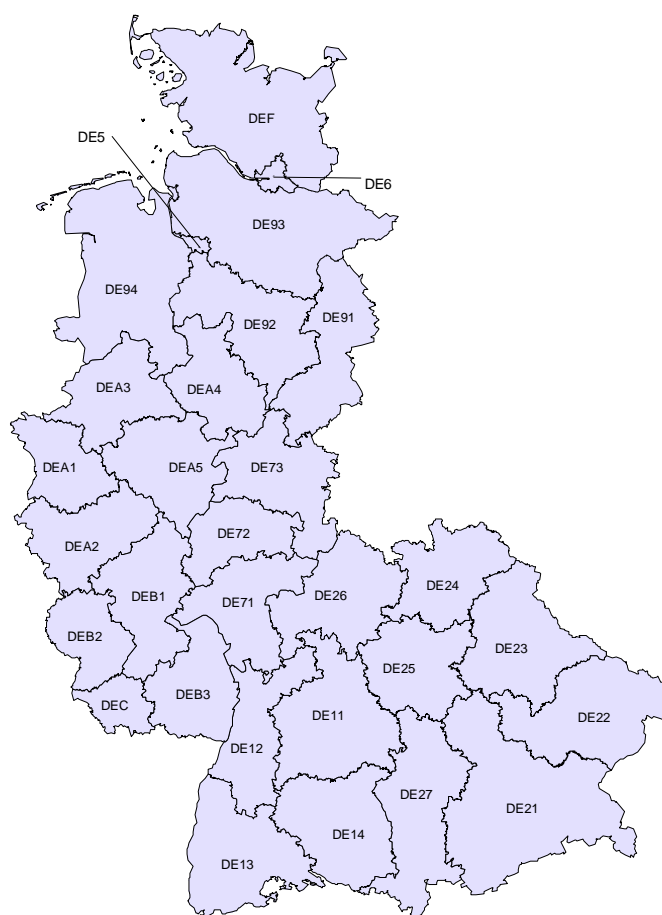


Figure 2 : Les régions d'Allemagne

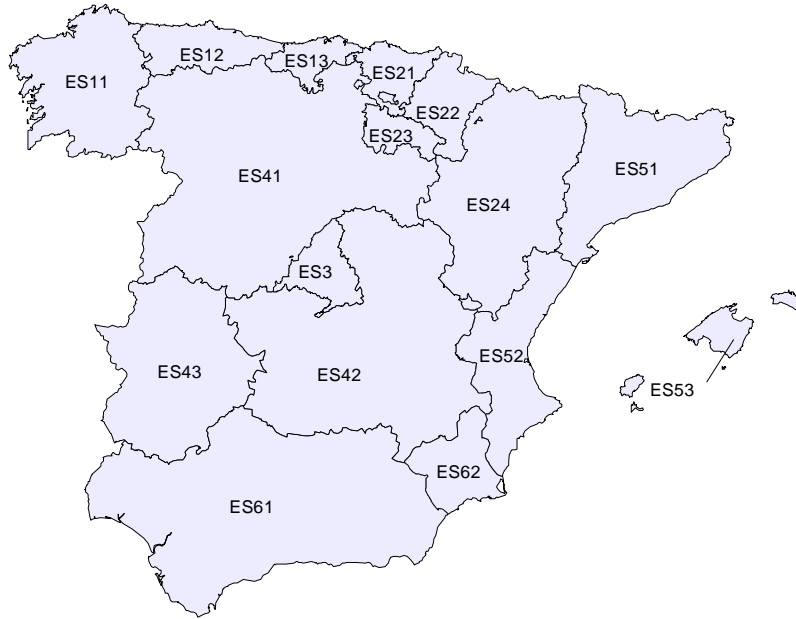


Figure 3 : Les régions d'Espagne

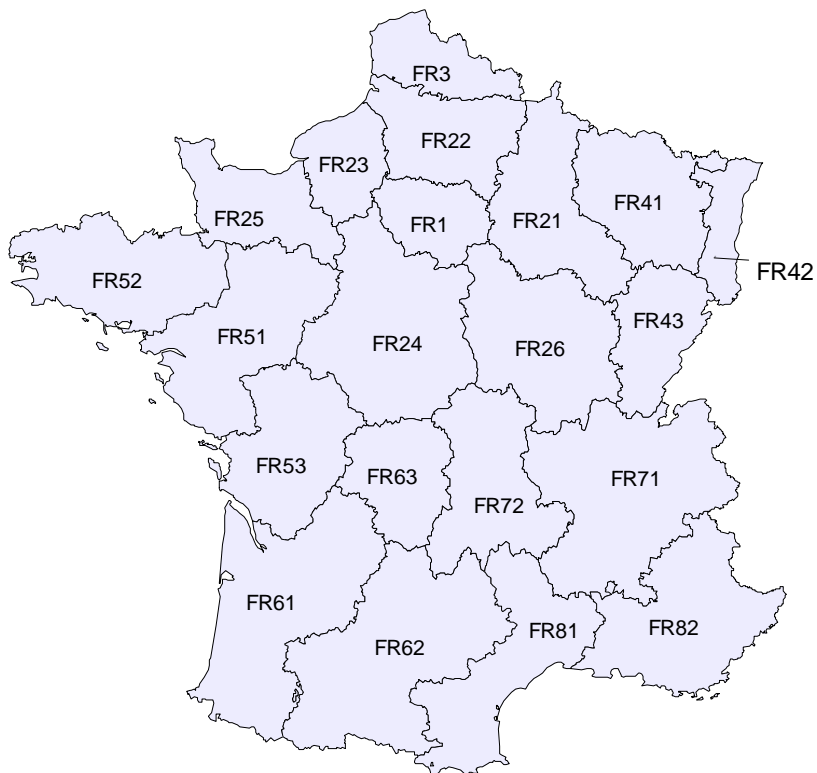


Figure 4 : Les régions de France



Figure 5 : Les régions de Grèce

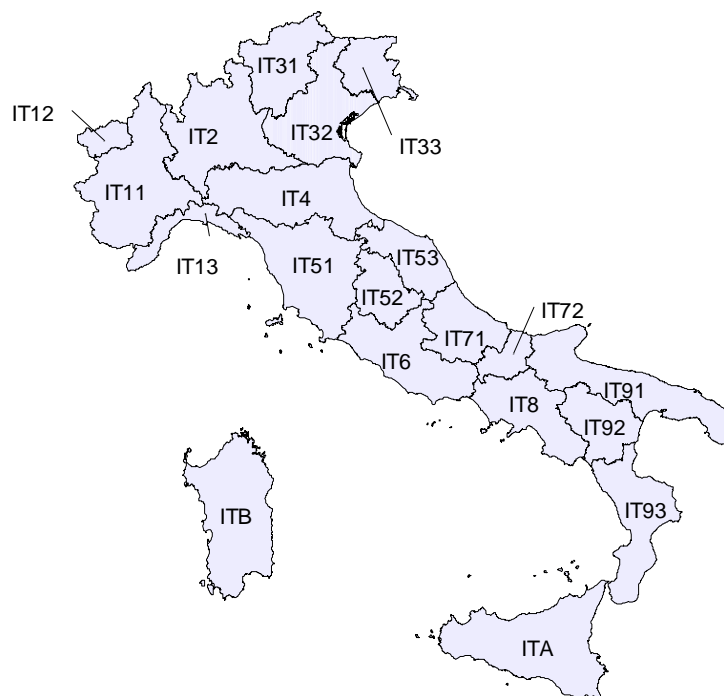


Figure 6 : Les régions d'Italie



Figure 7 : Les régions des Pays-Bas

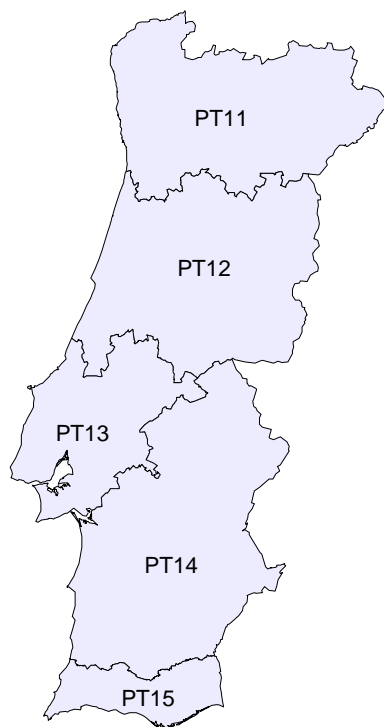


Figure 8 : Les régions du Portugal



Figure 9 : Les régions du Royaume-Uni

Annexe 3.2

Résultats pour les sections 32 et 33 avec les matrices des distances inverses ; PIB par tête en Ecus

Année	W(1)			W(2)			W(3)		
	I de Moran	Ecart-type	Stand.	I de Moran	Ecart-type	Stand.	I de Moran	Ecart-type	Stand.
1980	0,812	0,040333	20,310	0,775	0,034896	22,377	0,712	0,032944	21,842
1981	0,788	0,040367	19,689	0,749	0,034959	21,638	0,686	0,033003	21,016
1982	0,774	0,040351	19,359	0,734	0,034933	21,211	0,670	0,032979	20,545
1983	0,803	0,040453	20,030	0,762	0,035017	21,965	0,696	0,033058	21,267
1984	0,788	0,040452	19,668	0,747	0,035010	21,548	0,684	0,033052	20,906
1985	0,795	0,040419	19,851	0,755	0,034993	21,782	0,691	0,033030	21,136
1986	0,813	0,040418	20,270	0,771	0,034966	22,212	0,708	0,033000	21,674
1987	0,819	0,040430	20,396	0,777	0,034969	22,380	0,715	0,033002	21,892
1988	0,804	0,040494	20,041	0,766	0,035033	22,062	0,766	0,035033	22,062
1989	0,791	0,040403	19,737	0,751	0,034965	21,681	0,695	0,032997	21,276
1990	0,790	0,040359	19,755	0,750	0,034904	21,685	0,694	0,032937	21,293
1991	0,782	0,040321	19,564	0,741	0,034864	21,467	0,688	0,032899	21,127
1992	0,791	0,040280	19,823	0,751	0,034797	21,780	0,698	0,032831	21,464
1993	0,809	0,040344	20,226	0,771	0,034835	22,328	0,715	0,032866	21,978
1994	0,816	0,040367	20,390	0,778	0,034864	22,531	0,721	0,032894	22,140
1995	0,817	0,040346	20,421	0,777	0,034819	22,510	0,720	0,032848	22,145

Note : L'espérance de la statistique I de Moran est constante pour toutes les années : $E(I) = -0,007$

Tableau 1 : Evolution de la statistique de Moran pour les PIB par tête en logarithmes et en Ecus sur la période 1980-1995

1980		1995	
Getis-Ord 10 plus grandes valeurs	Getis-Ord 10 plus petites valeurs	Getis-Ord 10 plus grandes valeurs	Getis-Ord 10 plus petites valeurs
Champagne-Ardenne	Extremadura (1)	Unterfranken	Voreio Aigaio (1)
Franche-Comté	Lisboa (1)	Arnsberg	Extremadura (1)
Bourgogne	Andalucia (1)	Rheinhessen-Pfalz	Dytiki Makedonia (1)
Unterfranken	Alentejo (1)	Karlsruhe	Thessalia (1)
Münster	Centro (1)	Darmstadt	Anatoliki Makedonia (1)
Arnsberg	East Anglia	Mittelfranken	Kentriki Makedonia (1)
Lorraine	Galicia	Münster	Andalucia (1)
Rheinhessen-Pfalz	Algarve (1)	Stuttgart	Lisboa (1)
Karlsruhe	Voreio Aigaio (1)	Champagne-Ardenne	Dytiki Ellada (1)
Picardie	Dytiki Makedonia (1)	Koblenz	Centro (1)

Tableau 2 : Valeurs extrêmes de la statistique Getis-Ord pour W(1) ; PIB par tête en logarithmes et en Ecus pour 1980 et 1995

Intervalle	Type0	I	II	IIIA	IIIB	IV	Cohésion	Flux
1 an	Total	0,019	0,023	0,004	0,000	0,954		
	Hausse	0,003	0,015	0,001	0,000	-	0,958	0,042
	Baisse	0,015	0,008	0,003	0,000	-		
5ans	Total	0,036	0,044	0,022	0,000	0,898		
	Hausse	0,005	0,021	0,005	0,000	-	0,920	0,080
	Baisse	0,031	0,023	0,017	0,000	-		
10 ans	Total	0,048	0,046	0,057	0,000	0,849		
	Hausse	0,007	0,025	0,010	0,000	-	0,906	0,094
	Baisse	0,041	0,021	0,047	0,000	-		
15 ans	Total	0,043	0,022	0,043	0,000	0,891		
	Hausse	0,007	0,000	0,000	0,000	-	0,935	0,065
	Baisse	0,036	0,022	0,043	0,000	-		

Tableau 3 : Cohésion et flux dans les diagrammes de Moran pour $W(1)$;
 PIB par tête en logarithmes et en Ecus sur la période 1980-1995

	1980		1995	
	Région	Résidus studentisés	Région	Résidus studentisés
Résidu-type excédant 2 en valeur absolue	Lüneburg	3,275735	Lüneburg	3,237612
	Vlaams Brabant	2,916681	Vlaams Brabant	2,215845
	Galicja	-2,029131	Darmstadt	-2,107749
	Stereia Ellada (1)	-2,539532	Madrid	-2,202453
	Alentejo (1)	-3,004722	Luxembourg	-2,374028
	Hamburg	-3,154712	Ile-de-France	-2,479044
	Bruxelles	-3,406122	Alentejo (1)	-2,721763
			Bruxelles	-2,781087
			Hamburg	-3,202301
Levier excédant 4/N	Lisboa (1)	0,072428	Ipeiros (1)	0,052553
	Centro (1)	0,065610	Hamburg	0,046399
	Algarve (1)	0,062048	Voreio Aigaio (1)	0,040424
	East Anglia	0,058095	Algarve (1)	0,038027
	Voreio Aigaio (1)	0,038353	Darmstadt	0,037616
	Hamburg	0,036314	Lisboa (1)	0,034994
	Extremadura (1)	0,035489	Centro (1)	0,032597
	Ipeiros (1)	0,035076	Dytiki Ellada (1)	0,031850
	Bruxelles	0,032278	Oberbayern	0,031539
	Alentejo (1)	0,031164	Luxembourg	0,030974
	Ionia Nisia (1)	0,029664	Peloponnisos (1)	0,030740
	Anatoliki Makedonia (1)	0,029641	Bremen	0,029290

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 4 : Observations extrêmes du diagramme de Moran pour $W(1)$;
 PIB par tête en logarithmes et en Ecus pour 1980 et 1995

Valeurs extrêmes du diagramme de Moran		Valeurs extrêmes Getis-Ord	
Résidus studentisés	Levier	10 plus grandes valeurs	10 plus petites valeurs
-	East Anglia	0,105763	Extremadura (1) Northern Ireland (1)
	Lisboa	0,102492	Alentejo (1) Scotland
	Centro (1)	0,089878	Lisboa (1) Wales
	Stereia Ellada (1)	0,065942	Algarve (1) North
	Alentejo	0,064531	Andalucia (1) South West
	Luxembourg	0,055558	Centro (1) West Midlands
	Algarve (1)	0,054656	Galicia (1) Yorkshire and Humberside
	Picardie (1)	0,030390	Norte (1) East Midlands
		Asturias (1) North West	
		Oberfranken Bourgogne	

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 5 : Valeurs extrêmes du diagramme de Moran et des statistiques de Getis-Ord pour $W(1)$; PIB par tête en Ecus ; taux de croissance annuel moyen

Code	Région	G_j 1980	I_j 1980	G_j 1995	I_j 1995	G_j croiss.	I_j croiss.	Code	Région	G_j 1980	I_j 1980	G_j 1995	I_j 1995	G_j croiss.	I_j croiss.
BELGIQUE															
Be1	Bruxelles	+	HH	+	HH	+	BH	Fr51	Pays de la Loire	+	HH	+	HH	-*	BB
Be21	Antwerpen	+*	HH	+*	HH	-	BB	Fr52	Bretagne	+	HH	+	HH	-*	BB
Be22	Limburg (B)	+*	HH	+*	HH	-	HB	Fr53	Poitou-Charentes	+	HH	+	HH	-*	BB
Be23	Oost-Vlaanderen	+*	HH	+*	HH	-	BB	Fr61	Aquitaine	+	HH	-	HB	-	BB
Be24	Vlaams Brabant	+*	HH	+	HH	-	HB	Fr62	Midi-Pyrénées	+	HH	+	HH	-	BB
Be25	West-Vlaanderen	+*	HH	+*	HH	-	HB	Fr63	Limousin	+	HH	+	HH	-*	BB
Be31	Brabant wallon	+	HH	+	HH	-	BB	Fr71	Rhône-Alpes	+*	HH	+	HH	-*	BB
Be32	Hainaut	+*	HH	+*	BH	-	BB	Fr72	Auvergne	+*	HH	+	HH	-*	BB
Be33	Liège	+*	HH	+*	HH	+	BH	Fr81	Languedoc-Roussillon	+	HH	+	BH	-	BB
Be34	Luxembourg (B)	+*	HH	+*	HH	+	HH	Fr82	PACA	+	HH	+	HH	-	BB
Be35	Namur	+*	HH	+*	BH	-	BB	GRECE							
ALLEMAGNE								Gr11	Anatoliki Makedonia (1)	-**	BB	-**	BB	-	HB
De11	Stuttgart	+**	HH	+**	HH	+	HH	Gr12	Kentriki Makedonia (1)	-**	BB	-**	BB	-	BB
De12	Karlsruhe	+**	HH	+**	HH	+	HH	Gr13	Dytiki Makedonia (1)	-**	BB	-**	BB	-	BB
De13	Freiburg	+*	HH	+**	HH	+	HB	Gr14	Thessalia (1)	-**	BB	-**	BB	-*	BB
De14	Tübingen	+*	HH	+**	HH	+	HH	Gr21	Ipeiros (1)	-*	BB	-**	BB	-	BB
De21	Oberbayern	+*	HH	+*	HH	+*	HH	Gr22	Ionia Nisia (1)	-*	BB	-**	BB	-	HB
De22	Niederbayern	+*	HH	+**	HH	+*	HH	Gr23	Dytiki Ellada (1)	-*	BB	-**	BB	-*	BB
De23	Oberpfalz	+*	HH	+**	HH	+*	HH	Gr24	Sterea Ellada (1)	-*	BB	-**	BB	-	BB
De24	Oberfranken	+*	HH	+**	HH	+*	HH	Gr25	Peloponnisos (1)	-*	BB	-**	BB	-	BB
De25	Mittelfranken	+**	HH	+**	HH	+*	HH	Gr3	Attiki (1)	-*	BB	-**	BB	-*	BB
De26	Unterfranken	+**	HH	+**	HH	+*	HH	Gr41	Voreio Aigaio (1)	-**	BB	-**	BB	-	HB
De27	Schwaben	+**	HH	+**	HH	+	HH	Gr42	Notio Aigaio (1)	-*	BB	-**	BB	-	HB
De5	Bremen	+*	HH	+**	HH	+	HH	Gr43	Kriti (1)	-*	BB	-**	BB	-	HB
De6	Hamburg	+	HH	+	HH	+	HH	ITALIE							
De71	Darmstadt	+**	HH	+**	HH	+	HH	It11	Piemonte	+	HH	+	HH	-	BB
De72	Giessen	+**	HH	+**	HH	+	HH	It12	Valle d'Aosta	+	HH	+	HH	-	BB
De73	Kassel	+*	HH	+**	HH	+	HH	It13	Liguria	+	HH	+	HH	-	HB
De91	Braunschweig	+**	HH	+**	HH	+	HH	It21	Lombardia	+	HH	+*	HH	+	BH
De92	Hannover	+*	HH	+**	HH	+	HH	It31	Trentino – Alto Adige	+	HH	+	HH	+	HH
De93	Lüneburg	+*	HH	+**	HH	+	HH	It32	Veneto	+	HH	+	HH	+	HH
De94	Weser-Ems	+**	HH	+**	HH	-	HB	It33	Friuli – Venezia Giulia	+	HH	+	HH	+	HH
Dea1	Düsseldorf	+*	HH	+**	HH	-	BB	It4	Emilia – Romagna	+	HH	+	HH	+	BH
Dea2	Köln	+**	HH	+**	HH	+	HH	It51	Toscana	+	HH	+	HH	-	BB
Dea3	Münster	+**	HH	+**	HH	-	HB	It52	Umbria	-	BB	-	BB	-	BB
Dea4	Detmold	+*	HH	+**	HH	+	HH	It53	Marche	-	BB	-	BB	-	BB
Dea5	Arnsberg	+**	HH	+**	HH	+	BH	It6	Lazio	-	BB	-	HB	-	HB
Deb1	Koblenz	+**	HH	+**	HH	+	HH	It71	Abruzzo (1)	-	BB	-	BB	+	HH
Deb2	Trier	+*	HH	+**	HH	+*	HH	It72	Molise (1)	-	BB	-	BB	+	HH
Deb3	Rheinhessen-Pfalz	+**	HH	+**	HH	+	BH	It8	Campania (1)	-*	BB	-*	BB	-	BB
Dec	Saarland	+**	HH	+**	HH	+	HH	It91	Puglia (1)	-	BB	-	BB	-	BB
Def	Schleswig-Holstein	+*	HH	+**	HH	+	HH	It92	Basilicata (1)	-	BB	-	BB	-	BB
Dk	DANEMARK	+*	HH	+*	HH	+	HH	It93	Calabria (1)	-*	BB	-**	BB	-	HB
ESPAÑE								Ita	Sicilia (1)	-*	BB	-*	BB	-	BB
Es11	Galicía (1)	-**	BB	-**	BB	+**	BH	Itb	Sardegna (1)	-	BB	-	BB	-	HB
Es12	Asturias (1)	-**	BB	-*	BB	+	BH	Lu	LUXEMBOURG	+	HH	+*	HH	+	HB
Es13	Cantabria	-*	BB	-	BB	+	BH	PAYS-BAS							
Es21	Pais Vasco	-	BB	-	BB	+	HH	NI12	Friesland	+*	HH	+*	HH	-	BB
Es22	Navarra	-	BB	-	BB	+	HH	NI13	Drenthe	+*	HH	+*	HH	-	BB
Es23	La Rioja	-	BB	-	BB	+	HH	NI2	Oost Nederland	+**	HH	+**	HH	-	BB
Es24	Aragon	-	BB	-	BB	+	HH	NI31	Utrecht	+*	HH	+*	HH	-	HB
Es3	Madrid	-**	BB	-*	BB	+	HH	NI32	Noord-Holland	+*	HH	+*	HH	-	BB
Es41	Castilla-Leon (1)	-**	BB	-*	BB	+*	HH	NI33	Zuid-Holland	+*	HH	+*	HH	-	BB
Es42	Castilla-la Mancha (1)	-**	BB	-*	BB	+*	HH	NI34	Zeeland	+*	HH	+*	HH	-	BB
Es43	Extremadura (1)	-**	BB	-**	BB	+**	HH	NI41	Noord-Brabant	+*	HH	+*	HH	-	HB
Es51	Cataluna	-	BB	-	BB	+	HH	NI42	Limburg (NL)	+*	HH	+*	HH	-	BB
Es52	Valenciana (1)	-*	BB	-*	BB	+	HH	PORTUGAL							
Es53	Islas Baleares	-	BB	-	BB	+	HH	Pt11	Norte (1)	-**	BB	-**	BB	+**	HH
Es61	Andalucia (1)	-**	BB	-**	BB	+**	HH	Pt12	Centro (1)	-**	BB	-**	BB	+**	HH
Es62	Murcia (1)	-*	BB	-*	BB	+	HH	Pt13	Lisboa e vale do Tejo (1)	-**	BB	-**	BB	+**	HH
FRANCE								Pt14	Alentejo (1)	-**	BB	-*	BB	+**	HH
Fr1	Ile-de-France	+**	HH	+*	HH	-**	BB	Pt15	Algarve (1)	-**	BB	-**	BB	+**	HH
Fr21	Champagne-Ardenne	+**	HH	+**	HH	-*	BB	ROYAUME-UNI							
Fr22	Picardie	+**	HH	+**	HH	-*	BB	Uk1	North	+	HH	-	BB	-*	BB
Fr23	Haute-Normandie	+**	HH	+*	HH	-*	BB	Uk2	Yorkshire and Humberside	+	HH	-	BH	-**	BB
Fr24	Centre	+**	HH	+*	HH	-*	BB	Uk3	East Midlands	+	HH	-	HH	-*	BB
Fr25	Basse-Normandie	+*	HH	+	HH	-*	BB	Uk4	East Anglia	+*	HH	+	BB	-**	BB
Fr26	Bourgogne	+**	HH	+**	HH	-**	BB	Uk5	South East	+*	HH	+	BB	-**	BB
Fr3	Nord-Pas-De-Calais	+*	HH	+*	HH	-	BB	Uk6	South West	+	HH	-	BB	-**	BB
Fr41	Lorraine	+**	HH	+**	HH	-	BB	Uk7	West Midlands	+	HH	-	BB	-**	BB
Fr42	Alsace	+*	HH	+**	HH	+	BH	Uk8	North West	+	HH	-	BB	-*	BB
Fr43	Franche-Comté	+**	HH	+**	HH	-	BB	Uk9	Wales	+	BH	-	BB	-*	BB
								Uka	Scotland	-	HB	-	BB	-*	BB
								Ukb	Northern Ireland (1)	+	BH	-	BB	-**	BB

Notes : (*) significatif à 5%, basée sur une approximation normale pour Getis-Ord et sur 10 000 permutations pour LISA
(**) significatif à 5% Sidak pour Getis-Ord et à 5% Bonferroni pour LISA

Tableau 6 : Statistiques Getis-Ord et LISA pour la matrice $W(1)$;
PIB par tête en logarithmes et en Ecus pour 1980, 1995 et pour le taux de croissance annuel moyen

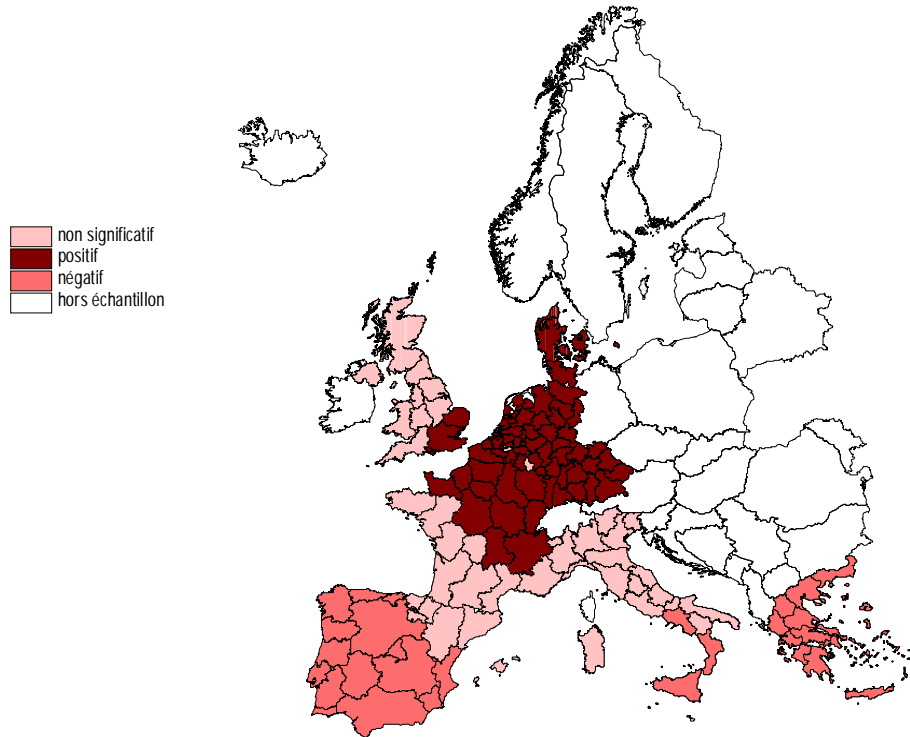


Figure 1 : Statistiques de Getis-Ord significatives pour $W(1)$;
PIB par tête en logarithmes et en Ecus pour 1980

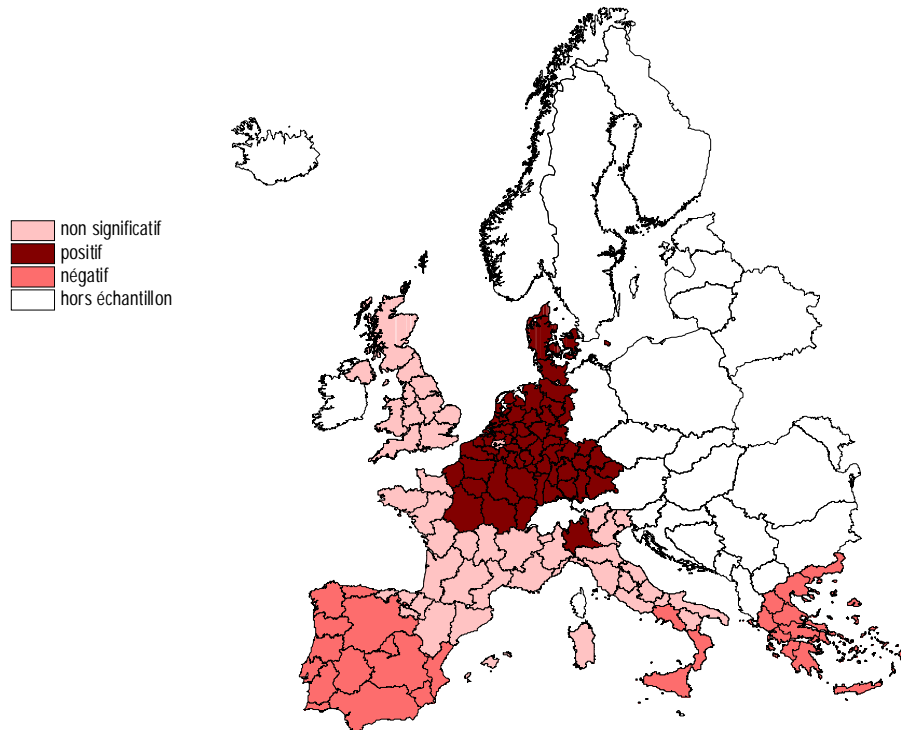


Figure 2 : Statistiques de Getis-Ord significatives pour $W(1)$;
PIB par tête en logarithmes et en Ecus pour 1995

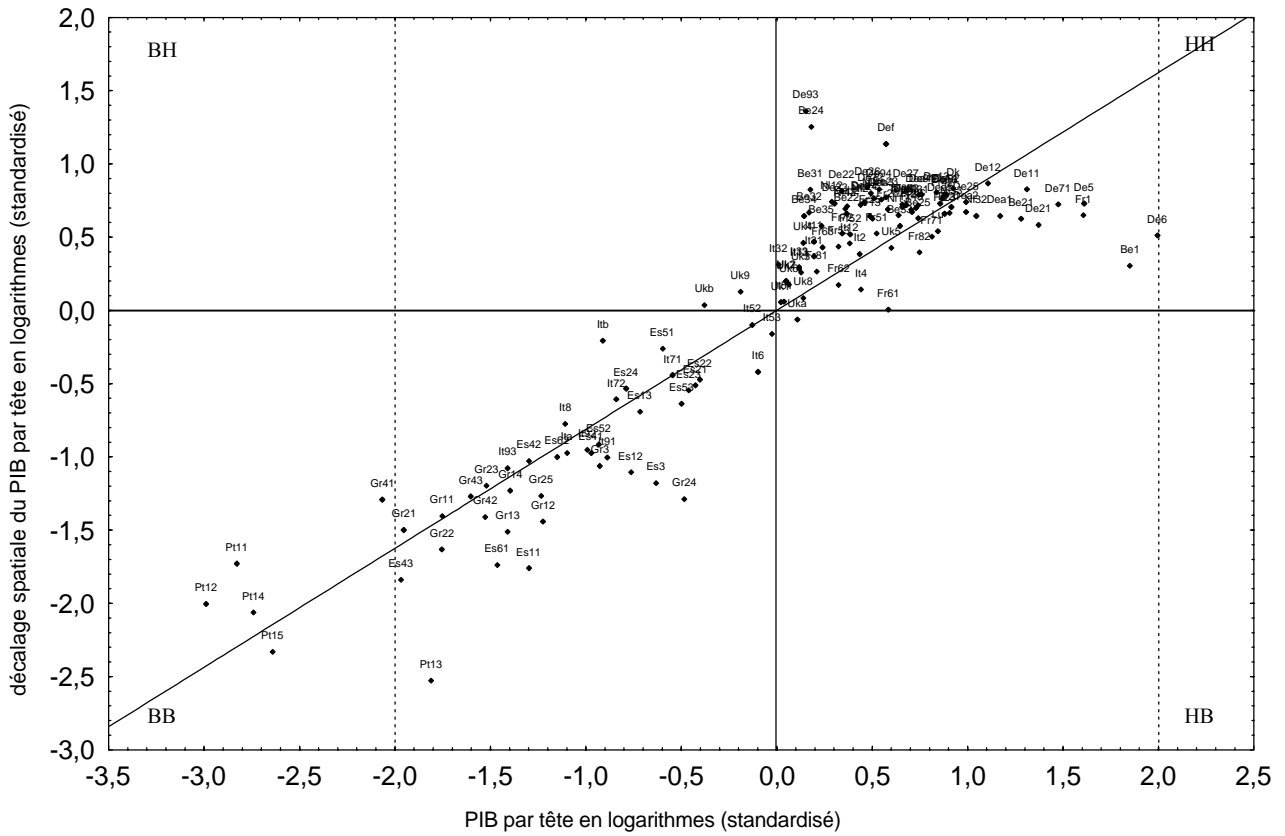


Figure 3 : Diagramme de Moran pour $W(1)$; PIB par tête en Ecus pour 1980

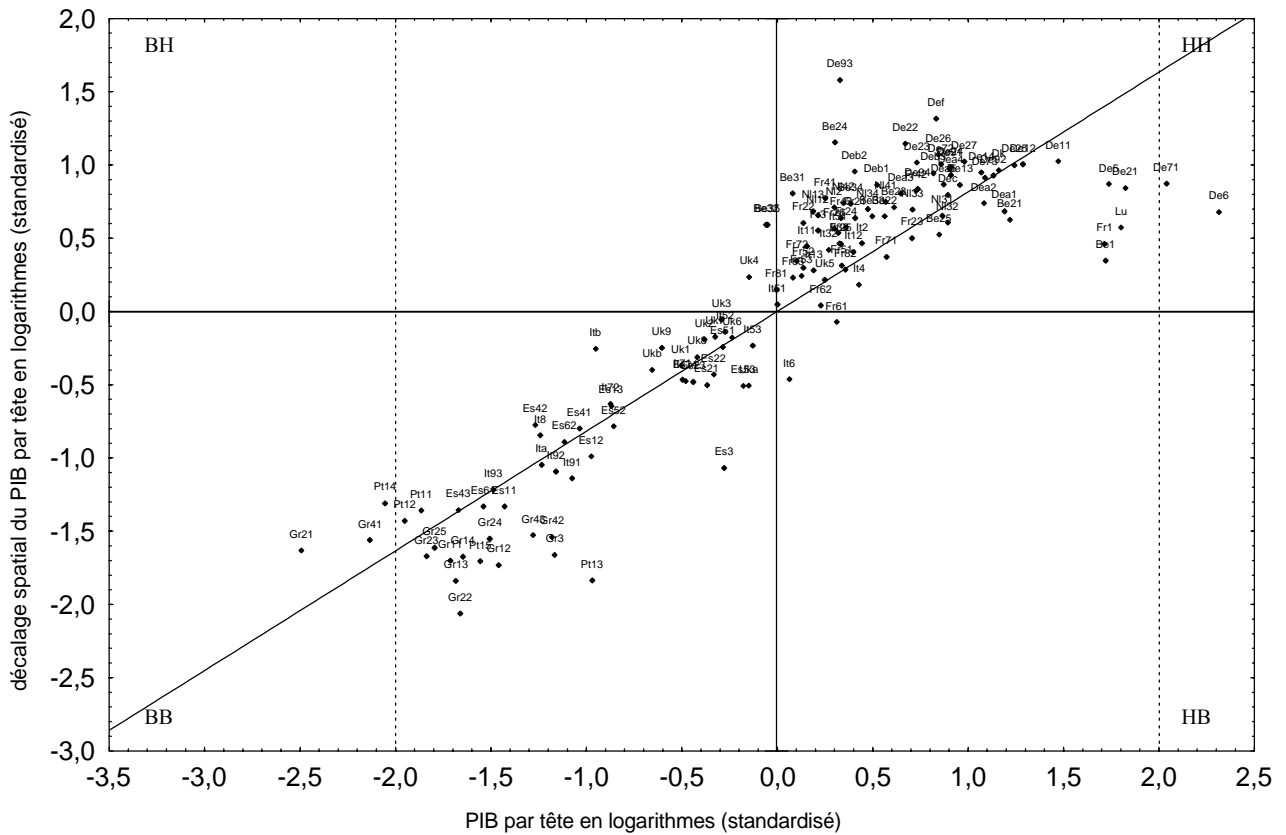


Figure 4 : Diagramme de Moran pour $W(1)$; PIB par tête en Ecus pour 1995

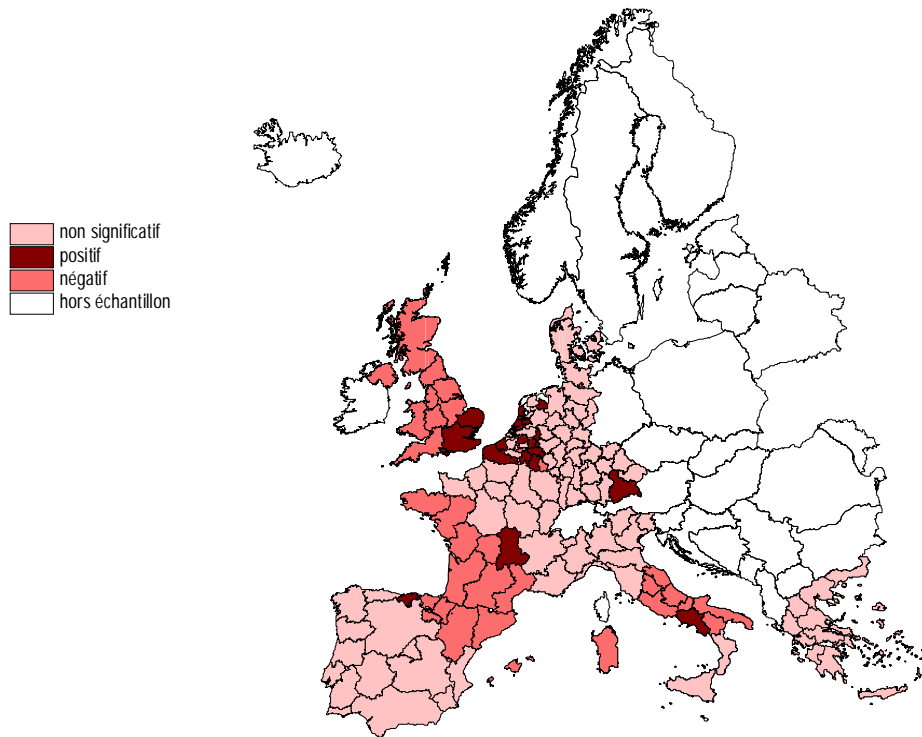


Figure 5 : Statistiques de Getis-Ord significatives pour $W(1)$;
Taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus

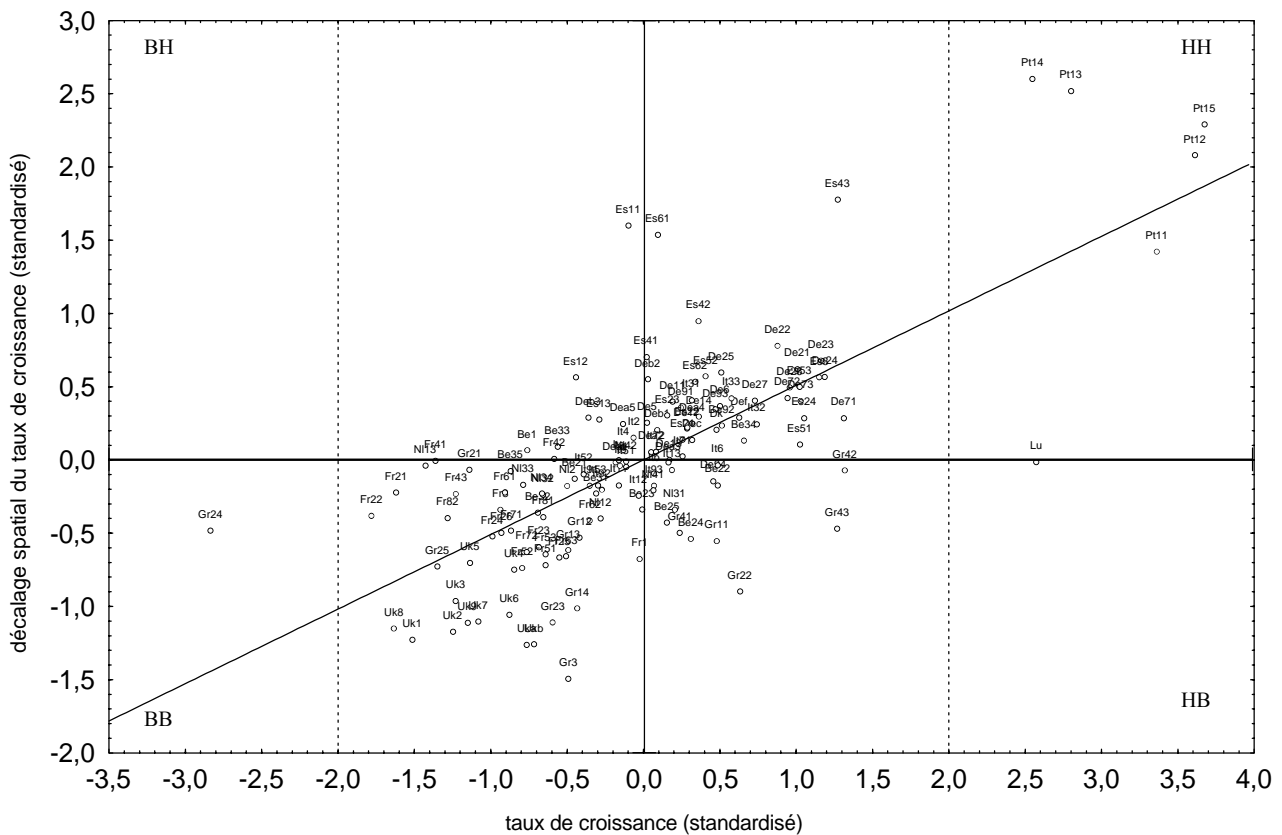


Figure 6 : Diagramme de Moran pour le taux de croissance annuel moyen
entre 1980 et 1995 pour $W(1)$; PIB par tête en Ecus

Annexe 3.3

Résultats pour les sections 32 et 33 avec les matrices des distances inverses ; PIB par tête en SPA

Année	W(1)			W(2)			W(3)		
	I de Moran	Ecart-type	Stand.	I de Moran	Ecart-type	Stand.	I de Moran	Ecart-type	Stand.
1980	0,641	0,040368	16,046	0,606	0,034917	17,545	0,553	0,032950	17,003
1981	0,625	0,040345	15,653	0,587	0,034901	17,015	0,535	0,032932	16,454
1982	0,629	0,040371	15,744	0,591	0,034928	17,134	0,539	0,032959	16,554
1983	0,638	0,040422	15,966	0,604	0,034969	17,464	0,552	0,032996	16,945
1984	0,644	0,040435	16,098	0,609	0,034958	17,628	0,558	0,032986	17,118
1985	0,642	0,040354	16,074	0,607	0,034895	17,598	0,555	0,032916	17,079
1986	0,637	0,040544	15,895	0,602	0,035066	17,365	0,550	0,033073	16,848
1987	0,634	0,040518	15,829	0,600	0,03505	17,313	0,550	0,033058	16,838
1988	0,625	0,040537	15,583	0,591	0,035061	17,055	0,541	0,033071	16,579
1989	0,615	0,040446	15,380	0,580	0,034987	16,786	0,532	0,032991	16,327
1990	0,643	0,040348	16,107	0,610	0,034898	17,666	0,560	0,032912	17,227
1991	0,643	0,040213	16,170	0,608	0,034763	17,684	0,560	0,032785	17,285
1992	0,650	0,040218	16,334	0,615	0,034769	17,887	0,566	0,032791	17,475
1993	0,627	0,040244	15,752	0,595	0,034800	17,288	0,548	0,032816	16,927
1994	0,626	0,040230	15,746	0,592	0,034786	17,229	0,546	0,032803	16,844
1995	0,624	0,040217	15,680	0,587	0,034766	17,097	0,541	0,032782	16,719

Note : L'espérance de la statistique I de Moran est constante pour toutes les années : $E(I) = -0,007$

Tableau 1 : Evolution de la statistique de Moran pour les PIB par tête en logarithmes et en SPA sur la période 1980-1995

1980		1995	
Getis-Ord 10 plus grandes valeurs	Getis-Ord 10 plus petites valeurs	Getis-Ord 10 plus grandes valeurs	Getis-Ord 10 plus petites valeurs
Franche-Comté	Extremadura (1)	Unterfranken	Dytiki Makedonia (1)
Bourgogne	Andalucia (1)	Lombardia	Voreio Aigaio (1)
Lombardia	Voreio Aigaio (1)	Rheinhessen-Pfalz	Thessalia (1)
Unterfranken	Dytiki Makedonia (1)	Karlsruhe	Anatoliki Makedonia (1)
Karlsruhe	Lisboa (1)	Stuttgart	Kentriki Makedonia (1)
Champagne-Ardenne	Anatoliki Makedonia (1)	Lüneburg	Extremadura (1)
Rheinhessen-Pfalz	Kentriki Makedonia (1)	Schwaben	Ionia Nisia (1)
Lüneburg	Centro (1)	Franche-Comté	Andalucia (1)
Lorraine	Alentejo (1)	Arnsberg	Dytiki Ellada (1)
Schwaben	East Anglia	Mittelfranken	Centro (1)

Tableau 2 : Valeurs extrêmes de la statistique Getis-Ord pour W(1) ; PIB par tête en logarithmes et en SPA pour 1980 et 1995

Intervalle	Type0	I	II	IIIA	IIIB	IV	Cohésion	Flux
1 an	Total	0,028	0,028	0,001	0,000	0,943	0,945	0,055
	Hausse	0,007	0,019	0,000	0,000	-		
	Baisse	0,021	0,008	0,001	0,000	-		
5ans	Total	0,053	0,061	0,013	0,000	0,873	0,886	0,114
	Hausse	0,012	0,040	0,002	0,000	-		
	Baisse	0,042	0,021	0,011	0,000	-		
10 ans	Total	0,087	0,064	0,014	0,002	0,832	0,847	0,151
	Hausse	0,023	0,036	0,001	0,001	-		
	Baisse	0,064	0,028	0,013	0,001	-		
15 ans	Total	0,116	0,101	0,014	0,000	0,768	0,783	0,217
	Hausse	0,029	0,051	0,000	0,000	-		
	Baisse	0,087	0,051	0,014	0,000	-		

Tableau 3 : Cohésion et flux dans les diagrammes de Moran pour $W(1)$;
 PIB par tête en logarithmes et en SPA sur la période 1980-1995

		1980		1995	
		Région	Résidus studentisés	Région	Résidus studentisés
Résidu-type excédant 2 en valeur Absolue	Lüneburg		3,4617901	Lüneburg	-2,940089
	Vlaams Brabant		3,05981867	Vlaams Brabant	3,324574
	Stereia Ellada (1)		-2,72951457	Attiki (1)	2,372115
	Hamburg		-3,01058336	Luxembourg	-2,072648
	Bruxelles		-3,37543051	Ile-de-France	-2,106307
	Alentejo (1)		-3,45907291	Madrid	-2,127090
Lever excédant 4/N	Voreio Aigaio (1)		0,053695	Ipeiros (1)	0,030798
	Hamburg		0,052013	Hamburg	0,040955
	Lisboa (1)		0,050041	Voreio Aigaio (1)	0,044988
	Extremadura (1)		0,048691	Darmstadt	0,030022
	Ipeiros (1)		0,047648	Bruxelles	0,031672
	Bruxelles		0,043457	Luxembourg	0,065033
	Centro (1)		0,041845	Ile-de-France	0,038955
	Ionias Nisia (1)		0,037876	Extremadura (1)	0,030518
	Anatoliki Makedonia (1)		0,037838	Oberbayern	0,035889
	Algarve (1)		0,037740	Dytiki Ellada (1)	0,044826
	Ile-de-France		0,035943	Bremen	0,059825
	Bremen		0,033375	Algarve (1)	0,031306
	East Anglia		0,033346	Andalucia (1)	0,034872
	Kriti (1)		0,031384	Peloponnisos (1)	0,043646

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 4 : Observations extrêmes du diagramme de Moran pour $W(1)$;
 PIB par tête en logarithmes et en SPA pour 1980 et 1995

Valeurs extrêmes du diagramme de Moran				Valeurs extrêmes Getis-Ord	
Résidus studentisés		Lever		10 plus grandes valeurs	10 plus petites valeurs
Algarve (1)	3,254746	Luxembourg (Lux)	0,089036	Extremadura (1)	Bourgogne
Alentejo (1)	2,639222	Stereia Ellada (1)	0,081831	Lisboa (1)	Ile-de-France
Andalucia (1)	2,594280	Notio Aigaio (1)	0,054936	Algarve (1)	Auvergne
Galicia	2,202333	East Anglia	0,054338	Andalucia (1)	Limousin
Ionian Nisia (1)	-2,290450	Kriti (1)	0,052348	Alentejo (1)	Rhône-Alpes
Attiki (1)	-3,171888	Picardie	0,052143	Trier	Centre
		Lisboa (1)	0,051240	East Anglia	Poitou-Charentes
		Champagne-Ardenne	0,044446	Centro (1)	Pays de la Loire
		Centro (1)	0,039800	Castilla-La-Mancha (1)	Franche-Comté
		Lorraine	0,033541	Liège	Attiki (1)
		PACA	0,030651		

Note : (1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Tableau 5 : Valeurs extrêmes du diagramme de Moran et des statistiques de Getis-Ord pour $W(1)$; PIB par tête en SPA ; taux de croissance annuel moyen

Code	Région	G _i	I _i	G _i	I _i	G _i	I _i	Code	Région	G _i	I _i	G _i	I _i	G _i	I _i
		1980	1980	1995	1995	croiss.	croiss.			1980	1980	1995	1995	croiss.	croiss.
BELGIQUE								Fr51	Pays de la Loire	+	HH	+	HH	-	BB
Be1	Bruxelles	-	HB	+	HH	+	BH	Fr52	Bretagne	+	HH	+	BH	-*	BB
Be21	Anvers	+	HH	+	HH	+	HH	Fr53	Poitou-Charentes	+	HH	-	BB	-*	BB
Be22	Limburg (B)	+	BH	+	HH	+	HH	Fr61	Aquitaine	-	HB	-	BB	-	BB
Be23	Oost Vlaanderen	+	HH	+	HH	+	HH	Fr62	Midi-Pyrénées	+	HH	-	BB	-*	BB
Be24	Vlaams Brabant	+	BH	+	HH	-	HB	Fr63	Limousin	+	BH	-	BB	-**	BB
Be25	West Vlaanderen	+	HH	+	HH	-	HB	Fr71	Rhône-Alpes	+	HH	+	HH	-**	BB
Be31	Brabant wallon	+	BH	+	BH	+	HH	Fr72	Auvergne	+	HH	+	BH	-**	BB
Be32	Hainaut	+	BH	+	BH	-	BB	Fr81	Languedoc-Roussillon	+	BH	-	BB	-*	BB
Be33	Liège	+	HH	+	HH	+	BH	Fr82	PACA	+	HH	+	HH	-*	BB
Be34	Luxembourg (B)	+	BH	+	HH	+	HH	GRECE							
Be35	Namur	+	BH	+	BH	+	BH	Gr11	Anatoliki Makedonia (1)	-**	BB	-**	BB	-	HB
ALLEMAGNE								Gr12	Kentriki Makedonia (1)	-**	BB	-**	BB	-	HB
De11	Stuttgart	+	HH	+	HH	+	BH	Gr13	Dytiki Makedonia (1)	-**	BB	-**	BB	-	HB
De12	Karlsruhe	+	HH	+	HH	-	BB	Gr14	Thessalia (1)	-**	BB	-**	BB	-	HB
De13	Freiburg	+	HH	+	HH	-	BB	Gr21	Ipeiros (1)	-**	BB	-**	BB	+	BH
De14	Tübingen	+	HH	+	HH	+	BH	Gr22	Ionia Nisia (1)	-**	BB	-**	BB	-	HB
De21	Oberbayern	+	HH	+	HH	+	HH	Gr23	Dytiki Ellada (1)	-*	BB	-**	BB	-	BB
De22	Niederbayern	+	BH	+	HH	+	HH	Gr24	Stereia Ellada (1)	-*	BB	-**	BB	+	BH
De23	Oberpfalz	+	BH	+	HH	+	HH	Gr25	Peloponnisos (1)	-*	BB	-**	BB	-	BB
De24	Oberfranken	+	HH	+	HH	+	HH	Gr3	Attiki (1)	-*	BB	-**	BB	-*	HB
De25	Mittelfranken	+	HH	+	HH	+	HH	Gr41	Voreio Aigaiio (1)	-**	BB	-**	BB	+	HH
De26	Unterfranken	+	HH	+	HH	+	HH	Gr42	Notio Aigaiio (1)	-**	BB	-*	BB	+	HH
De27	Schwaben	+	HH	+	HH	+	HH	Gr43	Kriti (1)	-**	BB	-**	BB	+	HH
De5	Bremen	+	HH	+	HH	+	BH	ITALIE							
De6	Hamburg	+	HH	+	HH	+	HH	It11	Piemonte	+	HH	+	HH	-	BB
De71	Darmstadt	+	HH	+	HH	+	HH	It12	Valle d'Aosta	+	HH	+	HH	-	BB
De72	Giessen	+	HH	+	HH	+	HH	It13	Liguria	+	HH	+	HH	-	HB
De73	Kassel	+	HH	+	HH	+	HH	It2	Lombardia	+	HH	+	HH	-	BB
De91	Braunschweig	+	HH	+	HH	+	BH	It31	Trentino – Alto Adige	+	HH	+	HH	+	HH
De92	Hannover	+	HH	+	HH	-	HB	It32	Veneto	+	HH	+	HH	+	HH
De93	Lüneburg	+	BH	+	BH	+	HH	It33	Friuli – Venezia Giulia	+	HH	+	HH	+	HH
De94	Weser-Ems	+	HH	+	HH	-	HB	It4	Emilia – Romagna	+	HH	+	HH	-	BB
Dea1	Düsseldorf	+	HH	+	HH	+	BH	It51	Toscana	+	HH	+	HH	-	BB
Dea2	Köln	+	HH	+	HH	+	BH	It52	Umbria	+	HH	+	HH	-	BB
Dea3	Münster	+	HH	+	HH	-	BB	It53	Marche	+	HH	+	HH	-	BB
Dea4	Detmold	+	HH	+	HH	+	BH	It6	Lazio	+	HH	-	HB	-	HB
Dea5	Arnsberg	+	HH	+	HH	+	BH	It71	Abruzzo (1)	-	BB	-	BB	-	HB
Deb1	Koblenz	+	HH	+	BH	+	BH	It72	Molise (1)	-	BB	-	BB	-	BB
Deb2	Trier	+	HH	+	BH	+	BH	It8	Campania (1)	-	BB	-	BB	-	BB
Deb3	Rheinhausen-Pfalz	+	HH	+	HH	+	BH	It91	Puglia (1)	-	BB	-	BB	-	BB
Dec	Saarland	+	HH	+	HH	+	BH	It92	Basilicata (1)	-	BB	-	BB	-	BB
Def	Schleswig-Holstein	+	HH	+	HH	+	HH	It93	Calabria (1)	-*	BB	-*	BB	-	BB
Dk	DANEMARK	+	HH	+	HH	+	HH	Ita	Sicilia (1)	-*	BB	-**	BB	-	BB
ESPAGNE								Itb	Sardegna (1)	+	BH	+	BH	-	BB
Es11	Galicía (1)	-**	BB	-*	BB	+	BH	Lu	LUXEMBOURG	+	HH	+	HH	+	HH
Es12	Asturias (1)	-*	BB	-*	BB	+	BH	PAYS-BAS							
Es13	Cantabria	-	BB	-	BB	-	BB	NI12	Friesland	+	BH	+	BH	-	HB
Es21	Pais Vasco	-	BB	-	BB	-	HB	NI13	Drenthe	+	HH	+	BH	+	BH
Es22	Navarra	-	BB	-	HB	+	HH	NI2	Oost Nederland	+	HH	+	HH	+	HH
Es23	La Rioja	-	BB	-	BB	+	BH	NI31	Utrecht	+	HH	+	HH	+	HH
Es24	Aragon	-	BB	-	BB	+	HH	NI32	Noord-Holland	+	HH	+	HH	+	BH
Es3	Madrid	-*	BB	-*	HB	+	HH	NI33	Zuid-Holland	+	HH	+	HH	+	BH
Es41	Castilla-Leon (1)	-*	BB	-*	BB	+	BH	NI34	Zeeland	+	HH	+	HH	+	BH
Es42	Castilla-la Mancha (1)	-*	BB	-	BB	+	HH	NI41	Noord-Brabant	+	HH	+	HH	+	HH
Es43	Extremadura (1)	-**	BB	-**	BB	+	HH	NI42	Limburg (NL)	+	BH	+	HH	+	HH
Es51	Cataluna	-	BB	-	HB	-	HB	PORTUGAL							
Es52	Valenciana (1)	-*	BB	-	BB	+	HH	Pt11	Norte (1)	-**	BB	-*	BB	+	HH
Es53	Islas Baleares	-	BB	-	HB	+	HH	Pt12	Centro (1)	-**	BB	-*	BB	+	HH
Es61	Andalucía (1)	-**	BB	-**	BB	+	BH	Pt13	Lisboa e vale do Tejo (1)	-**	BB	-*	BB	+	HH
Es62	Murcia (1)	-*	BB	-*	BB	+	HH	Pt14	Alentejo (1)	-*	BB	-	BB	+	HH
FRANCE								Pt15	Algarve (1)	-**	BB	-*	BB	+	HH
Fr1	Ile-de-France	+	HH	+	HH	-**	BB	ROYAUME-UNI							
Fr21	Champagne-Ardenne	+	HH	+	HH	-	BB	Uk1	North	-	BH	-	BB	-	BB
Fr22	Picardie	+	HH	+	BH	-	BB	Uk2	Yorkshire and Humberside	+	HH	-	BB	-	BB
Fr23	Haute-Normandie	+	HH	+	HH	-*	BB	Uk3	East Midlands	+	HH	+	BH	-	BB
Fr24	Centre	+	HH	+	HH	-**	BB	Uk4	East Anglia	+	HH	+	HH	-	HB
Fr25	Basse-Normandie	+	HH	+	HH	-*	BB	Uk5	South East	+	BH	+	HH	-	BB
Fr26	Bourgogne	+	HH	+	BH	-**	BB	Uk6	South West	+	BH	-	BB	-	HB
Fr3	Nord-Pas-De-Calais	+	HH	+	BH	+	BH	Uk7	West Midlands	+	HB	-	BB	-	BB
Fr41	Lorraine	+	HH	+	BH	-	BB	Uk8	North West	-	BH	-	BB	-	BB
Fr42	Alsace	+	HH	+	HH	-	BB	Uk9	Wales	+	HB	-	BB	-	BB
Fr43	Franche-Comté	+	HH	+	HH	-*	BB	Uka	Scotland	-	BB	-	HB	-	HB
								Ukb	Northern Ireland (1)	-	BH	-	BB	-	HB

Notes : (*) significatif à 5%, basée sur une approximation normale pour Getis-Ord et sur 10 000 permutations pour LISA
(**) significatif à 5% Sidák pour Getis-Ord et à 5% Bonferroni pour LISA

Tableau 5 : Statistiques Getis-Ord et LISA pour la matrice W(1) ;
PIB par tête en logarithmes et en SPA pour 1980, 1995 et pour le taux de croissance annuel moyen

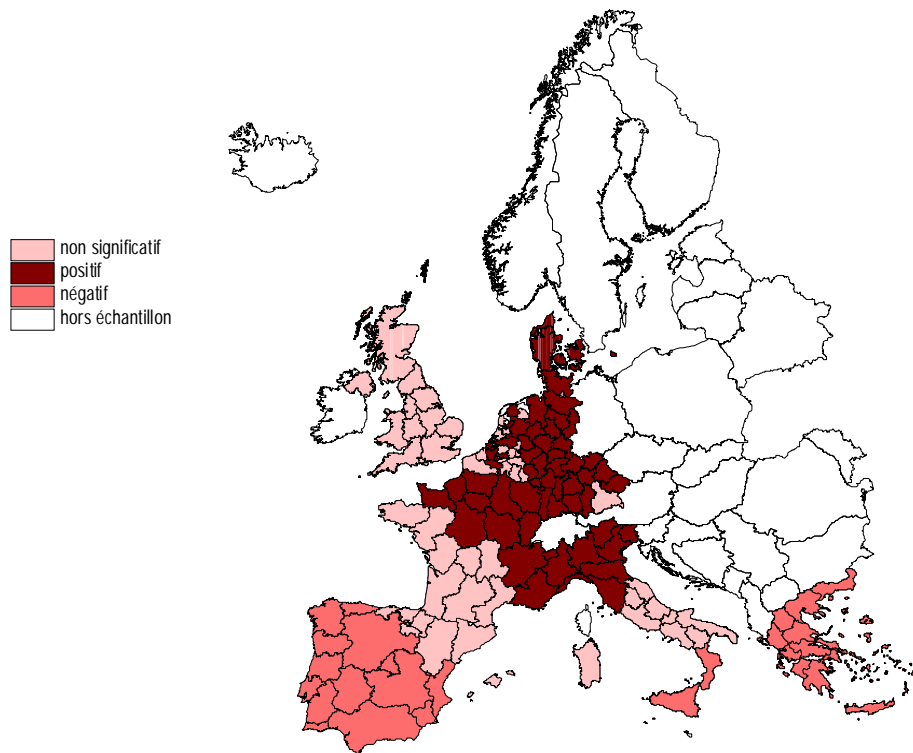


Figure 1 : Statistiques de Getis-Ord significatives pour $W(1)$;
PIB par tête en logarithmes et en SPA pour 1980

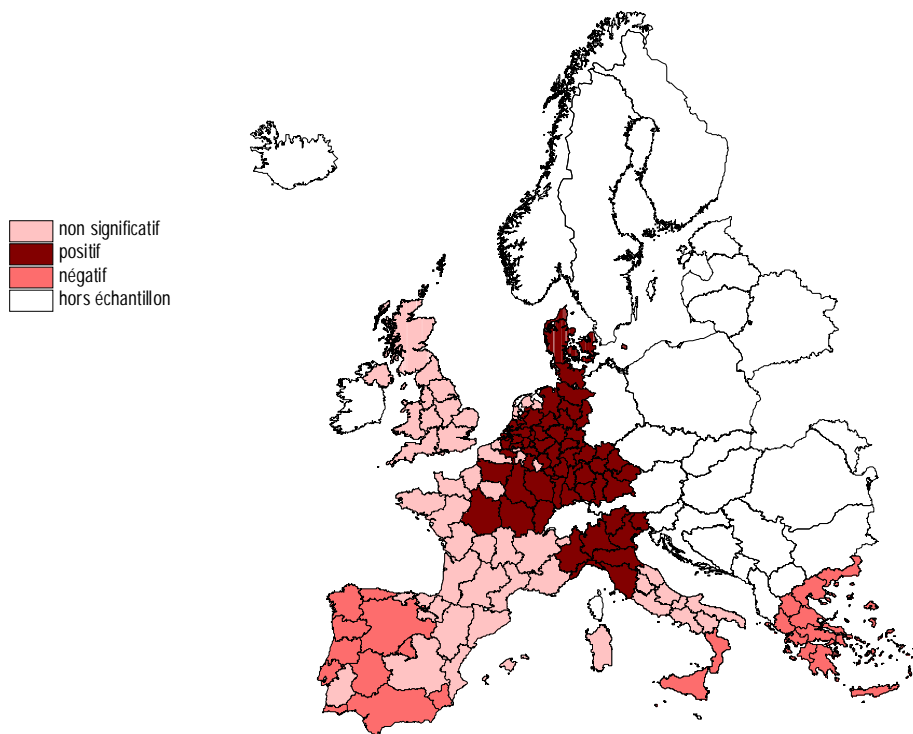


Figure 2 : Statistiques de Getis-Ord significatives pour $W(1)$;
PIB par tête en logarithmes et en SPA pour 1995

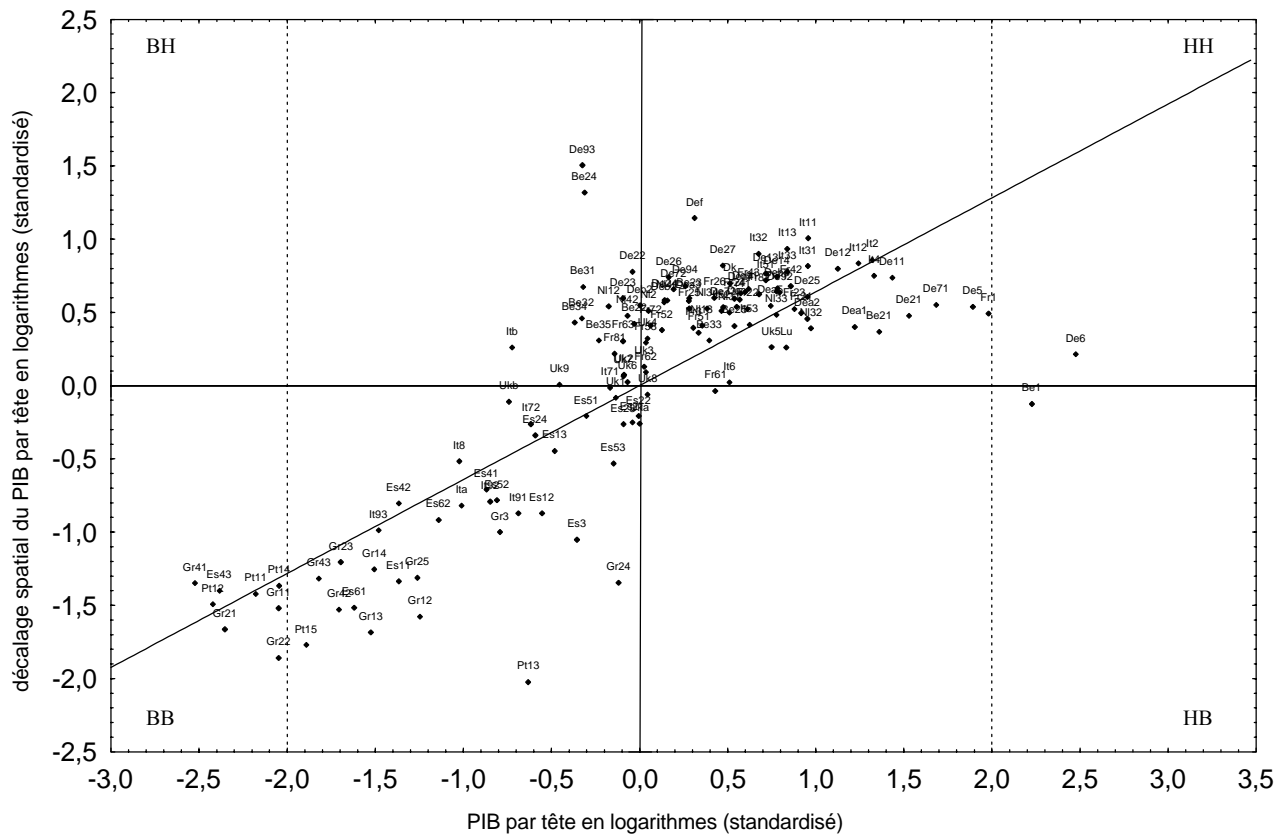


Figure 3 : Diagramme de Moran pour $W(1)$; PIB par tête en SPA pour 1980

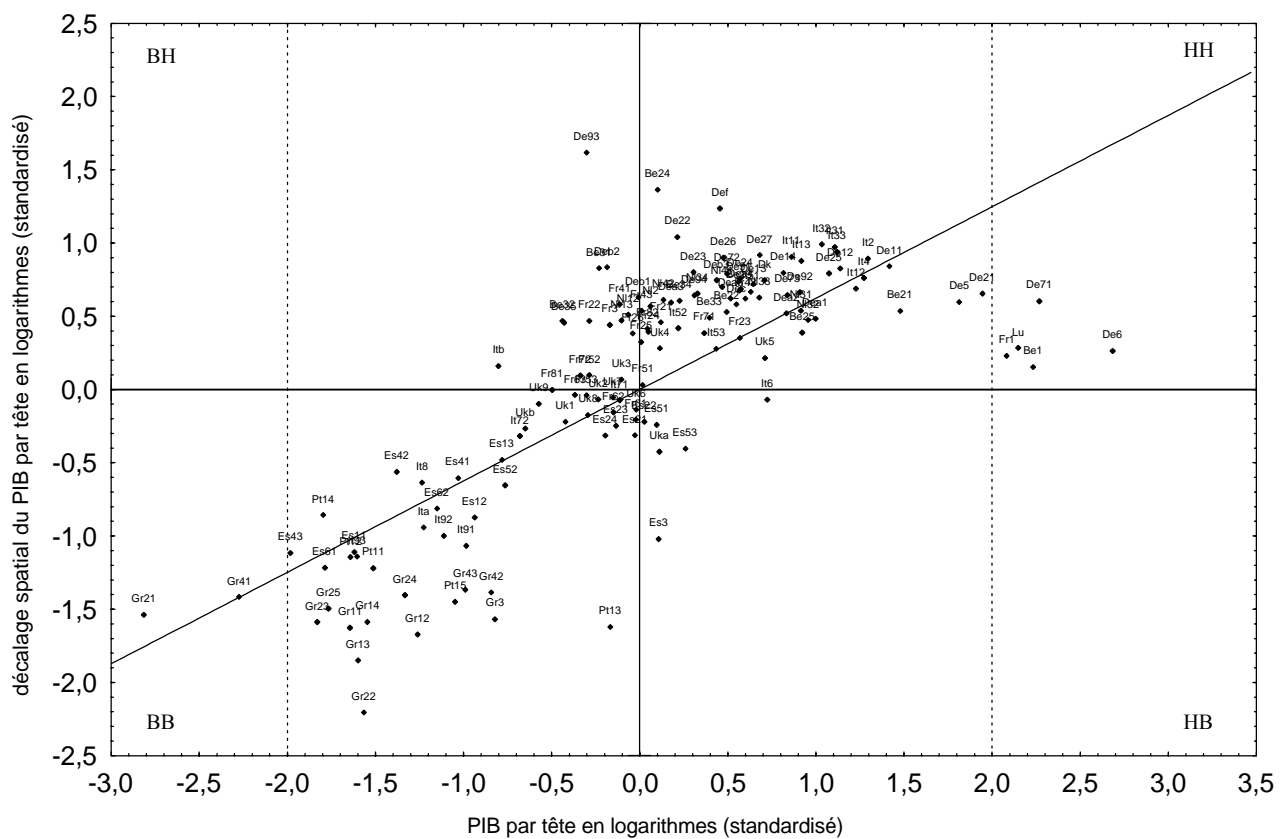


Figure 4 : Diagramme de Moran pour $W(1)$; PIB par tête et en SPA pour 1995

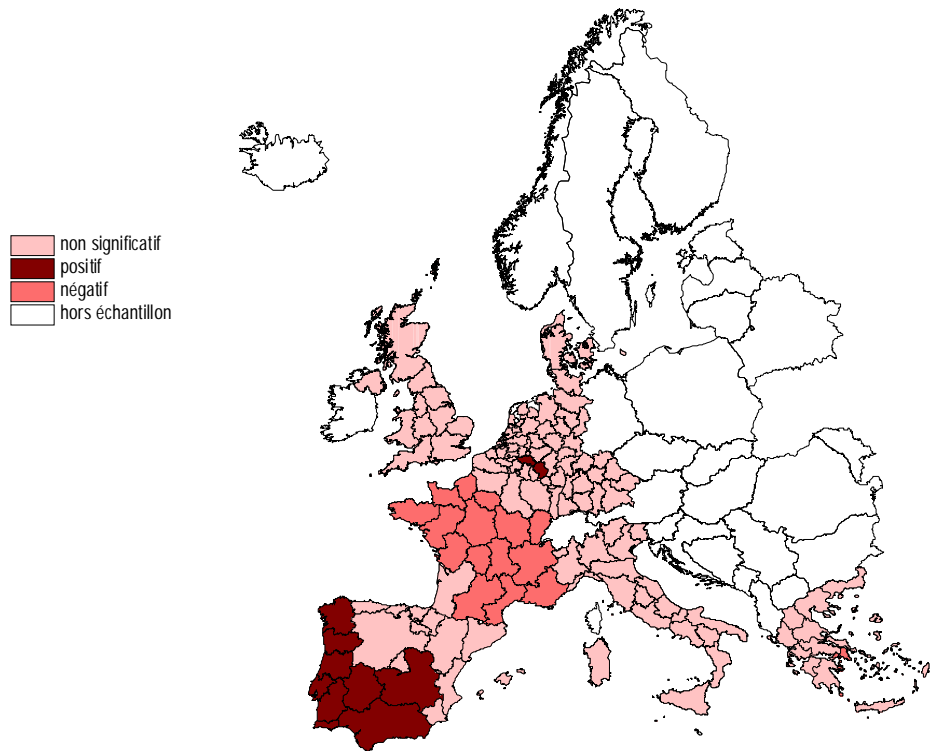


Figure 5 : Statistiques de Getis-Ord significatives pour $W(1)$;
Taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en SPA

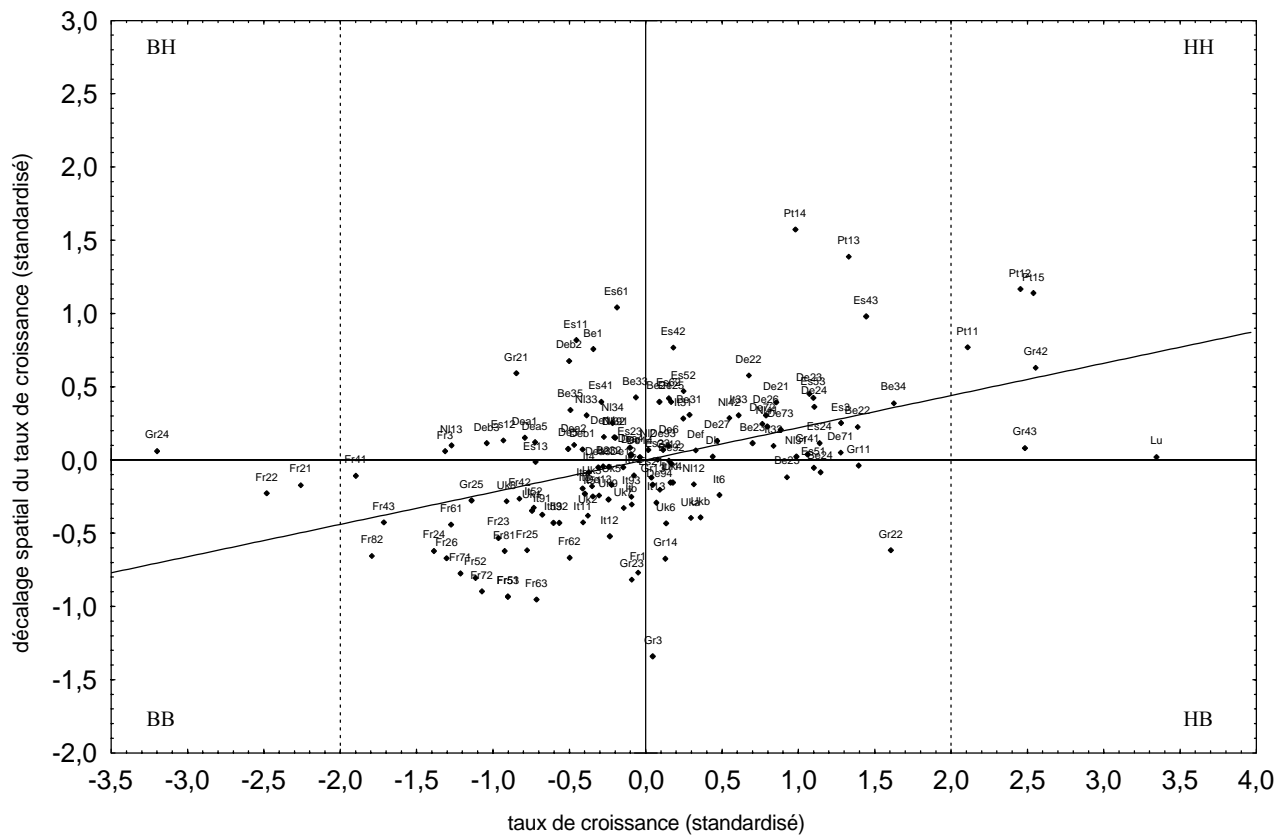


Figure 6 : Diagramme de Moran pour le taux de croissance annuel moyen
entre 1980 et 1995 pour $W(1)$; PIB par tête en SPA

Annexe 3.4

Statistiques $G_i(k)$ et LISA pour $W(10)$; PIB par tête en logarithmes et en Ecus ; 1980-1995

Code	Région	$G_i(10)$ (Ord and Getis 1995)				LISA (Anselin 1995)						
		Signif+	Signif-	Années 5% s.	Années 5% Sidák s.	Signif	HH	BH	BB	HB	Années 5% s.	Années 5% Bonf. s.
	BELGIQUE											
Be1	Bruxelles	0 (0)	0 (0)			1 (0)	1				80	80
Be21	Anvers	2 (0)	0 (0)	80;95		6 (0)	6				80-81;87; 93-95	
Be22	Limburg (B)	5 (0)	0 (0)	80-81;93-95		12 (1)	12 (1)				80-83;85-88;92-95	
Be23	Oost Vlaanderen	2 (0)	0 (0)	80;95		4 (0)	4				80-81;94-95	
Be24	Vlaams Brabant	1 (0)	0 (0)	80		4 (0)	4				80-81;94-95	
Be25	West Vlaanderen	1 (0)	0 (0)	80		5 (0)	5				80-81;93-95	
Be31	Brabant Wallon	0 (0)	0 (0)			5 (0)	2	3			80;95 / 81;93-94	
Be32	Hainaut	2 (0)	0 (0)	80-81		4 (1)	2 (1)	2			80-81 / 94-95	80
Be33	Liège	2 (0)	0 (0)	94-95		4 (0)	4				80;93-95	
Be34	Luxembourg (B)	4 (0)	0 (0)	80;93-95		9 (0)	9				80-81;86-88;92-95	
Be35	Namur	3 (0)	0 (0)	80;94-95		5 (0)	2	3			80-81 / 93-95	
	ALLEMAGNE											
De11	Stuttgart	16 (14)	0 (0)	80-95	82-95	16 (16)	16 (16)				80-95	80-95
De12	Karlsruhe	16 (16)	0 (0)	80-95	80-95	16 (16)	16 (16)				80-95	80-95
De13	Freiburg	16 (16)	0 (0)	80-95	80-95	16 (16)	16 (16)				80-95	80-95
De14	Tübingen	16 (16)	0 (0)	80-95	80-95	16 (16)	16 (16)				80-95	80-95
De21	Oberbayern	16 (5)	0 (0)	80-95	91-95	16 (15)	16 (15)				80-95	80; 82-95
De22	Niederbayern	16 (9)	0 (0)	80-95	87-95	16 (14)	16 (14)				80-95	82-95
De23	Oberpfalz	16 (16)	0 (0)	80-95	80-95	16 (16)	16 (16)				80-95	80-95
De24	Oberfranken	16 (15)	0 (0)	80-95	80;82-95	16 (16)	16 (16)				80-95	80-95
De25	Mittelfranken	16 (14)	0 (0)	80-95	82-95	16 (16)	16 (16)				80-95	80-95
De26	Unterfranken	16 (12)	0 (0)	80-95	83-84;86-95	16 (16)	16 (16)				80-95	80-95
De27	Schwaben	16 (12)	0 (0)	80-95	83-84;86-95	16 (16)	16 (16)				80-95	80-95
De5	Bremen	15 (0)	0 (0)	80-88;90-95		16 (8)	16 (8)				80-95	80-84; 93-95
De6	Hamburg	16 (1)	0 (0)	80-95	95	16 (13)	16 (13)				80-95	80-88; 91; 93-95
De71	Darmstadt	16 (3)	0 (0)	80-95	93-95	16 (15)	16 (15)				80-95	80-91; 93-95
De72	Giessen	16 (9)	0 (0)	80-95	83;86-88;91-95	16 (16)	16 (16)				80-95	80-95
De73	Kassel	16 (3)	0 (0)	80-95	93-95	16 (14)	16 (14)				80-95	80; 82-91; 93-95
De91	Braunschweig	16 (7)	0 (0)	80-95	80;83;91-95	16 (16)	16 (16)				80-95	80-95
De92	Hannover	16 (7)	0 (0)	80-95	80;82-83;92-95	16 (15)	16 (15)				80-95	80-91;93-95
De93	Lüneburg	16 (11)	0 (0)	80-95	80-85;91-95	16 (16)	16 (16)				80-95	80-95
De94	Weser-Ems	16 (3)	0 (0)	80-95	82;94-95	16 (14)	16 (14)				80-95	80-88;90-95
Dea1	Düsseldorf	4 (0)	0 (0)	80;93-95		14 (0)	14				80-90;93-95	
Dea2	Köln	3 (0)	0 (0)	93-95		14 (0)	14				80-81;83;85-95	
Dea3	Münster	11 (0)	0 (0)	80-87;93-95		16 (4)	16 (4)				80-95	80; 83; 94-95
Dea4	Detmold	16 (3)	0 (0)	80-95	93-95	16 (12)	16 (12)				80-95	80-85; 87-88; 91; 93-95

Chapitre 3 – Analyse exploratoire des disparités régionales de PIB par tête

Code	Région	Signif+	Signif-	Années 5% s.	Années 5% Sidák s.	Signif	HH	BH	BB	HB	Années 5% s.	Années 5% Bonf. s.
Dea5	Arnsberg	16 (3)	0 (0)	80-95	93-95	16 (15)	16 (15)				80-95	80-91; 93-95
Deb1	Koblenz	16 (3)	0 (0)	80-95	93-95	16 (8)	16 (8)				80-95	80; 86-88; 91; 93-95
Deb2	Trier	4 (0)	0 (0)	80;93-95		12 (0)	12				80-81;85-90;92-95	
Deb3	Rheinhesen-Pfalz	16 (7)	0 (0)	80-95	86-87;91-95	16 (16)	16 (16)				80-95	80-95
Dec	Saarland	16 (1)	0 (0)	80-95	95	16 (11)	16 (11)				80-95	80-81;85-88;90-91;93-95
Def	Schleswig-Holstein	16 (10)	0 (0)	80-95	80-85;92-95	16 (16)	16 (16)				80-95	80-95
Dk	DANEMARK	16 (3)	0 (0)	80-95	93-95	16 (14)	16 (14)				80-95	80-88;90-91; 93-95
	ESPAGNE											
Es11	Galicie (1)	0 (0)	16 (16)	80-95	80-95	16 (15)			16 (15)		80-95	80-91; 93-95
Es12	Asturias (1)	0 (0)	16 (16)	80-95	80-95	16 (15)			16 (15)		80-95	80-91; 93-95
Es13	Cantabria	0 (0)	14 (4)	80-90 ;93-95	82-85	16 (4)			16 (4)		80-95	82-85
Es21	Pais Vasco	0 (0)	5 (0)	82-86		9 (0)			9		81-87 ;94-95	
Es22	Navarra	0 (0)	3 (0)	83-85		7 (0)			7		81-87	
Es23	La Rioja	0 (0)	8 (0)	80-87		11(1)			11 (1)		80-88;94-95	83
Es24	Aragon	0 (0)	2 (0)	83-84		4 (0)			4		82-85	
Es3	Madrid	0 (0)	16 (12)	80-95	80-88;93-95	16 (13)			14 (13)	2	80-90;93-95/91-92	80-89; 93-95
Es41	Castilla-Leon (1)	0 (0)	16 (11)	80-95	80-88 ;94-95	16 (11)			16 (11)		80-95	80-88; 94-95
Es42	Castilla-la Mancha (1)	0 (0)	14 (9)	80-90;93-95	80-87;94	16 (9)			16 (9)		80-95	80-87; 94
Es43	Extremadura (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Es51	Cataluna	0 (0)	0 (0)			0 (0)						
Es52	Valenciana (1)	0 (0)	12 (6)	80-88;93-95	81-86	13 (6)			13 (6)		80-89;93-95	81-86
Es53	Islas Baleares	0 (0)	6 (0)	80-85		10 (0)			10		80-87;94-95	
Es61	Andalucia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Es62	Murcia (1)	0 (0)	14 (9)	80-90 ;93-95	80-88	16 (11)			16 (11)		80-95	80-88 ;94-95
	FRANCE											
Fr1	Ile-de-France	0 (0)	0 (0)			2 (0)	2				80-81	
Fr21	Champagne-Ardenne	1 (0)	0 (0)	80		9 (0)	9				80-82 ;85-87 ;93-95	
Fr22	Picardie	3 (0)	0 (0)	80-81;95		12 (2)	12 (2)				80-83;85-87;93-95	80-81
Fr23	Haute-Normandie	3 (0)	0 (0)	80-82		10 (2)	10 (2)				80-89	80-81
Fr24	Centre	3 (0)	0 (0)	80-82		8 (0)	8				80-87	
Fr25	Basse-Normandie	4 (0)	0 (0)	80-82 ;85		10 (1)	10 (1)				80-89	81
Fr26	Bourgogne	7 (0)	0 (0)	80-86		11 (3)	11(3)				80-90	80-82
Fr3	Nord-Pas-De-Calais	1 (0)	0 (0)	80		4 (0)	4				80-81;94-95	
Fr41	Lorraine	16 (0)	0 (0)	80-95		16 (7)	16 (7)				80-95	80-81 ;86-87 ; 93-95
Fr42	Alsace	16 (3)	0 (0)	80-95	93-95	16 (15)	16 (15)				80-95	80-91 ; 93-95
Fr43	Franche-Comté	14 (0)	0 (0)	80-93		16 (11)	16 (11)				80-95	80-90
Fr51	Pays de la Loire	3 (0)	0 (0)	80-82		8 (0)	8				80-87	
Fr52	Bretagne	3 (0)	0 (0)	80-82		10 (0)	10				80-89	
Fr53	Poitou-Charentes	2 (0)	0 (0)	81-82		8 (0)	8				80-87	
Fr61	Aquitaine	0 (0)	0 (0)			0 (0)						
Fr62	Midi-Pyrénées	0 (0)	0 (0)			0 (0)						
Fr63	Limousin	3 (0)	0 (0)	80-82		9 (0)	9 (0)				80-88	
Fr71	Rhône-Alpes	0 (0)	0 (0)			2 (0)	2 (0)				86-87	
Fr72	Auvergne	0 (0)	0 (0)			4 (0)	4 (0)				80-82 ;86	

Chapitre 3 – Analyse exploratoire des disparités régionales de PIB par tête

Code	Région	Signif+	Signif-	Années 5% s.	Années 5% Sidák s.	Signif	HH	BH	BB	HB	Années 5% s.	Années 5% Bonf. s.
Fr81	Languedoc-Roussillon	0 (0)	0 (0)			0 (0)	0 (0)					
Fr82	PACA	0 (0)	0 (0)			9 (0)	9 (0)				83-91	
	GRECE											
Gr11	Anatoliki Makedonia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr12	Kentriki Makedonia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr13	Dytiki Makedonia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr14	Thessalia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr21	Ipeiros (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr22	Ionia Nisia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr23	Dytiki Ellada (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr24	Stereia Ellada (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr25	Peloponnisos (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr3	Attiki (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr41	Voreio Aigaio (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr42	Notio Aigaio (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr43	Kriti (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
	Italie											
It11	Piemonte	10 (0)	0 (0)	83-92		14 (4)	14 (4)				81-94	88-91
It12	Valle d'Aosta	13 (0)	0 (0)	80-92		16 (3)	16 (3)				80-95	86 ; 89-90
It13	Liguria	5 (0)	0 (0)	88-92		10 (3)	10 (3)				83-92	89-91
It2	Lombardia	9 (1)	0 (0)	84-92	91	12 (4)	12 (4)				83-94	88-91
It31	Trentino – Alto Adige	13 (3)	0 (0)	83-95	90-92	15 (8)	15 (8)				81-95	84-91
It32	Veneto	9 (1)	0 (0)	84-92	91	12 (4)	12 (4)				83-94	88-91
It33	Friuli – Venezia Giulia	11 (1)	0 (0)	84-94	91	14 (5)	14 (5)				82-95	87-91
It4	Emilia – Romagna	5 (0)	0 (0)	88-92		9 (3)	9 (3)				84-92	89-91
It51	Toscana	4 (0)	0 (0)	89-92		7 (0)	7				86-92	
It52	Umbria	0 (0)	0 (0)			0 (0)						
It53	Marche	0 (0)	0 (0)			0 (0)						
It6	Lazio	0 (0)	0 (0)			0 (0)						
It71	Abruzzo (1)	0 (0)	0 (0)			0 (0)						
It72	Molise (1)	0 (0)	0 (0)			1 (0)			1		95	
It8	Campania (1)	0 (0)	3 (0)	80-81;95		4 (0)			4		80-82;95	
It91	Puglia (1)	0 (0)	16 (6)	80-95	80-82;93-95	16 (6)			16 (6)		80-95	80-82; 93-95
It92	Basilicata (1)	0 (0)	6 (0)	80-82;93-95		6 (1)			6 (1)		80-82;93-95	95
It93	Calabria (1)	0 (0)	16 (7)	80-95	80-82;92-95	16 (6)			16 (6)		80-95	80-82; 93-95
Ita	Sicilia (1)	0 (0)	6 (3)	80-82;93-95	80-81;95	8 (3)			8(3)		80-83;85;93-95	80-81; 95
Itb	Sardegna (1)	0 (0)	0 (0)			0 (0)						
Lu	LUXEMBOURG	0 (0)	0 (0)			6 (0)	6				80-81;86;93-95	
	PAYS-BAS											
NI12	Friesland	12 (0)	0 (0)	80-88;93-95		16 (6)	10 (6)	6			80-84;86-87;93-95/85;88-92	80-84; 95
NI13	Drenthe	14 (0)	0 (0)	80-88;90;92-95		16 (0)	15 (10)	1			80-91;93-95 / 92	80-84; 86-87; 93-95
NI2	Oost Nederland	5 (0)	0 (0)	80-83;95		12 (0)	12				80-88;93-95	
NI31	Utrecht	4 (0)	0 (0)	80-82;95		14 (1)	14 (1)				80-90;93-95	80
NI32	Noord-Holland	5 (0)	0 (0)	80-83;95		12 (1)	12 (1)				80-88;93-95	80

Code	Région	Signif+	Signif-	Années 5% s.	Années 5% Sidák s.	Signif	HH	BH	BB	HB	Années 5% s.	Années 5% Bonf. s.
NI33	Zuid-Holland	2 (0)	0 (0)	80;95		5 (0)	5				80-81;93-95	
NI34	Zeeland	4 (0)	0 (0)	80-81;94-95		7 (1)	7(1)				80-82;87;93-95	80
NI41	Noord-Brabant	4 (0)	0 (0)	80-81;94-95		12 (0)	12				80-83;86-90;93-95	
NI42	Limburg (NL)	3 (0)	0 (0)	80 ;94-95		5 (0)	5				80-81;87 ;93-95	
	PORTUGAL											
Pt11	Norte (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Pt12	Centro (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Pt13	Lisboa e vale do Tejo (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Pt14	Alentejo (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Pt15	Algarve (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
	ROYAUME-UNI											
Uk1	North	0 (0)	0 (0)			0 (0)						
Uk2	Yorkshire and Humberside	0 (0)	0 (0)			0 (0)						
Uk3	East Midlands	0 (0)	0 (0)			0 (0)						
Uk4	East Anglia	0 (0)	0 (0)			1 (0)	1				81	
Uk5	South East	0 (0)	0 (0)			0 (0)						
Uk6	South West	0 (0)	0 (0)			0 (0)						
Uk7	West Midlands	0 (0)	0 (0)			0 (0)						
Uk8	North West	0 (0)	0 (0)			0 (0)						
Uk9	Wales	0 (0)	0 (0)			0 (0)						
Uka	Scotland	0 (0)	0 (0)			0 (0)						
Ukb	Northern Ireland (1)	0 (0)	0 (0)			0 (0)						
	Signif. tot. 5%	686	509	1195		1463	911	15	535	2		
	% versus total de 2208	31,07	23,05	54,12		66,26	41,26	0,68	24,23	0,09		
	% versus signif. total de 5%	57,41	42,59				62,27	1,03	36,57	0,14		
	Signif. tot. 5% Sidák ou Bonf.	228	419	647		(913)	(492)	(0)	421	(0)		
	% versus total de 2208	10,33	18,98	29,30		41,35	22,28	0	19,07	0		
	% versus signif. total de 5%	35,24	64,76				53,89	0	46,11	0		

Notes : Signif+ et Signif- : nombre d'années durant lesquelles les statistiques $G_i(k)$ sont significativement positives ou négatives au niveau de 5% (entre parenthèses au pseudo-niveau de significativité de Sidák). L'inférence statistique est basée sur l'approximation asymptotique normale.

Signif. : nombre d'années durant lesquelles les statistiques LISA sont significatives négatives au niveau de 5% (entre parenthèses au pseudo-niveau de significativité de Bonferroni). L'inférence statistique est basée sur 10 000 permutations.

HH, BH, BB et BH : nombre d'années durant lesquelles la statistique locale est significative et respectivement dans le quadrant HH, BH, BB et BH du diagramme de Moran.

(1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

Annexe 3.5

Statistiques $G_i(k)$ et LISA pour $W(10)$; PIB par tête en logarithmes et en SPA ; 1980-1995

Code	Région	$G_i(10)$ (Ord and Getis 1995)				LISA (Anselin 1995)						
		Signif+	Signif-	Années 5% s.	Années 5% Sidák s.	Signif	HH	BH	BB	HB	Années 5% s.	Années 5% Bonf. s.
	BELGIQUE											
Be1	Bruxelles	0 (0)	0 (0)			0 (0)						
Be21	Anvers	0 (0)	0 (0)			3 (0)	3				93-95	
Be22	Limburg (B)	3 (0)	0 (0)	93-95		7 (0)	5	2			85;92-95 / 80-81	
Be23	Oost Vlaanderen	0 (0)	0 (0)			3 (0)	3				93-95	
Be24	Vlaams Brabant	0 (0)	0 (0)			2 (0)	2				93-94	
Be25	West Vlaanderen	0 (0)	0 (0)			3 (0)	3				93-95	
Be31	Brabant Wallon	0 (0)	0 (0)			3 (0)		3			93-95	
Be32	Hainaut	0 (0)	0 (0)			0 (0)						
Be33	Liège	0 (0)	0 (0)			3 (0)	3				93-95	
Be34	Luxembourg (B)	2 (0)	0 (0)	93-94		4 (0)	4				92-95	
Be35	Namur	3 (0)	0 (0)	80;94-95		5 (0)	2	3			80-81 / 93-95	
	ALLEMAGNE											
De11	Stuttgart	16 (5)	0 (0)	80-95	91-95	16 (14)	16 (14)				80-95	81;83-95
De12	Karlsruhe	16 (15)	0 (0)	80-95	81-95	16 (16)	16 (16)				80-95	80-95
De13	Freiburg	16 (12)	0 (0)	80-95	80-88;90-92	16 (16)	16 (16)				80-95	80-95
De14	Tübingen	16 (16)	0 (0)	80-95	80-95	16 (16)	16 (16)				80-95	80-95
De21	Oberbayern	16 (0)	0 (0)	80-95		16 (5)	16 (5)				80-95	91-95
De22	Niederbayern	16 (6)	0 (0)	80-95	90-95	16 (15)	13 (13)	3 (2)			83-95 / 80-82	83-95 / 81-82
De23	Oberpfalz	16 (15)	0 (0)	80-95	81-95	16 (16)	13 (13)	3 (3)			81;84-95 / 80;82-83	81;84-95 / 80;82-83
De24	Oberfranken	16 (6)	0 (0)	80-95	90-95	16 (13)	16 (13)				80-95	83-95
De25	Mittelfranken	16 (6)	0 (0)	80-95	90-95	16 (16)	16 (16)				80-95	80-95
De26	Unterfranken	16 (5)	0 (0)	80-95	91-95	16 (13)	16 (13)				80-95	83-95
De27	Schwaben	16 (6)	0 (0)	80-95	90-95	16 (16)	16 (16)				80-95	80-95
De5	Bremen	0 (0)	0 (0)			11 (0)	11				80-85;91-95	
De6	Hamburg	0 (0)	0 (0)			11 (0)	11				80-85;91-95	
De71	Darmstadt	5 (0)	0 (0)	85-86;91-92;94		16 (0)	16				80-95	
De72	Giessen	16 (0)	0 (0)	80-95		16 (3)	16 (3)				80-95	86;91-92
De73	Kassel	5 (0)	0 (0)	91-95		16 (0)	16				80-95	
De91	Braunschweig	16 (0)	0 (0)	80-95		16 (6)	16 (6)				80-95	80;91-95
De92	Hannover	13 (0)	0 (0)	80-86;90-95		16 (2)	16 (2)				80-95	80-81
De93	Lüneburg	16 (0)	0 (0)	80-95		16 (11)		16 (11)			80-95	80-85;91-95
De94	Weser-Ems	12 (0)	0 (0)	80-86;91-95		16 (1)	16 (1)				80-95	81
Dea1	Düsseldorf	0 (0)	0 (0)			3 (0)	3				93-95	
Dea2	Köln	0 (0)	0 (0)			3 (0)	3				93-95	
Dea3	Münster	0 (0)	0 (0)			4 (0)	4				80-81;94-95	
Dea4	Detmold	7 (0)	0 (0)	80-81;91-95		13 (0)	13				80-85;88;90-95	

Chapitre 3 – Analyse exploratoire des disparités régionales de PIB par tête

Code	Région	Signif+	Signif-	Années 5% s.	Années 5% Sidák s.	Signif	HH	BH	BB	HB	Années 5% s.	Années 5% Bonf. s.
Dea5	Arnsberg	11 (0)	0 (0)	80-81;84-86;90-95		16 (0)	16				80-95	
Deb1	Koblenz	11 (0)	0 (0)	85-95		16 (0)	14	2			80-92;94 / 93;95	
Deb2	Trier	0 (0)	0 (0)			0 (0)						
Deb3	Rheinhesse-Pfalz	16 (2)	0 (0)	80-95	91-92	16 (16)	16 (16)				80-95	80-95
Dec	Saarland	16 (0)	0 (0)	80-95		16 (0)	16				80-95	
Def	Schleswig-Holstein	14 (0)	0 (0)	80-87;90-95		16 (6)	16 (6)				80-95	80-85
Dk	DANEMARK	11 (0)	0 (0)	80-85;91-95		16 (0)	16				80-95	
	ESPAGNE											
Es11	Galicie (1)	0 (0)	16 (13)	80-95	80-92	16 (16)			16 (16)		80-95	80-95
Es12	Asturias (1)	0 (0)	16 (15)	80-95	80-92;94-95	16 (16)			16 (16)		80-95	80-95
Es13	Cantabria	0 (0)	1 (0)	85		13 (0)			13		80-88 ;90;92;94-95	
Es21	Pais Vasco	0 (0)	0 (0)			0 (0)						
Es22	Navarra	0 (0)	0 (0)			0 (0)						
Es23	La Rioja	0 (0)	0 (0)			0 (0)						
Es24	Aragon	0 (0)	0 (0)			0 (0)						
Es3	Madrid	0 (0)	16 (8)	80-95	80-87	16 (9)				9(2)	80-86 / 87-95	80-86 / 87-88
Es41	Castilla-Leon (1)	0 (0)	16 (1)	80-95	85	16 (6)			16 (6)		80-95	81-86
Es42	Castilla-la Mancha (1)	0 (0)	15 (1)	80-90;92-95	82	16 (3)			16 (3)		80-95	81-83
Es43	Extremadura (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Es51	Cataluna	0 (0)	0 (0)			0 (0)						
Es52	Valenciana (1)	0 (0)	7 (0)	80-86		11 (0)			11		80-88;94-95	
Es53	Islas Baleares	0 (0)	0 (0)			2 (0)			1	1	81 / 82	
Es61	Andalucia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Es62	Murcia (1)	0 (0)	15 (4)	80-90;92-95	80-83	16 (6)			16 (6)		80-95	80-83 ;85-86
	FRANCE											
Fr1	Ile-de-France	0 (0)	0 (0)			0 (0)						
Fr21	Champagne-Ardenne	0 (0)	0 (0)			0 (0)						
Fr22	Picardie	3 (0)	0 (0)	80-82		15 (0)	12	3			80-90;92 / 93-95	
Fr23	Haute-Normandie	2 (0)	0 (0)	80;82		9 (0)	9				80-85 ;87;89-90	
Fr24	Centre	0 (0)	0 (0)			5 (0)	5				80-84	
Fr25	Basse-Normandie	0 (0)	0 (0)			10 (0)	10				80-89	
Fr26	Bourgogne	5 (0)	0 (0)	80-84		11 (0)	11				80-90	
Fr3	Nord-Pas-De-Calais	0 (0)	0 (0)			0 (0)						
Fr41	Lorraine	13 (0)	0 (0)	80-92		16 (0)	13	3			80-92 / 93-95	
Fr42	Alsace	16 (0)	0 (0)	80-95		16 (13)	16 (13)				80-95	80-92
Fr43	Franche-Comté	13 (0)	0 (0)	80-92		16 (4)	16 (4)				80-95	80-83
Fr51	Pays de la Loire	0 (0)	0 (0)			5 (0)	5				80-84	
Fr52	Bretagne	0 (0)	0 (0)			8 (0)	8				80-87	
Fr53	Poitou-Charentes	0 (0)	0 (0)			2 (0)	2				82-83	
Fr61	Aquitaine	0 (0)	0 (0)			0 (0)						
Fr62	Midi-Pyrénées	0 (0)	0 (0)			0 (0)						
Fr63	Limousin	0 (0)	0 (0)			3 (0)	2	1			82-83 / 84	
Fr71	Rhône-Alpes	0 (0)	0 (0)			5 (0)	5				80-84	
Fr72	Auvergne	0 (0)	0 (0)			0 (0)						

Chapitre 3 – Analyse exploratoire des disparités régionales de PIB par tête

Code	Région	Signif+	Signif-	Années 5% s.	Années 5% Sidák s.	Signif	HH	BH	BB	HB	Années 5% s.	Années 5% Bonf. s.
Fr81	Languedoc-Roussillon	0 (0)	0 (0)			0 (0)						
Fr82	PACA	6 (0)	0 (0)	80-84;89		12 (0)	12				80-91	
	GRECE											
Gr11	Anatoliki Makedonia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr12	Kentriki Makedonia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr13	Dytiki Makedonia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr14	Thessalia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr21	Ipeiros (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr22	Ionia Nisia (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr23	Dytiki Ellada (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr24	Stereia Ellada (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr25	Peloponnisos (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr3	Attiki (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr41	Voreio Aigaio (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr42	Notio Aigaio (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Gr43	Kriti (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
	ITALIE											
It11	Piemonte	16 (11)	0 (0)	80-95	80-85;87-91	16 (16)	16 (16)				80-95	80-95
It12	Valle d'Aosta	13 (0)	0 (0)	80-92		16 (8)	16 (8)				80-95	80-85 ;88-89
It13	Liguria	16 (8)	0 (0)	80-95	80-83;85;87-89	16 (16)	16 (16)				80-95	80-95
It2	Lombardia	16 (16)	0 (0)	80-95	80-95	16 (16)	16 (16)				80-95	80-95
It31	Trentino – Alto Adige	16 (16)	0 (0)	80-95	80-95	16 (16)	16 (16)				80-95	80-95
It32	Veneto	16 (16)	0 (0)	80-95	80-95	16 (16)	16 (16)				80-95	80-95
It33	Friuli – Venezia Giulia	16 (15)	0 (0)	80-95	81-95	16 (16)	16 (16)				80-95	80-95
It4	Emilia – Romagna	16 (13)	0 (0)	80-95	80-82;85-92;94-95	16 (16)	16 (16)				80-95	80-95
It51	Toscana	16 (0)	0 (0)	80-95		16 (12)	16 (12)				80-95	80-82;85-92;95
It52	Umbria	0 (0)	0 (0)			0 (0)						
It53	Marche	0 (0)	0 (0)			0 (0)						
It6	Lazio	0 (0)	0 (0)			0 (0)						
It71	Abruzzo (1)	0 (0)	0 (0)			0 (0)						
It72	Molise (1)	0 (0)	0 (0)			0 (0)						
It8	Campania (1)	0 (0)	0 (0)			0 (0)						
It91	Puglia (1)	0 (0)	16 (10)	80-95	80;82-83;89-95	16 (13)			16 (13)		80-95	80-85;89-95
It92	Basilicata (1)	0 (0)	3 (0)	93-95		8 (0)			8		80-82;90;92-95	
It93	Calabria (1)	0 (0)	16 (16)	80-95	80-95	16 (15)			16 (15)		80-95	80-86;88-95
Ita	Sicilia (1)	0 (0)	14 (0)	80-83;85-86;88-95		16 (0)			16		80-95	
Itb	Sardegna (1)	0 (0)	0 (0)			0 (0)						
Lu	LUXEMBOURG	0 (0)	0 (0)			0 (0)						
	PAYS-BAS											
NI12	Friesland	3 (0)	0 (0)	80-81;95		14 (0)		14			80-87;90-95	
NI13	Drenthe	3 (0)	0 (0)	80-81;95		16 (0)	8	8			80-86;90 / 87-89;91-95	
NI2	Oost Nederland	0 (0)	0 (0)			0 (0)						
NI31	Utrecht	0 (0)	0 (0)			3 (0)	3				90;94-95	
NI32	Noord-Holland	0 (0)	0 (0)			1 (0)	1				95	

Chapitre 3 – Analyse exploratoire des disparités régionales de PIB par tête

Code	Région	Signif+	Signif-	Années 5% s.	Années 5% Sidák s.	Signif	HH	BH	BB	HB	Années 5% s.	Années 5% Bonf. s.
NI33	Zuid-Holland	0 (0)	0 (0)			2 (0)	2				94-95	
NI34	Zeeland	3 (0)	0 (0)	93-95		6 (0)	6				80;90;92-95	
NI41	Noord-Brabant	3 (0)	0 (0)	93-95		8 (0)	8				80;88-95	
NI42	Limburg (NL)	0 (0)	0 (0)			3 (0)	2	1			94-95 / 93	
	PORTUGAL											
Pt11	Norte (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Pt12	Centro (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Pt13	Lisboa e vale do Tejo (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Pt14	Alentejo (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
Pt15	Algarve (1)	0 (0)	16 (16)	80-95	80-95	16 (16)			16 (16)		80-95	80-95
	ROYAUME-UNI											
Uk1	North	0 (0)	0 (0)			0 (0)						
Uk2	Yorkshire and Humberside	0 (0)	0 (0)			0 (0)						
Uk3	East Midlands	0 (0)	0 (0)			0 (0)						
Uk4	East Anglia	0 (0)	0 (0)			0 (0)						
Uk5	South East	0 (0)	0 (0)			0 (0)						
Uk6	South West	0 (0)	0 (0)			0 (0)						
Uk7	West Midlands	0 (0)	0 (0)			0 (0)						
Uk8	North West	0 (0)	0 (0)			0 (0)						
Uk9	Wales	0 (0)	0 (0)			0 (0)						
Uka	Scotland	0 (0)	0 (0)			0 (0)						
Ukb	Northern Ireland (1)	0 (0)	0 (0)			0 (0)						
	Signif. tot. 5%	564	471	1035		1289	728	63	488	10		
	% versus total de 2208	25,54	21,33	46,87		58,38	32,97	2,85	22,10	0,45		
	% versus signif. total de 5%	54,49	45,51				56,48	4,89	37,86	0,78		
	Signif. tot. 5% Sidák ou Bonf.	189	388	577		(754)	(334)	(16)	(402)	(2)		
	% versus total de 2208	8,56	17,57	26,13		34,15	15,13	0,72	18,21	0,09		
	% versus signif. total de 5%	32,76	67,24				44,30	2,12	53,32	0,27		

Notes : Signif+ et Signif- : nombre d'années durant lesquelles les statistiques $G_i(k)$ sont significativement positives ou négatives au niveau de 5% (entre parenthèses au pseudo-niveau de significativité de Sidák). L'inférence statistique est basée sur l'approximation asymptotique normale.

Signif. : nombre d'années durant lesquelles les statistiques LISA sont significatives négatives au niveau de 5% (entre parenthèses au pseudo-niveau de significativité de Bonferroni). L'inférence statistique est basée sur 10 000 permutations.

HH, BH, BB et BH : nombre d'années durant lesquelles la statistique locale est significative et respectivement dans le quadrant HH, BH, BB et BH du diagramme de Moran.

(1) dénote les régions éligibles à l'objectif 1 des fonds structurels sur la période 1989-1995.

CHAPITRE 4

Analyse économétrique spatiale de la β -convergence entre les régions européennes

Une partie des résultats empiriques présentés dans ce chapitre est tirée des articles suivants :

Baumont C., Ertur C., Le Gallo J. (2002b) Spatial convergence clubs and the European growth process, 1980-1995, dans : Fingleton B. (ed.), *European Regional Growth*, Springer, Berlin, à paraître.

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

Chapitre 4

Analyse économétrique spatiale de la β -convergence entre les régions européennes

La convergence des régions européennes a souvent été étudiée dans la littérature récente de macroéconomie et de science régionale. Il apparaît que la vitesse de convergence entre les régions européennes est très faible malgré le processus d'intégration européenne et le taux de croissance plus élevé de certaines régions pauvres (Barro et Sala-I-Martin, 1991, 1995 ; Abraham et Von Rompuy, 1995 ; Armstrong, 1995a, 1995b ; Sala-I-Martin, 1996b). En outre, sur la période 1980-1995, nous avons montré dans le chapitre 3 que la distribution géographique des régions européennes est fortement concentrée suivant un schéma de polarisation Nord-Sud. Ces observations peuvent être reliées à différents résultats des théories de l'économie géographique car ces dernières aboutissent le plus souvent à des schémas de type centre-périphérie et de développement inégal. Par ailleurs, elles mettent en lumière le rôle joué par les effets de débordement géographiques dans les mécanismes de croissance.

Ces résultats ont trois implications majeures sur les études empiriques de croissance régionale. Tout d'abord, ils impliquent que les régions similaires en termes de PIB par tête ont tendance à se regrouper. En conséquence, les estimations économétriques basées sur des données géographiques doivent tenir compte du fait que les phénomènes économiques peuvent être spatialement autocorrélés. Ensuite, si les effets de débordement géographiques influencent effectivement le processus de croissance, il est intéressant d'estimer leurs impacts et la façon dont les performances économiques des régions interagissent entre elles. Enfin, la polarisation de l'économie européenne suggère l'existence éventuelle de différents régimes de croissance parmi les régions européennes et doit être testée.

Ces différentes questions empiriques sont traitées dans le cadre du modèle de β -convergence entre les régions européennes. Nous considérons que la prise en compte de l'autocorrélation et de l'hétérogénéité spatiales constitue une voie alternative permettant d'améliorer ce modèle.

Le point de départ de notre argumentation est économétrique. En effet, la présence de l'autocorrélation et l'hétérogénéité spatiales entre les observations ne permet pas d'obtenir une inférence statistique fiable lorsqu'elle est basée sur les Moindres Carrés Ordinaires (MCO). Par conséquent, ces deux effets devraient être systématiquement testés dans les modèles mobilisant des données géographiques de la même façon que la corrélation temporelle est systématiquement testée dans les modèles de séries temporelles. En testant et

en estimant le modèle de β -convergence avec les techniques de l'économétrie spatiale, nous évitons les biais dans l'inférence statistique dus à l'autocorrélation spatiale et nous obtenons des estimations plus fiables de la vitesse de convergence.

Parallèlement, notre démarche permet d'améliorer le modèle de β -convergence dans trois directions. Tout d'abord, sur le plan méthodologique, l'autocorrélation spatiale conduit à une spécification minimale de β -convergence conditionnelle. Face à une approche qui consisterait à rechercher et à trouver l'ensemble des déterminants de la différenciation régionale des états réguliers en incluant des variables explicatives additionnelles dans un modèle de β -convergence conditionnelle, nous considérons à la place que l'autocorrélation spatiale absorbe tous ces effets (Fingleton, 1999). Ensuite, nous sommes en mesure d'estimer l'ampleur des effets de débordement géographiques sur le processus de croissance régionale et nous pouvons décrire le processus de diffusion spatiale correspondant. Enfin, le traitement de l'hétérogénéité spatiale, sous la forme de régimes, permet de capter la polarisation Nord-Sud observée dans les régions européennes et de l'interpréter comme des clubs de convergence spatiaux.

Il apparaît ainsi que l'économétrie spatiale est un outil important pour souligner les liens entre les théories de la croissance et les théories de l'économie géographique. Les études empiriques que nous présentons dans ce chapitre s'inscrivent dans ce contexte mais n'ont pas pour ambition de tester formellement les modèles d'économie géographique. En effet, les modèles estimés ne sont pas des équations réduites tirées d'un modèle de l'économie géographique. Il s'agit simplement de modèles de β -convergence dans lesquels l'autocorrélation et l'hétérogénéité spatiales sont introduites. Notre objectif est plutôt d'évaluer l'importance du rôle joué par la géographie dans le processus de croissance et de convergence. Par ailleurs, les effets de débordement géographiques que nous estimons sont plus généraux que les interactions de marché ou les externalités de connaissance traitées dans les modèles d'économie géographique. En d'autres termes, nous proposons une méthode d'appréciation des effets de débordement géographiques, mais nous n'entrons pas dans la « boîte noire ». Enfin, pour des raisons techniques, les spécifications que nous utilisons supposent que les effets de débordement géographiques sont globaux. Elles ne nous permettent donc pas de déterminer le caractère local ou global de ces effets, qui sont au centre des modèles de la synthèse géographie-croissance (Kubo, 1995 ; Martin et Ottaviano, 1999).

L'étude empirique de β -convergence sur les régions européennes réalisée dans ce chapitre illustre ces différents points en suivant une progression qui consiste à s'intéresser dans un premier temps à l'autocorrélation spatiale (section 41) et dans un second temps à l'hétérogénéité spatiale (section 42). Cette procédure permet d'illustrer les techniques et les apports propres à chaque effet spatial⁶⁶.

⁶⁶ Les résultats des estimations et des tests ont été obtenus à l'aide des logiciels SpaceStat 1.90 (Anselin, 1999) et Gauss 3.2.26 (Aptech). Les cartes ont été réalisées avec Arcview 3.2 (Esri).

SECTION 41 : **β -CONVERGENCE ET AUTOCORRELATION SPATIALE**

Le but de cette section est de montrer comment l'autocorrélation spatiale est introduite dans le modèle de β -convergence. Pour différentes raisons que nous explicitons dans un premier temps, nous nous situons dans le cadre du modèle de β -convergence absolue (paragraphe 411). Nous montrons alors comment l'autocorrélation spatiale est introduite dans ce modèle à l'aide des spécifications les plus communément utilisées (paragraphe 412). Ensuite, nous détaillons les résultats empiriques obtenus à l'aide de cette méthodologie. En utilisant l'échantillon de 138 régions sur la période 1980-1995, nous montrons que le modèle de β -convergence absolue est mal spécifié en raison de l'omission d'une autocorrélation spatiale des erreurs. Après avoir déterminé le modèle spatial le plus approprié, à l'aide des tests de spécification et des règles de décision suggérées par Anselin et Florax (1995b), nos estimations font apparaître un effet de débordement géographique significatif (paragraphe 413). Ce modèle implique enfin un processus de diffusion spatiale des chocs aléatoires que nous évaluons par simulation (paragraphe 414).

411 Raisons et implications du choix du modèle de β -convergence absolue

Notre point de départ est le modèle de β -convergence absolue en coupe transversale estimé par les MCO, que nous avons présenté dans le chapitre 1 (paragraphe 121.1, p. 37) :

$$g_T = \alpha S + \beta y_0 + \varepsilon \quad [4.1a]$$

$$\varepsilon \sim N(0, \sigma_\varepsilon^2 I) \quad [4.1b]$$

où g_T est le vecteur de dimension $(N,1)$ contenant les taux de croissance annuels moyens des PIB par tête des régions entre la date 0 et la date T , y_0 est le vecteur des niveaux de PIB par tête en logarithmes à la date initiale, S est le vecteur somme de dimension $(N,1)$, α et β sont les paramètres inconnus à estimer.

Le manque de données constitue toujours un problème important dans le contexte européen. Des données harmonisées et fiables permettant des comparaisons régionales sont rares, en particulier pour le début de la période. Ainsi, il est impossible de pouvoir associer à chaque région de l'échantillon une série de variables de conditionnement homogènes pour pouvoir estimer un modèle de β -convergence conditionnelle. Cette situation contraste avec le

nombre de variables disponibles pour les études internationales (Summers et Heston, 1988). Ainsi, il n'est pas possible d'estimer, sur notre échantillon, un modèle de β -convergence conditionnelle avec capital humain dans le cadre proposé par Mankiw *et al.* (1992).

Face à cette absence d'informations sur les variables de conditionnement, de nombreux auteurs ont utilisé, dans le cadre des régions européennes, des variables muettes nationales (Armstrong, 1995a ; Neven et Gouyette, 1994, 1995 ; Fagerberg et Verspagen, 1996 ; Button et Pentecost, 1999 ; Jean-Pierre, 1999). Ces dernières sont introduites dans le but de capter les différences dans les états réguliers des régions européennes. Les auteurs qui introduisent des variables muettes nationales dans le modèle de β -convergence supposent alors que chaque pays est caractérisé par un état régulier différent et que les régions convergent vers leur état régulier *national* respectif. Dans ce cas, ainsi que le souligne Armstrong (1995b), la vitesse de convergence estimée à partir du coefficient β dans [4.1a] mesure la convergence « intra-pays », c'est-à-dire la convergence entre les régions d'un même pays, et non plus la convergence entre toutes les régions de l'échantillon, quelle que soit leur appartenance nationale.

Nous avons choisi de ne pas appliquer cette méthodologie pour deux raisons. D'une part, l'introduction de variables muettes nationales implique des états réguliers nationaux différents mais une vitesse de convergence constante vers ces états. En d'autres termes, chaque région converge vers son état régulier national au même rythme que l'ensemble des autres régions de l'échantillon. Cette procédure apparaît alors être une procédure hybride et il serait sans doute plus satisfaisant dans ce cas de ne pas contraindre la vitesse de convergence à être constante pour toutes les régions (Barro et Sala-I-Martin, 1995). D'autre part, notre problématique est l'étude de la convergence entre toutes les régions européennes, ce qui exclut l'utilisation de variables muettes nationales.

D'autres auteurs ont utilisé, au niveau international, des variables muettes frontalières, c'est-à-dire des variables muettes égales à 1 si deux pays ont une frontière commune et 0 sinon (Chua, 1993 ; Barro et Sala-I-Martin, 1995 ; Ades et Chua, 1997 ; Easterly et Levine, 1997). Cette procédure n'est pas pertinente pour les régions européennes puisqu'elle conduirait à exclure les îles de notre échantillon et à supposer que les régions grecques et les régions anglaises n'interagissent pas avec les régions du continent (cf. paragraphe 313 du chapitre 3, p. 214). Ce type de variables muettes conduit en outre à n'avoir qu'une vue très restrictive des effets liés à la distance puisque seul le passage de la frontière compte. En revanche, l'utilisation des techniques de l'économétrie spatiale, combinées à la mobilisation de matrices de poids basées sur la distance, permet de capter l'effet de cette dernière sur l'influence des effets de débordement géographiques.

Dans ce contexte, nous ne cherchons pas à trouver l'ensemble des déterminants de la différenciation des états réguliers régionaux en incluant diverses variables explicatives dans le modèle de β -convergence. Au contraire, le point de départ de notre analyse est le modèle de

β -convergence absolue dans lequel l'autocorrélation spatiale est introduite à l'aide de diverses spécifications. Dans le cadre des régions européennes, l'introduction de l'autocorrélation spatiale est particulièrement pertinente. En effet, les observations seront vraisemblablement géographiquement corrélées en raison de divers processus qui connectent les zones, comme les processus de diffusion ou de commerce. Plusieurs facteurs économiques, stimulés par le processus d'intégration régionale, comme la mobilité de la force de travail, la mobilité du capital, la diffusion de la technologie et de la connaissance, la baisse des coûts de transport et de transaction sont à l'œuvre et affectent directement les interactions régionales. Ces effets sont particulièrement importants lorsqu'on travaille sur des régions à un niveau aussi désagrégé que le niveau NUTS 2. En effet, les régions à ce niveau ne sont pas des unités économiques auto-contenues et sont parfois affectées par les flux de migrants journaliers.

Tous ces éléments induisent une organisation particulière des activités dans l'espace. Ils sont eux-mêmes engagés dans des processus complexes, sont difficiles à évaluer et ne sont probablement pas tous identifiés. Par conséquent, nous ne cherchons ni à introduire ces facteurs explicatifs dans l'équation de convergence comme variables explicatives, ni à les introduire dans la matrice de poids, qui doit rester purement exogène (cf. chapitre 2, paragraphe 212.2, p. 103). Au contraire, nous considérons que la dépendance spatiale capte tous ces effets simultanément et absorbe toutes les différences dans les états réguliers régionaux (Fingleton, 1999).

Un dernier aspect concerne les problèmes d'estimation induits par l'utilisation des MCO dans le cadre des modèles de β -convergence. Comme nous l'avons souligné dans le chapitre 1 (paragraphe 121.22, p. 42), ce modèle est soumis à de nombreuses critiques, concernant la corrélation entre le PIB initial et le terme d'erreur ou encore le biais dû aux variables omises. Nous sommes conscients des limites induites par l'utilisation des MCO comme point de départ de notre analyse. Cependant, il est délicat d'utiliser des modèles prenant en compte simultanément la dimension spatiale et la dimension temporelle. Compte tenu des caractéristiques de notre échantillon, pour lequel le nombre de régions ($N = 138$) est très supérieur au nombre de périodes ($T = 15$), il n'est pas possible d'estimer une matrice des variances-covariances contemporaine dans le cadre d'un modèle SUR. En outre, les propriétés des méthodes d'estimation des panels dynamiques ne sont pas bien connues en échantillon fini, alors que l'introduction de l'autocorrélation spatiale dans ce cadre particulier n'a pas encore été effectuée. Finalement, puisque notre but est d'estimer les effets de débordement géographiques, nous avons donc choisi de considérer le modèle de β -convergence absolue estimé par les MCO comme point de départ et de l'amender en fonction des tests d'autocorrélation spatiale effectués.

412 L'autocorrélation spatiale dans le modèle de β -convergence absolue

L'introduction de l'autocorrélation spatiale dans le modèle de β -convergence permet d'interpréter les différents modèles obtenus comme des modèles de β -convergence conditionnelle minimaux et de mettre en évidence des effets de débordement géographiques. Différents types de modèles peuvent être utilisés pour traiter l'autocorrélation spatiale dans les observations (cf. paragraphe 221 du chapitre 2, p. 120). Appliqués au modèle de β -convergence, ils donnent lieu à différentes interprétations des coefficients associés aux effets de débordement géographiques. Nous étudions le cas des 3 modèles les plus communément utilisés : le modèle avec variable endogène décalée (paragraphe 412.1), le modèle spatial régressif croisé (paragraphe 412.2) et enfin le modèle avec autocorrélation spatiale des erreurs (paragraphe 412.3).

412.1 Le modèle avec variable endogène décalée

L'autocorrélation spatiale des observations est captée à l'aide de la variable endogène décalée des taux de croissance annuels moyens Wg_T , introduite dans le modèle [4.1]. Le modèle structurel s'écrit dans ce cas de la façon suivante :

$$g_T = \alpha S + \beta y_0 + \rho Wg_T + u \quad [4.2a]$$

$$u \sim N(0, \sigma^2 I) \quad [4.2b]$$

où ρ est le paramètre spatial autorégressif indiquant l'ampleur de l'interaction entre les observations, compte tenu du schéma spatial introduit d'une façon exogène dans la matrice de poids standardisée W . Pour les autres termes, la même notation que pour le modèle [4.1] est utilisée. La variable spatiale endogène décalée Wg_T est le vecteur de décalage spatial : pour une région i du vecteur g_T , la ligne correspondante de Wg_T contient la moyenne pondérée des taux de croissance annuels moyens dans les régions voisines.

Rappelons que l'estimation de ce modèle par les MCO produit des estimateurs non-convergeants à cause de la présence du régresseur stochastique Wg_T . En effet, ce dernier est toujours corrélé avec le vecteur d'erreurs u , même si les erreurs sont identiquement et indépendamment distribuées. Ce modèle doit donc être estimé par la méthode du maximum de vraisemblance ou par la méthode des variables instrumentales.

Cette spécification peut être interprétée de deux façons. Du point de vue de la convergence, elle fournit des informations sur la nature de la convergence à l'œuvre entre les régions de l'échantillon à travers le paramètre β , lorsque les effets spatiaux sont pris en compte. Du point de vue de l'économie géographique, elle permet de capter un effet de

débordement géographique : elle indique la façon dont le taux de croissance du PIB par tête dans une région est affecté par celui des régions voisines à travers le paramètre ρ , après avoir conditionné sur les niveaux de PIB par tête initiaux.

En outre, ce modèle peut se réécrire de la façon suivante sous forme réduite :

$$(I - \rho W)g_T = \alpha S + \beta y_0 + u \quad [4.3a]$$

soit
$$g_T = \alpha(I - \rho W)^{-1}S + \beta(I - \rho W)^{-1}y_0 + (I - \rho W)^{-1}u \quad [4.3b]$$

Cette expression indique, qu'en moyenne conditionnelle, le taux de croissance annuel moyen d'une région n'est pas seulement affecté par son propre niveau de PIB par tête initial mais également par ceux de toutes les régions voisines, à travers la transformation spatiale inverse $(I - \rho W)^{-1}$. Il s'agit du multiplicateur spatial ou de l'effet d'interaction globale. Concernant le processus des erreurs, cette expression indique qu'un choc aléatoire dans une région spécifique affecte non seulement le taux de croissance annuel moyen de cette région, mais a également un impact sur le taux de croissance de toutes les autres régions à travers la même transformation spatiale inverse.

412.2 Le modèle régressif croisé

Une autre façon de traiter l'autocorrélation spatiale est d'introduire des variables exogènes décalées :

$$g_T = \alpha S + \beta y_0 + WZ\gamma + u \quad [4.4a]$$

$$u \sim N(0, \sigma^2 I) \quad [4.4b]$$

où Z est une matrice de dimension (N, h) comportant les observations sur h variables explicatives. L'influence des h variables spatiales décalées est reflétée par le paramètre γ . Cette spécification générale permet de traiter explicitement un autre type d'effet de débordement géographique que le modèle précédent. En outre, elle peut être interprétée comme un *modèle de convergence conditionnelle intégrant des variables d'environnement spatial* qui peuvent affecter les taux de croissance. L'ensemble des variables explicatives dans Z peut inclure, ou non, la variable y_0 . Ce modèle peut s'estimer par les MCO.

Un cas particulier intéressant apparaît lorsque Z inclut simplement la variable des PIB par tête initiaux en logarithmes, y_0 . On obtient alors :

$$g_T = \alpha S + \beta y_0 + \gamma W y_0 + u \quad [4.5a]$$

$$u \sim N(0, \sigma^2 I) \quad [4.5b]$$

Dans cette spécification, seul le taux de croissance annuel moyen spatialement décalé affecte le taux de croissance d'une région. Ce modèle peut alors être interprété comme la spécification *minimale* permettant de capter un effet spatial exogène décalé dans un modèle de β -convergence. Ainsi, il fournit à la fois un effet « direct » et un effet spatialement décalé des niveaux de PIB par tête initiaux sur le taux de croissance d'une région.

412.3 Le modèle avec autocorrélation des erreurs

Cette spécification est utilisée lorsqu'on pense que l'autocorrélation spatiale est induite par des variables omises. Dans ce cas, le processus des erreurs est tel que la covariance entre les différents termes d'erreurs, associés aux différentes régions, n'est pas nulle. Lorsque les erreurs suivent un processus spatial autorégressif, le modèle est le suivant :

$$g_T = \alpha S + \beta y_0 + \varepsilon \quad [4.6a]$$

$$\varepsilon = \lambda W\varepsilon + u \quad \text{et} \quad u \sim N(0, \sigma^2 I) \quad [4.6b]$$

où λ est le paramètre scalaire qui exprime l'intensité de la corrélation spatiale entre les résidus de la régression. Rappelons que l'utilisation des MCO en présence d'erreurs non-sphériques produit des estimateurs centrés mais inefficients. En outre, l'inférence basée sur les MCO n'est pas fiable puisque l'estimateur de la matrice des variances-covariances est biaisé.

Ce modèle possède deux propriétés intéressantes.

La première se réfère à l'effet de diffusion de chocs aléatoires. En effet, puisque $\varepsilon = \lambda W\varepsilon + u$, alors $\varepsilon = (I - \lambda W)^{-1}u$ et le modèle peut se réécrire de la façon suivante :

$$g_T = \alpha S + \beta y_0 + (I - \lambda W)^{-1}u \quad [4.7]$$

Deuxièmement, ce modèle est réécrit d'une autre façon, qui peut être interprétée comme un *modèle minimal de β -convergence conditionnelle* intégrant deux variables d'environnement spatial. En effet, en prémultipliant l'équation précédente par $(I - \lambda W)$, il vient :

$$(I - \lambda W)g_T = \alpha(I - \lambda W)S + \beta(I - \lambda W)y_0 + u \quad [4.8]$$

On obtient alors :

$$g_T = \alpha(I - \lambda W)S + \beta y_0 + \lambda Wg_T + \gamma Wy_0 + u \quad [4.9]$$

avec la restriction $\gamma = -\lambda\beta$. Ce modèle est appelé modèle de Durbin spatial contraint et s'estime par la méthode du maximum de vraisemblance. La restriction $\gamma + \lambda\beta = 0$ est testée à l'aide du test du facteur commun. Si cette restriction ne peut pas être rejetée, alors le modèle [4.9] se réduit au modèle [4.6].

Le modèle [4.6] inclut les modèles [4.2] et [4.5] dans le sens où il incorpore à la fois la variable endogène décalée Wg_T ainsi que la variable exogène décalée Wy_0 . Il révèle deux types d'effets de débordement géographiques. En effet, le taux de croissance annuel moyen d'une région i peut être influencé par les taux de croissance annuels moyens de ses régions voisines à travers la variable endogène décalée. Il peut aussi être influencé par le niveau du PIB par tête initial de ses régions voisines, à travers la variable exogène décalée.

Les effets de débordement géographiques impliqués par cette spécification sont globaux, au contraire de la spécification moyenne mobile (Haining, 1978 ; Cliff et Ord, 1981) ou la spécification suggérée par Kelejian et Robinson (1993, 1995) qui impliquent des effets de débordement locaux. Cependant, de nombreux problèmes d'estimation associés à ces derniers modèles ne nous permettent pas de discriminer entre le caractère global et local des effets de débordement. Nous nous limitons, par conséquent, aux modèles autorégressifs.

A l'aide de l'analyse exploratoire des données spatiales, nous avons montré dans le chapitre 3 que les niveaux de PIB par tête et les taux de croissance annuels moyens étaient caractérisés par une forte autocorrélation spatiale positive pour notre échantillon de 138 régions sur la période 1980-1995. Ainsi, il est probable que le modèle de β -convergence absolue, qui relie le taux de croissance annuel moyen au PIB par tête initial en logarithmes, sera également affecté par une autocorrélation spatiale, sous la forme d'une autocorrélation des erreurs ou d'une variable décalée, endogène et/ou exogène. Nous présentons maintenant les différentes analyses économétriques que nous avons réalisées pour détecter et traiter cette autocorrélation spatiale.

413 Convergence et effets de débordement géographiques

Pour déterminer la spécification de l'autocorrélation spatiale la plus appropriée à notre étude, nous estimons et testons le modèle de β -convergence absolue par les MCO en utilisant plusieurs matrices de poids (paragraphe 413.1). Dans un deuxième temps, nous recherchons parmi les spécifications précédentes celle qui est la plus appropriée à la forme d'autocorrélation spatiale présente dans notre échantillon (paragraphe 413.2).

L'ensemble des résultats de régression trouvés dans ce paragraphe est fourni dans les tableaux 4.2a, 4.2b, 4.3a et 4.3b (cf. p. 307-310).

413.1 Une forte autocorrélation spatiale dans le modèle de β -convergence absolue

Nous reprenons les matrices de poids utilisées dans le chapitre précédent, soit la matrice des plus proches voisins et la matrice basée sur les distances inverses. Nous rappelons par commodité, les formes générales de ces matrices.

* Pour la matrice des plus proches voisins :

$$\begin{cases} w_{ij}^*(k) = 0 & \text{si } i = j \quad \forall k \\ w_{ij}^*(k) = 1 & \text{si } d_{ij} \leq d_i(k) \\ w_{ij}^*(k) = 0 & \text{si } d_{ij} > d_i(k) \end{cases} \quad k = 10, 15 \text{ et } 20 \quad [4.10]$$

où d_{ij} est la distance sphérique entre les centroïdes des régions i et j , $d_i(k)$ est la valeur-seuil définie pour chaque région, de telle sorte que chaque région a exactement k voisins.

* Pour la matrice basée sur les distances inverses :

$$\begin{cases} w_{ij}^*(k) = 0 & \text{si } i = j \quad \forall k \\ w_{ij}^*(k) = 1/d_{ij}^2 & \text{si } d_{ij} \leq D(k) \\ w_{ij}^*(k) = 0 & \text{si } d_{ij} > D(k) \end{cases} \quad k = 1, \dots, 3 \quad [4.11]$$

où $D(1) = Q_1$, $D(2) = Me$ et $D(3) = Q_3$, Q_1 , Me et Q_3 sont respectivement le premier quartile (516 km), la médiane (953 km) et le troisième quartile (1501 km) de la distribution des distances.

Contrairement au chapitre précédent, nous présentons dans le texte les résultats obtenus à l'aide des matrices $W(1)$ et $W(10)$. En effet, certains tests effectués à l'aide de cette dernière posent problème et nous préférons ainsi fournir l'ensemble des conclusions pour les deux types de matrices.

Prenons comme point de départ le modèle de β -convergence absolue adapté à notre échantillon :

$$g_T = S\alpha + \beta y_0 + \varepsilon \quad [4.12a]$$

$$\varepsilon \sim N(0, \sigma_\varepsilon^2 I) \quad [4.12b]$$

où g_T est le vecteur de dimension $(N,1)$ des taux de croissance annuels moyens des PIB par tête pour chaque région i entre 1995 et 1980, $N = 138$, $T = 15$, y_0 est le vecteur contenant les

observations des PIB par tête pour toutes les régions en 1980, α et β sont les paramètres inconnus à estimer, S est le vecteur somme de dimension $(N,1)$ et ε est le vecteur des erreurs avec les propriétés habituelles. Les valeurs des PIB par tête sont exprimées en logarithmes.

Les résultats des estimations par les MCO de ce modèle sont donnés dans la seconde colonne du tableau 4.2a pour les données en Ecus et dans la seconde colonne du tableau 4.2b pour les données en SPA. Les écarts-types estimés des coefficients ont été calculés à l'aide de la procédure de White (1980), afin de tenir compte d'une hétéroscédasticité de forme inconnue (cf. paragraphe 232.1 du chapitre 2, p. 168).

On observe que les coefficients estimés $\hat{\alpha}$ et $\hat{\beta}$ sont quasiment identiques dans les deux séries de données. Ainsi, le coefficient associé avec le PIB par tête initial est significatif et négatif, $\hat{\beta} = -0,0079$ dans les deux cas. Ceci confirme l'hypothèse de convergence pour notre échantillon de régions européennes. La vitesse de convergence associée à cette estimation est égale à 0,84% et la demi-vie est de 88 ans. Ces résultats indiquent que le processus de convergence est effectivement très faible.

La vitesse de convergence est bien inférieure aux 2% trouvés habituellement dans la littérature traitant de la convergence internationale (Barro et Sala-I-Martin, 1995). Le tableau 4.1 reprend les résultats des principales études effectuées sur la β -convergence, absolue et conditionnelle, entre les régions européennes. On remarque que sur la période qui nous intéresse, 1980-1995, la vitesse de convergence estimée est souvent relativement faible et rarement supérieure à 2%. Nos résultats sont ainsi en conformité avec ceux trouvés au niveau européen.

Pour les deux séries de données, on note que le test de Jarque-Bera (1987) rejette la normalité des résidus à 5%, le rejet étant plus important pour les PIB par tête en SPA. Concernant l'hétéroscédasticité, les résultats diffèrent selon la variable étudiée. Pour les PIB par tête en Ecus, le test de White (1980) et le test de Breusch-Pagan (1979) conduisent tous deux au rejet de l'hypothèse d'homoscédasticité. Pour les PIB par tête en SPA, au contraire, les deux tests ne permettent pas de rejeter l'homoscédasticité. Dans cette section, nous ne traitons pas davantage des problèmes d'hétérogénéité spatiale potentiellement présents dans notre échantillon. Ils seront détaillés dans la section suivante. Nous traitons uniquement l'autocorrélation spatiale en considérant, en première approximation, que l'hétéroscédasticité trouvée pour les PIB par tête en Ecus est impliquée par la présence d'autocorrélation spatiale (Anselin, 1988a ; Anselin et Griffith, 1988).

Pour détecter l'autocorrélation spatiale et déterminer sa forme, cinq tests ont été effectués : le test I de Moran adapté aux résidus de régression (Cliff et Ord, 1981). Ce test est très puissant contre toutes les formes de dépendance spatiale mais ne permet pas de les discriminer. Dans ce but, nous utilisons les deux tests du multiplicateur de Lagrange ainsi que leurs versions robustes : LM_{LAG} , LM_{ERR} , RLM_{LAG} et RLM_{ERR} (cf. paragraphes 223.3 et 223.4 du

chapitre 2). LM_{ERR} permet de tester une autocorrélation spatiale des erreurs et LM_{LAG} permet de tester une variable endogène décalée. Leurs versions robustes RLM_{ERR} et RLM_{LAG} ont une bonne puissance contre leur alternative spécifique.

Les résultats sont présentés dans les tableaux 4.2a et 4.2b pour $W(1)$ (seconde colonne) et dans les tableaux 4.3a et 4.3b pour $W(10)$ (seconde colonne). Dans tous les cas, le test de Moran indique la présence d'une certaine forme d'autocorrélation spatiale. Les règles de décision, suggérées par Anselin et Florax (1995b) et présentées dans le chapitre 2 (paragraphe 224.2, p. 147), indiquent toujours la présence d'une autocorrélation spatiale des erreurs plutôt que la présence d'une variable endogène décalée. En effet, dans tous les cas, LM_{ERR} (respectivement RLM_{ERR}) est plus significatif que LM_{LAG} (respectivement RLM_{LAG}). Le modèle avec autocorrélation spatiale des erreurs apparaît donc être la spécification la plus appropriée, à la fois pour les PIB par tête en Ecus et en SPA.

En conclusion, le modèle de β -convergence absolue souffre d'une mauvaise spécification due à l'omission de l'autocorrélation spatiale des erreurs. En réalité, chaque région n'est pas indépendante des autres, comme cela est fréquemment supposé dans les études antérieures effectuées au niveau régional. L'inférence statistique que nous avons réalisée sur ce modèle n'est pas fiable. Le modèle de β -convergence absolue doit donc être modifié pour intégrer explicitement cette forme particulière de dépendance spatiale.

413.2 Le modèle avec autocorrélation spatiale des erreurs

Considérons tout d'abord les résultats des estimations par la méthode du maximum de vraisemblance pour le modèle avec autocorrélation spatiale des erreurs [4.6], obtenus avec $W(1)$. Ils sont fournis dans la troisième colonne du tableau 4.2a en utilisant les PIB par tête en Ecus et dans le tableau 4.2b pour les PIB par tête en SPA (troisième colonne).

On note que les coefficients sont tous significatifs. Le coefficient associé au niveau de PIB par tête initial $\hat{\beta}$ est supérieur à celui du modèle de β -convergence absolue estimé par les MCO : la vitesse de convergence augmente ainsi pour atteindre 1,23% pour les PIB par tête en Ecus et 0,98% pour les PIB par tête en SPA. La demi-vie est réduite à 62 et 75 ans respectivement. Ce résultat suggère que les effets de proximité favorisent légèrement le processus de convergence régional mais que celui-ci reste encore relativement faible. En outre, l'estimation de ce modèle par la méthode des moments généralisés (quatrième colonne) conduit pratiquement aux mêmes résultats sur le paramètre de convergence dans les deux cas.

Cette première implication sur la vitesse de convergence peut sembler quantitativement négligeable, mais nous devons souligner qu'il s'agit là de la seule manière appropriée d'estimer un modèle de β -convergence une fois que l'autocorrélation spatiale des erreurs est détectée. C'est aussi la seule manière de réaliser des tests d'inférence statistique fiables.

Auteurs	Base de données	Pays	Niveau NUTS	Méthode(s) d'estimation	Type de convergence	Période	Vitesse de convergence
Barro et Sala-I-Martin (1991)	Eurostat-Regio + autres	Allemagne, Royaume-Uni, Italie, France, Pays-Bas, Belgique, Danemark	NUTS 1 – NUTS 2	MCO	Absolue	1950 – 1960	1,06%
					Conditionnelle		1,05%-2,06%*
					Absolue	1960 – 1970	3,67%
					Conditionnelle		2,41%-2,79%*
					Absolue	1970 – 1980	0,35%
					Conditionnelle		1,39%-1,84%*
Absolue	1980 – 1985	9,53%					
Conditionnelle		1,11%-1,16%*					
Armstrong (1995a)	Eurostat-Regio + autres	UE 12	NUTS 1-NUTS2	MCO	Absolue	1950 – 1970	2,11%
					Absolue	1970 – 1990	1,17%
					Absolue	1975 – 1992	1,03%
Armstrong (1995b)	Cambridge Econometrics + autres	UE 12	NUTS 1-NUTS2	MCO	Absolue	1950 – 1960	0,94%-1,01%*
					Conditionnelle		1%-2,1%**
					Absolue	1960 – 1970	3%-3,71%*
					Conditionnelle		2,34%-2,87%**
					Absolue	1970 – 1980	1,31%-2,01%*
					Conditionnelle		1,09%-1,86%**
Absolue	1980 – 1990	1,19%-2,46%*					
Conditionnelle		0,73%-1,03%**					
Barro et Sala-I-Martin (1995)	Eurostat-Regio + autres	Allemagne, Royaume-Uni, Italie, France, Pays-Bas, Belgique, Danemark, Espagne	NUTS 1 – NUTS 2	MCO	Conditionnelle	1950 – 1960	1,8%-3,4%*
					Conditionnelle	1960 – 1970	2%-2,3%*
					Conditionnelle	1970 – 1980	2%-2,2%*
					Conditionnelle	1980 – 1990	0,7%-1%*

Auteurs	Base de données	Pays	Niveau NUTS	Méthode(s) d'estimation	Type de convergence	Période	Vitesse de convergence
Dewhurst et Mutis-Gaitan (1995)	Eurostat-Regio	UE 12	NUTS 1	MCO	Conditionnelle	1981 – 1991	0,85% - 1,21%*
Neven et Gouyette (1994, 1995)	Eurostat-Regio	UE 12	NUTS 2	MCO	Absolue	1980 – 1989	0,53%
					Conditionnelle		1,11%
					Absolue	1980 – 1985	25,1%
					Conditionnelle		2,01%
Absolue	1985 – 1989	0,77%					
Conditionnelle		0,42%					
Fagerberg et Verspagen (1996)	Eurostat-Regio + autres	Allemagne, Royaume-Uni, Italie, France, Pays-Bas, Belgique	NUTS 1 – NUTS 2	MCO-White	Absolue	1950 – 1970	4,34%
					Conditionnelle		2,39%
					Absolue	1970 – 1990	2,39%
					Conditionnelle		0,64%
Absolue	1980 – 1990	0,62%					
Conditionnelle		-					
Sala-I-Martin (1996a)	Eurostat-Regio + autres	Allemagne, Royaume-Uni, Italie, France, Pays-Bas, Belgique, Danemark, Espagne	NUTS 1	MCO	Absolue	1950 – 1990	1,5%
Button et Pentecost (1999)	Eurostat-Regio	UE 12	NUTS 1 – NUTS 2	MCO-White	Absolue	1975 – 1988	3,06%
					Absolue	1975 – 1981	2,76%
					Absolue	1981 – 1988	3,75%
Fingleton (1999)	Cambridge Econometrics	UE 12	NUTS 2	MCO	Absolue Conditionnelle Absolue conditionnelle	1975 – 1995	0,25%
				MCO			0,62%-1,27%*
				ML – AS***			0,79%
				ML – AS***			0,87%-1,15%*

Auteurs	Base de données	Pays	Niveau NUTS	Méthode(s) d'estimation	Type de convergence	Période	Vitesse de convergence
Jean-Pierre (1999)	Eurostat-Regio	UE 12	NUTS 2	MCO-White	Absolue Conditionnelle	1980 – 1995	1,27% 1,05%
Beine et Jean-Pierre (2000)	Eurostat-Regio	UE 12 sauf Irlande	NUTS 1	MCO-White	Absolue Conditionnelle	1980 – 1995	0,63% 0,20%
Martin (2001)	Cambridge Econometrics	UE 16	NUTS 1 – NUTS 2	MCO	Absolue	1975 – 1986	0,72%
					Conditionnelle		0,98%
					Absolue	1986 – 1998	0,18%
Conditionnelle		0,20%					
Absolue	1975 – 1998	0,41%					
Conditionnelle		0,54%					
Maurseth (2001)	Eurostat-Regio	UE 12 sauf Irlande	NUTS 1 – NUTS 2	MCO-White	Conditionnelle	1980 – 1994	1,1%-0,7%****
						1980 – 1988	1,3%-0,6%****
						1988 – 1994	1,4%-1,4%****
Baumont <i>et al.</i> (2002a)	Eurostat-Regio	UE 12 sauf Royaume-Uni, Irlande et Grèce	NUTS 2	MCO ML – AS***	Absolue	1980 – 1995	1,70% 1,96%

Tableau 4.1 : Convergence dans les régions européennes

Notes : (*) La vitesse de convergence est variable selon les variables de conditionnement introduites dans l'équation de convergence.

(**) Des échantillons de tailles différentes ont été utilisés.

(***) Estimation du maximum de vraisemblance d'un modèle avec autocorrélation spatiale des erreurs.

(****) Régressions effectuées avec ou sans Groningen.

L'estimation fait ressortir une forte autocorrélation spatiale positive des erreurs : $\hat{\lambda} = 0,783$ en Ecus et $\hat{\lambda} = 0,531$ en SPA. Le test du facteur commun indique que la restriction $\gamma + \lambda\beta = 0$ ne peut pas être rejetée de telle sorte que le modèle avec autocorrélation spatiale des erreurs peut être réécrit comme un modèle spatial de Durbin contraint (modèle [4.9]) :

$$g_T = \alpha(I - \lambda W)S + \beta y_0 + \lambda Wg_T + \gamma W y_0 + u$$

où $\gamma = -\lambda\beta$. La significativité de ce coefficient a été établie à l'aide de la méthode Delta asymptotique. Il apparaît qu'il n'est jamais significatif au seuil de 5%.

L'estimation du modèle avec autocorrélation spatiale des erreurs permet donc de quantifier un double effet de débordement. Le taux de croissance dans une région est significativement et positivement influencé par le taux de croissance des régions voisines ($\hat{\lambda} = 0,783$ en Ecus et $\hat{\lambda} = 0,531$ en SPA). En revanche, l'influence des PIB par tête initiaux des régions voisines n'est pas significative.

Ces résultats indiquent que les schémas d'association spatiale ne sont pas neutres quant aux performances économiques des régions européennes. Ainsi, plus une région est entourée de régions dynamiques avec des taux de croissance élevés, plus son taux de croissance sera élevé. Comme nous l'avons montré dans le chapitre 2, les régions concernées se situent principalement au Portugal et au sud de l'Allemagne. En revanche, les régions entourées de régions avec faibles taux de croissance sont elles-mêmes négativement affectées. Les régions françaises, anglaises ainsi que la région grecque Sterea Ellada se trouvent dans ce cas. En d'autres termes, l'environnement géographique compte, ce qui rejoint certaines prédictions des modèles de l'économie géographique.

Par ailleurs, la réécriture du modèle intégrant l'autocorrélation spatiale des erreurs aboutit à un modèle de convergence conditionnelle intégrant les variables d'environnement spatial. Ces dernières permettent de contrôler les effets de proximité spatiale de chaque région européenne dans le processus de la convergence régionale.

D'autres tests ont encore été effectués. Le test LM_{LAG}^* ne permet pas de rejeter l'hypothèse nulle d'absence d'une variable endogène décalée supplémentaire dans le modèle avec autocorrélation spatiale. Le test de Breusch-Pagan spatial demeure non-significatif lorsqu'on utilise les PIB par tête en SPA, ce qui indique une absence d'hétéroscédasticité contre y_0 . En outre, ce test n'est plus significatif au niveau de 5% lorsqu'on utilise les PIB par tête en Ecus. Par conséquent, dans ce dernier cas, l'hétéroscédasticité trouvée dans le modèle de β -convergence absolue estimé par les MCO peut être interprétée comme étant entièrement due à l'autocorrélation spatiale.

Compte tenu des critères d'information, ce modèle semble être meilleur que le précédent (Akaike, 1974 ; Schwarz, 1978) à la fois en Ecus et en SPA. Dans le seul but de

« vérifier » la règle de décision appliquée, nous avons estimé deux modèles supplémentaires, d'une part le modèle autorégressif incluant la variable endogène décalée [4.2], estimé par la méthode du maximum de vraisemblance, d'autre part le modèle régressif croisé incluant uniquement la variable exogène décalée du PIB initial [4.5], estimé par la méthode des MCO avec la correction de White. Les résultats des estimations sont fournis dans les cinquième et sixième colonnes des tableaux 4.2a et 4.2b.

Concernant le modèle avec variable endogène décalée, on note que le processus de convergence apparaît être encore plus faible. Ainsi, pour $W(1)$, la demi-vie atteint 157 ans pour les PIB par tête en Ecus et 107 ans pour les PIB par tête en SPA. Un effet de débordement géographique est toujours trouvé dans le sens où le taux de croissance du PIB par tête dans une région donnée est influencé par celui des régions voisines ($\hat{\rho} = 0,769$ en Ecus et $\hat{\rho} = 0,543$ en SPA). Ce modèle n'est pas satisfaisant. D'une part, pour les PIB par tête en Ecus, le test de Breusch-Pagan ajusté spatialement est toujours significatif et indique une hétéroscédasticité résiduelle. D'autre part, ce modèle n'est pas meilleur que le modèle avec autocorrélation spatiale des erreurs en termes des critères d'information.

Concernant le modèle régressif croisé, on observe qu'il n'y a pas d'effet de débordement géographique significatif associé à la variable initiale ($p = 0,437$ pour les PIB par tête en Ecus et $p = 0,865$ pour les PIB par tête en SPA ⁶⁷). En outre, pour les PIB par tête en Ecus, quelques problèmes subsistent sur la normalité et l'hétéroscédasticité des erreurs. Pour les PIB par tête en SPA, la normalité est toujours rejetée. Les tests d'autocorrélation spatiale indiquent la présence d'une autocorrélation spatiale des erreurs pour les PIB par tête en Ecus et en SPA. Ce dernier modèle semble ainsi mal spécifié et est le moins bon en termes des critères d'information.

Les estimations et les tests pour $W(10)$ sont présentés dans les tableaux 4.3a et 4.3b. Les résultats sont qualitativement similaires à ceux obtenus avec la matrice de distance. Ainsi, les tests effectués sur le modèle de β -convergence absolue estimé par les MCO indiquent toujours que la spécification la plus appropriée est le modèle avec autocorrélation spatiale des erreurs. Dans ce dernier, il apparaît que la vitesse de convergence est encore plus importante. Elle se monte à 1,62% pour les PIB par tête en Ecus (la demi-vie étant de 48 ans) et à 1,02% pour les PIB par tête en SPA (la demi-vie associée étant de 73 ans). Notons cependant que l'estimation du modèle avec autocorrélation spatiale des erreurs avec $W(10)$ pose problème pour le test du facteur commun. En effet, il apparaît dans ce dernier cas que le test du ratio de vraisemblance est négatif, indiquant que la valeur de la fonction de vraisemblance à l'optimum du modèle non-contraint est inférieure à la valeur de la fonction de vraisemblance à l'optimum du modèle contraint. Ce résultat, théoriquement impossible, pourrait provenir de problèmes de données et de multicollinéarité dans le modèle non-contraint. Nous avons donc reporté dans les tableaux les tests de Wald pour le facteur commun. Les résultats doivent cependant être interprétés avec précaution. La méthode d'estimation par les GMM est sans doute plus appropriée dans ce cas.

⁶⁷ p dénote la probabilité critique.

Modèle	β -convergence	Autocorrélation des erreurs		Variable endogène décalée	Régressif croisé
	MCO-White	MV	GMM	MV	MCO-White
$\hat{\alpha}$	0,129 (0,000)	0,158 (0,000)	0,159 (0,000)	0,053 (0,001)	0,123 (0,000)
$\hat{\beta}$	-0,0079 (0,002)	-0,0112 (0,000)	-0,0113 (0,000)	-0,0044 (0,005)	-0,0109 (0,006)
Vitesse de convergence	0,84% (0,000)	1,23% (0,000)	1,24% (0,000)	0,46% (0,002)	1,19% (0,021)
Demi-vie	88	62	61	157	64
$\hat{\lambda}$	-	0,783 (0,000)	0,822	-	-
$\hat{\rho}$	-	-	-	0,769 (0,000)	-
$\hat{\gamma}$	-	-	-	-	-0,0037 (0,437)
R^2 ou Sq.- Corr.*	0,13	0,13*	0,13*	0,54*	0,14
LIK	456,14	494,32	-	491,78	456,44
AIC	-908,27	-984,65	-	-977,57	-906,90
BIC	-902,42	-978,80	-	-968,79	-898,11
$\hat{\sigma}^2$	$7,996 \cdot 10^{-5}$	$4,088 \cdot 10^{-5}$	-	$4,263 \cdot 10^{-5}$	$8,019 \cdot 10^{-5}$
Tests					
JB	8,976 (0,011)	-	-	-	9,821 (0,007)
BP ou BP-S* ou KB** vs. y_0	14,786 (0,000)	3,324* (0,068)	-	6,972* (0,008)	8,849** (0,003)
White	29,903 (0,000)	-	-	-	37,41 (0,000)
I de Moran	13,425 (0,000)	-	-	-	13,310 (0,000)
LM_{ERR}	151,683 (0,000)	-	-	-	149,247 (0,000)
RLM_{ERR}	18,102 (0,000)	-	-	-	1,448 (0,228)
LM_{ERR}^*	-	-	-	2,198 (0,138)	-
LM_{LAG}	134,199 (0,000)	-	-	-	150,828 (0,000)
RLM_{LAG}	0,618 (0,432)	-	-	-	3,029 (0,0818)
LM_{LAG}^*	-	0,705 (0,401)	-	-	-
LR-facteur commun	-	0,151 (0,697)	-	-	-
$\hat{\gamma} = -\hat{\lambda}\hat{\beta}$	-	0,010 (0,869)	-	-	-

Tableau 4.2a : Résultats d'estimation avec $W(1)$ et PIB par tête en Ecus

Notes : les valeurs critiques sont entre parenthèses. MCO-White indique l'utilisation de la matrice des variances-covariances convergente en présence d'hétéroscédasticité dans l'estimation par les MCO (White, 1980). Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). JB est le test de normalité des résidus de Jarque-Bera (1987). BP est le test de Breusch-Pagan (1979) pour l'hétéroscédasticité, BP-S est la version ajustée spatialement (Anselin, 1988a) et KB est le test de Koenker-Bassett (1982) de l'hétéroscédasticité en présence de non-normalité. White est le test de White (1980) pour l'hétéroscédasticité. I de Moran est le test de Moran adapté aux résidus (Cliff et Ord, 1981). LM_{ERR} est le test du multiplicateur de Lagrange pour une autocorrélation des erreurs et RLM_{ERR} est sa version robuste. LM_{LAG} est le test du multiplicateur de Lagrange pour une variable endogène décalée et RLM_{LAG} est sa version robuste (Anselin *et al.*, 1996). LM_{ERR}^* est le test du multiplicateur de Lagrange pour une autocorrélation spatiale des erreurs dans le modèle avec variable endogène décalée et LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). LR-facteur commun est le test du ratio de vraisemblance du facteur commun (Burridge, 1981).

Modèle	β -convergence	Autocorrélation des erreurs		Variable endogène décalée	Régressif croisé
	MCO-White	MV	GMM	MV	MCO-White
$\hat{\alpha}$	0,130 (0,000)	0,141 (0,000)	0,141 (0,000)	0,084 (0,000)	0,128 (0,000)
$\hat{\beta}$	-0,0079 (0,000)	-0,0091 (0,000)	-0,0092 (0,000)	-0,0064 (0,001)	-0,0083 (0,023)
Vitesse de convergence	0,84% (0,000)	0,98% (0,000)	0,99%	0,68% (0,000)	0,89% (0,006)
Demi-vie	88	75	75	107	83
$\hat{\lambda}$	-	0,531 (0,000)	0,586	-	-
$\hat{\rho}$	-	-	-	0,543 (0,000)	-
$\hat{\gamma}$	-	-	-	-	-0,00068 (0,865)
R^2 ou Sq.- Corr.*	0,11	0,11*	0,11*	0,25*	0,11
LIK	499,27	507,77	-	507,80	499,29
AIC	-994,54	-1011,53	-	-1009,60	-992,57
BIC	-988,69	-1005,68	-	-1000,82	-983,79
$\hat{\sigma}^2$	$4,279.10^{-5}$	$3,588.10^{-5}$	-	$3,578.10^{-5}$	$4,310.10^{-5}$
Tests					
JB	12,469 (0,002)	-	-	-	12,351 (0,002)
BP ou BP-S* ou KB** vs. y_0	0,297** (0,586)	0,561* (0,454)	-	0,771* (0,380)	0,268** (0,604)
White	0,890 (0,641)	-	-	-	10,194 (0,067)
I de Moran	5,623 (0,000)	-	-	-	5,584 (0,000)
LM_{ERR}	25,121 (0,000)	-	-	-	24,626 (0,000)
RLM_{ERR}	3,309 (0,069)	-	-	-	2,338 (0,126)
LM_{ERR}^*	-	-	-	1,774 (0,183)	-
LM_{LAG}	21,842 (0,000)	-	-	-	24,780 (0,000)
RLM_{LAG}	0,030 (0,863)	-	-	-	2,493 (0,114)
LM_{LAG}^*	-	2,217 (0,136)	-	-	-
LR-facteur commun	-	0,536 (0,464)	-	-	-
$\hat{\gamma} = -\hat{\lambda}\hat{\beta}$	-	0,005 (0,937)	-	-	-

Tableau 4.2b : Résultats d'estimation avec $W(1)$ et PIB par tête en SPA

Notes : les valeurs critiques sont entre parenthèses. MCO-White indique l'utilisation de la matrice des variances-covariances convergente en présence d'hétéroscédasticité dans l'estimation par les MCO (White, 1980). Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaïke (1974). BIC est le critère d'information de Schwarz (1978). JB est le test de normalité des résidus de Jarque-Bera (1987). BP est le test de Breusch-Pagan (1979) pour l'hétéroscédasticité, BP-S est la version ajustée spatialement (Anselin, 1988a) et KB est le test de Koenker-Bassett (1982) de l'hétéroscédasticité en présence de non-normalité. White est le test de White (1980) pour l'hétéroscédasticité. I de Moran est le test de Moran adapté aux résidus (Cliff et Ord, 1981). LM_{ERR} est le test du multiplicateur de Lagrange pour une autocorrélation des erreurs et RLM_{ERR} est sa version robuste. LM_{LAG} est le test du multiplicateur de Lagrange pour une variable endogène décalée et RLM_{LAG} est sa version robuste (Anselin *et al.*, 1996). LM_{ERR}^* est le test du multiplicateur de Lagrange pour une autocorrélation spatiale des erreurs dans le modèle avec variable endogène décalée et LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). LR-facteur commun est le test du ratio de vraisemblance du facteur commun (Burridge, 1981).

Modèle	β -convergence	Autocorrélation des erreurs		Variable endogène décalée	Régressif croisé
	MCO-White	MV	GMM	MV	MCO-White
$\hat{\alpha}$	0,129 (0,000)	0,186 (0,000)	0,187 (0,000)	0,060 (0,001)	0,107 (0,000)
$\hat{\beta}$	-0,0079 (0,002)	-0,0144 (0,000)	-0,0145 (0,000)	-0,0053 (0,001)	-0,017 (0,008)
Vitesse de convergence	0,84% (0,000)	1,62% (0,000)	1,63% (0,000)	0,55% (0,000)	1,90% (0,000)
Demi-vie	88	48	47	131	41
$\hat{\lambda}$	-	0,802 (0,000)	0,813	-	-
$\hat{\rho}$	-	-	-	0,772 (0,000)	-
$\hat{\gamma}$	-	-	-	-	0,011 (0,022)
R^2 ou Sq.- Corr.*	0,13	0,13*	0,13*	0,49*	0,17
LIK	456,14	497,42	-	485,60	458,81
AIC	-908,27	-990,85	-	-965,21	-911,61
BIC	-902,42	-984,99	-	-956,43	-902,83
$\hat{\sigma}^2$	$7,996.10^{-5}$	$4,332.10^{-5}$	-	$4,748.10^{-5}$	$7,750.10^{-5}$
Tests					
JB	8,976 (0,011)	-	-	-	9,245 (0,010)
BP ou BP-S* ou KB** vs. y_0	14,786 (0,000)	3,315* (0,068)	-	14,124* (0,000)	5,237** (0,022)
White	29,903 (0,000)	-	-	-	34,364 (0,000)
I de Moran	13,937 (0,000)	-	-	-	13,449 (0,000)
LM_{ERR}	155,704 (0,000)	-	-	-	144,230 (0,000)
RLM_{ERR}	36,484 (0,000)	-	-	-	2,295 (0,130)
LM_{ERR}^*	-	-	-	4,343 (0,037)	-
LM_{LAG}	124,457 (0,000)	-	-	-	150,769 (0,000)
RLM_{LAG}	5,237 (0,022)	-	-	-	8,833 (0,003)
LM_{LAG}^*	-	0,027 (0,868)	-	-	-
Wald-facteur commun	-	0,189 (0,663)	-	-	-
$\hat{\gamma} = -\hat{\lambda}\hat{\beta}$	-	0,01 (0,823)	-	-	-

Tableau 4.3a : Résultats d'estimation avec $W(10)$ et PIB par tête en Ecus

Notes : les valeurs critiques sont entre parenthèses. MCO-White indique l'utilisation de la matrice des variances-covariances convergente en présence d'hétéroscédasticité dans l'estimation par les MCO (White, 1980). Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). JB est le test de normalité des résidus de Jarque-Bera (1987). BP est le test de Breusch-Pagan (1979) pour l'hétéroscédasticité, BP-S est la version ajustée spatialement (Anselin, 1988a) et KB est le test de Koenker-Bassett (1982) de l'hétéroscédasticité en présence de non-normalité. White est le test de White (1980) pour l'hétéroscédasticité. I de Moran est le test de Moran adapté aux résidus (Cliff et Ord, 1981). LM_{ERR} est le test du multiplicateur de Lagrange pour une autocorrélation des erreurs et RLM_{ERR} est sa version robuste. LM_{LAG} est le test du multiplicateur de Lagrange pour une variable endogène décalée et RLM_{LAG} est sa version robuste (Anselin *et al.*, 1996). LM_{ERR}^* est le test du multiplicateur de Lagrange pour une autocorrélation spatiale des erreurs dans le modèle avec variable endogène décalée et LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). Wald-facteur commun est le test du ratio de vraisemblance du facteur commun (Burrige, 1981).

Modèle	β -convergence	Autocorrélation des erreurs		Variable endogène décalée	Régressif croisé
	MCO-White	MV	GMM	MV	MCO-White
$\hat{\alpha}$	0,130 (0,000)	0,144 (0,000)	0,145 (0,000)	0,086 (0,000)	0,123 (0,000)
$\hat{\beta}$	-0,0079 (0,000)	-0,0095 (0,000)	-0,0096 (0,000)	-0,0064 (0,000)	-0,0093 (0,034)
Vitesse de convergence	0,84% (0,000)	1,02% (0,000)	1,04%	0,68% (0,000)	1,00% (0,009)
Demi-vie	88	73	72	107	74
$\hat{\lambda}$	-	0,527 (0,000)	0,563	-	-
$\hat{\rho}$	-	-	-	0,510 (0,000)	-
$\hat{\gamma}$	-	-	-	-	0,0022 (0,715)
R^2 ou Sq.- Corr.*	0,11	0,11*	0,11*	0,22*	0,11
LIK	499,27	508,76	-	505,97	499,41
AIC	-994,54	-1013,52	-	-1005,94	-992,82
BIC	-988,69	-1007,67	-	-997,16	-984,04
$\hat{\sigma}^2$	$4,279 \cdot 10^{-5}$	$3,683 \cdot 10^{-5}$	-	$3,727 \cdot 10^{-5}$	$4,302 \cdot 10^{-5}$
Tests					
JB	12,469 (0,002)	-	-	-	11,819 (0,003)
BP ou BP-S* ou KB** vs. y_0	0,297** (0,586)	0,569* (0,450)	-	1,016* (0,313)	0,210** (0,647)
White	0,890 (0,641)	-	-	-	13,091 (0,022)
I de Moran	5,811 (0,000)	-	-	-	5,749 (0,000)
LM_{ERR}	25,379 (0,000)	-	-	-	24,332 (0,000)
RLM_{ERR}	5,143 (0,023)	-	-	-	1,129 (0,288)
LM_{ERR}^*	-	-	-	1,655 (0,198)	-
LM_{LAG}	24,234 (0,000)	-	-	-	24,609 (0,000)
RLM_{LAG}	0,275 (0,560)	-	-	-	1,407 (0,236)
LM_{LAG}^*	-	0,305 (0,580)	-	-	-
Wald-facteur commun	-	0,014 (0,907)	-	-	-
$\hat{\gamma} = -\hat{\lambda}\hat{\beta}$	-	0,005 (0,938)	-	-	-

Tableau 4.3b : Résultats d'estimation avec $W(10)$ et PIB par tête en SPA

Notes : les valeurs critiques sont entre parenthèses. MCO-White indique l'utilisation de la matrice des variances-covariances convergente en présence d'hétéroscédasticité dans l'estimation par MCO (White, 1980). Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). JB est le test de normalité des résidus de Jarque-Bera (1987). BP est le test de Breusch-Pagan (1979) pour l'hétéroscédasticité, BP-S est la version ajustée spatialement (Anselin, 1988a) et KB est le test de Koenker-Bassett (1982) de l'hétéroscédasticité en présence de non-normalité. White est le test de White (1980) pour l'hétéroscédasticité. I de Moran est le test de Moran adapté aux résidus (Cliff et Ord, 1981). LM_{ERR} est le test du multiplicateur de Lagrange pour une autocorrélation des erreurs et RLM_{ERR} est sa version robuste. LM_{LAG} est le test du multiplicateur de Lagrange pour une variable endogène décalée et RLM_{LAG} est sa version robuste (Anselin *et al.*, 1996). LM_{ERR}^* est le test du multiplicateur de Lagrange pour une autocorrélation spatiale des erreurs dans le modèle avec variable endogène décalée et LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). Wald-facteur commun est le test du ratio de vraisemblance du facteur commun (Burrige, 1981).

Les résultats complets pour $W(2)$, $W(3)$, $W(15)$ et $W(20)$ sont présentés dans l'annexe 4.1 pour les PIB par tête en Ecus et dans l'annexe 4.2 pour les PIB par tête en SPA. Dans chacune des configurations, le modèle avec autocorrélation spatiale des erreurs est toujours la spécification la plus appropriée et la vitesse de convergence toujours plus élevée que dans le modèle de β -convergence absolue estimé par les MCO. Concernant les effets de débordement géographiques, la même configuration que précédemment apparaît systématiquement : le taux de croissance d'une région est influencé par celui des régions voisines mais pas par le PIB par tête initial des régions voisines. En outre, les tests ne permettent pas de rejeter l'hypothèse d'homoscédasticité contre la variable initiale y_0 . Le test de présence d'une variable endogène décalée supplémentaire dans le modèle avec autocorrélation des erreurs n'est jamais significatif à 5%. Enfin, le problème concernant le test du ratio de vraisemblance du facteur commun est toujours présent lorsque les régressions sont effectuées avec les matrices des plus proches voisins. Par conséquent, seul le test de Wald est reporté dans ce cas.

L'ensemble des résultats obtenus dans ce paragraphe indique que le modèle avec autocorrélation des erreurs apparaît être la spécification la plus appropriée comme l'avait indiqué la règle de décision d'Anselin et de Florax (1995b). Nous étudions maintenant les implications de ce modèle concernant la diffusion d'un choc aléatoire.

414 Analyse du processus de diffusion spatiale d'un choc aléatoire

Une question intéressante porte sur la façon dont un choc aléatoire affectant une région se diffuse à l'ensemble du système régional étudié. En d'autres termes, nous cherchons à évaluer si et comment un choc de croissance se produisant dans l'une des régions européennes peut influencer les taux de croissance des autres régions européennes. Inversement, chaque région est ainsi susceptible de recevoir de la part des autres régions européennes un effet positif de croissance. Ces impacts sont liés conjointement au dynamisme, évalué ici en termes de taux de croissance moyen, et aux positions relatives des régions émettrices et réceptrices. Cette question peut être traitée grâce au modèle économétrique intégrant l'autocorrélation spatiale des erreurs. A l'aide de simulations, nous pouvons, dans un premier temps, estimer le processus de diffusion spatiale de diffusion d'un choc aléatoire affectant une seule région en reprenant la procédure que Rey et Montouri (1999) ont employée sur les Etats américains (paragraphe 414.1). Dans un second temps, nous proposons une procédure originale permettant d'étudier et d'évaluer globalement l'impact reçu et émis par une région dans l'ensemble du système régional européen étudié (paragraphe 414.2).

414.1 Analyse du choc affectant une seule région

Formellement, considérons la forme réduite du modèle avec autocorrélation des erreurs (modèle [4.7]) :

$$g_T = \alpha S + \beta y_0 + (I - \lambda W)^{-1} u$$

$$u \sim N(0, \sigma^2 I)$$

Le choc aléatoire est fixé à deux fois l'écart-type estimé du modèle avec autocorrélation spatiale des erreurs (Rey et Montouri, 1999). Soit \hat{u}^i le vecteur contenant les erreurs estimées du modèle [4.7] avec un choc dans la région i . Ce vecteur de dimension $(N,1)$ s'écrit de la façon suivante :

$$\hat{u}^i = (\hat{u}_1 \quad \dots \quad \hat{u}_i + 2\hat{\sigma} \quad \dots \quad \hat{u}_N) \quad [4.13]$$

Notons y^{i*} , le vecteur de dimension $(N,1)$ contenant les observations sur les taux de croissance annuels moyens simulés après un choc dans la région i . Ce vecteur est calculé de la façon suivante :

$$y^{i*} = X\hat{\gamma} + (I - \hat{\lambda}W)^{-1} \hat{u}^i \quad [4.14]$$

avec $X_{(N,2)} = [S \quad y_0]$, $\hat{\gamma}_{(2,1)} = [\hat{\alpha} \quad \hat{\beta}]'$, $\hat{\alpha}$, $\hat{\beta}$ et $\hat{\lambda}$ sont les estimations obtenues à l'aide de la méthode du maximum de vraisemblance de α , β et λ dans le modèle avec autocorrélation des erreurs [4.6].

Pour évaluer l'impact d'un choc dans la région i sur les autres régions de l'échantillon, on calcule alors la différence, notée d^i , entre le vecteur des taux de croissance annuels moyens simulés y^{i*} et le vecteur des taux de croissance annuels moyens observés y . La variation relative entre y^{i*} et y est alors obtenue en divisant d^i par y , élément par élément.

Nous avons appliqué cette procédure pour un choc affectant la région Ile-de-France. Les programmes nécessaires ont été rédigés à l'aide de GAUSS 3.2.26.

Pour $W(1)$, l'impact relatif d'un choc sur cette région est représenté dans la figure 4.1a pour les PIB par tête en Ecus et dans la figure 4.1b pour les PIB par tête en SPA. L'impact relatif de ce choc est naturellement le plus élevé sur l'Ile-de-France : le taux de croissance annuel moyen simulé est, en Ecus, de 22,19% plus élevé que le taux de croissance annuel moyen observé sans le choc. En outre, on observe un schéma de diffusion de ce choc à toutes les autres régions de l'échantillon. L'ampleur de l'impact de ce choc est comprise entre 1,6% et 3,7% pour les régions entourant l'Ile-de-France et décroît graduellement pour les régions périphériques. L'impact médian du choc est de 0,230% en excluant l'Ile-de-France.

L'ampleur de l'effet de ce choc est moins élevée lorsqu'on utilise des PIB par tête en SPA : le choc est égal à 20,05% pour l'Ile-de-France et il est compris entre 0,75% et 1,94% pour les régions entourant l'Ile-de-France. L'impact médian est de 0,034% en excluant l'Ile-de-France. Cette différence est due au coefficient d'autocorrélation spatiale estimé moins élevé pour les PIB par tête en SPA.

Nous avons également appliqué cette procédure en utilisant $W(10)$. Les résultats obtenus sont similaires et représentés dans les figures 4.2a et 4.2b. L'impact médian est de 0,0625% pour les PIB par tête en Ecus et de 0,0028% pour les PIB par tête en SPA, l'Ile-de-France étant exclue.

414.2 Emission et réception au sein du système régional européen

Les résultats précédents sont généralisés à l'ensemble des 138 régions européennes. Nous nous intéressons maintenant à la capacité d'une région à émettre un choc de croissance dans le système global et à sa capacité à recevoir les chocs de croissance de la part de ce système. Nous cherchons plus précisément à détecter 2 types de régions : celles qui sont les plus influentes dans l'échantillon (c'est-à-dire les régions pour lesquelles la diffusion spatiale du choc est la plus importante) et celles qui sont les plus affectées par les chocs produits dans les autres régions. Nous essayons d'apprécier comment le dynamisme économique et la situation géographique des régions interviennent dans ces capacités à émettre et à recevoir. Après avoir présenté sous forme analytique la procédure que nous utilisons (paragraphe 414.21), nous détaillons les résultats obtenus par simulations (paragraphe 414.22).

414.21 Présentation analytique de la diffusion globale des chocs aléatoires

Notons Y^* , la matrice de dimension (N, N) , contenant les observations sur les taux de croissance annuels moyens simulés après un choc dans toutes les régions de l'échantillon. Cette matrice s'écrit de la façon suivante :

$$Y^* = [y^{1*} \dots y^{N*}] \quad [4.15]$$

Chaque vecteur y^{i*} est défini par l'équation [4.14] : $y^{i*} = X\hat{\gamma} + (I - \hat{\lambda}W)^{-1}\hat{u}^i \quad \forall i$. Par conséquent, Y^* se réécrit ainsi :

$$Y^* = [X\hat{\gamma} \dots X\hat{\gamma}] + A^{-1}[\hat{u}^1 \dots \hat{u}^N] \quad [4.16]$$

avec $A = (I - \hat{\lambda}W)$.

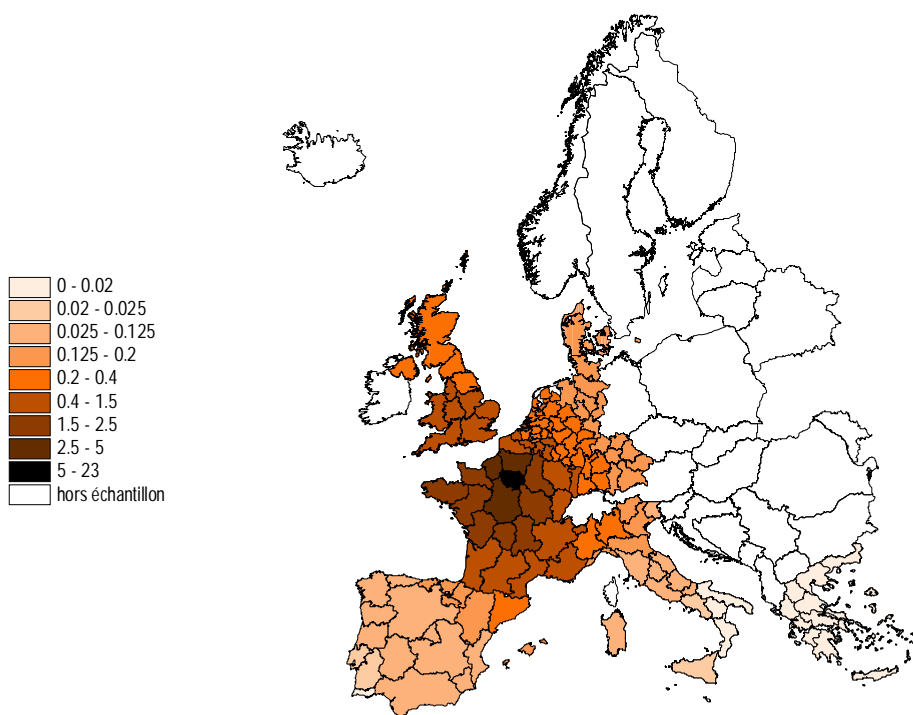


Figure 4.1a : Diffusion dans le modèle avec autocorrélation des erreurs avec $W(1)$ et PIB par tête en Ecus

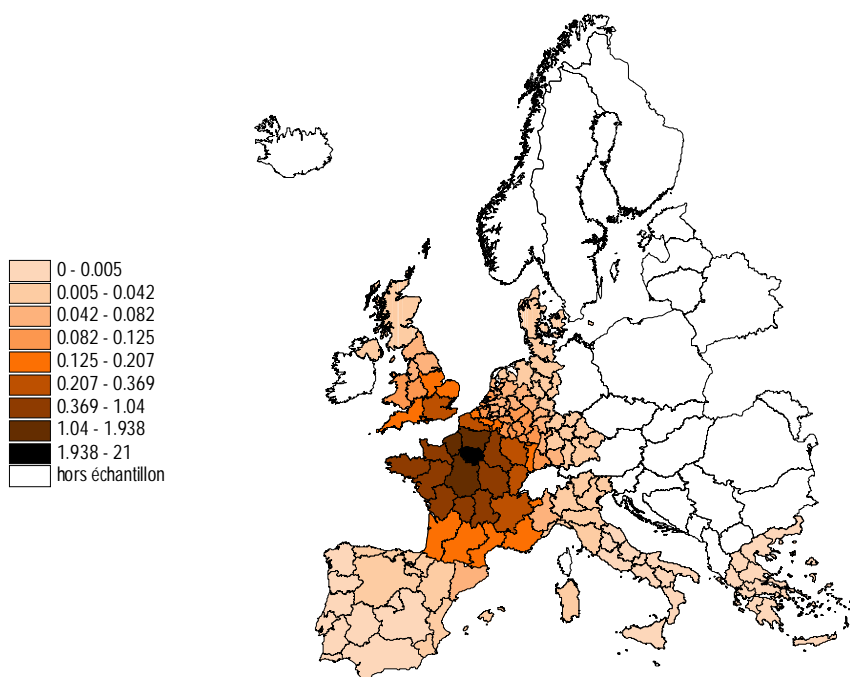


Figure 4.1b : Diffusion dans le modèle avec autocorrélation des erreurs avec $W(1)$ et PIB par tête en SPA

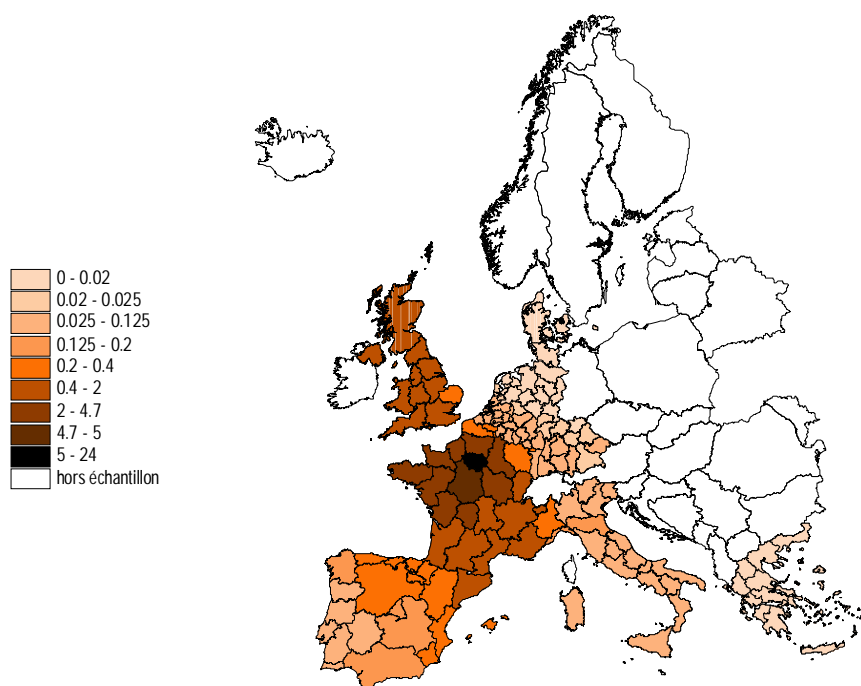


Figure 4.2a : Diffusion dans le modèle avec autocorrélation des erreurs avec $W(10)$ et PIB par tête en Ecus

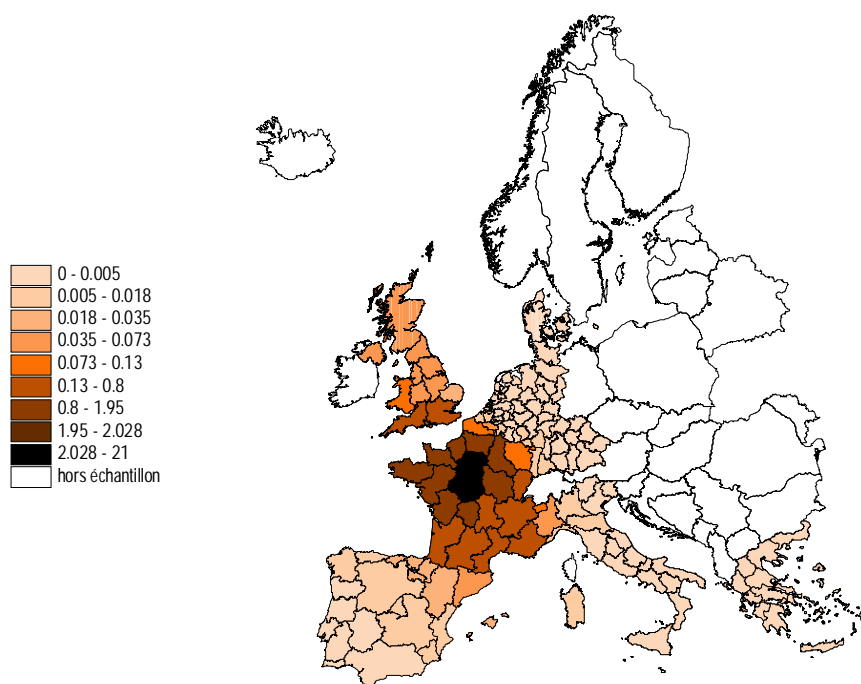


Figure 4.2b : Diffusion dans le modèle avec autocorrélation des erreurs avec $W(10)$ et PIB par tête en SPA

L'équation [4.16] est équivalente à :

$$Y^* = S' \otimes X \hat{\gamma} + A^{-1} \hat{U} \quad [4.17]$$

où \otimes est le produit Kronecker, \hat{U} est la matrice de dimension (N, N) définie par $\hat{U} = [\hat{u}^1 \dots \hat{u}^N]$. Chaque élément \hat{u}^i étant défini par l'équation [4.13], \hat{U} peut se réécrire de la façon suivante :

$$\hat{U} = \begin{pmatrix} \hat{u}_1 + 2\hat{\sigma} & \hat{u}_1 & \dots & \hat{u}_1 \\ \hat{u}_2 & \hat{u}_2 + 2\hat{\sigma} & \dots & \hat{u}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \hat{u}_n & \hat{u}_n & \dots & \hat{u}_n + 2\hat{\sigma} \end{pmatrix} \quad [4.18]$$

$$\Rightarrow \hat{U} = S' \otimes \hat{u} + 2\hat{\sigma}I \quad [4.19]$$

où $\hat{\sigma}$ est l'écart-type estimé du modèle avec autocorrélation des erreurs, estimé par la méthode du maximum de vraisemblance.

En combinant [4.17] et [4.19], on obtient finalement :

$$Y^* = S' \otimes X \hat{\gamma} + A^{-1} (S' \otimes \hat{u} + 2\hat{\sigma}I) \quad [4.20]$$

Cette expression produit une matrice de dimension (N, N) où chaque colonne i contient le taux de croissance annuel moyen simulé pour toutes les régions de l'échantillon après un choc dans la région i .

Pour évaluer l'impact de ces chocs sur les régions de l'échantillon, on calcule la différence, notée D , entre les matrices Y^* et Y . Puisque $Y = S' \otimes y$ et que $y = X \hat{\gamma} + A^{-1} \hat{u}$, on obtient :

$$D = 2\hat{\sigma}A^{-1} \quad [4.21]$$

Nous retrouvons ainsi le rôle de la matrice $A = (I - \lambda W)$ dans le processus de diffusion spatiale. Il y a diffusion d'un choc aléatoire dans toutes les régions de l'échantillon à cause de la transformation spatiale inverse $(I - \lambda W)^{-1}$ dans la forme réduite du modèle avec autocorrélation spatiale des erreurs.

Notons V la matrice de dimension (N, N) contenant la variation relative entre le taux de croissance annuel moyen simulé et le taux de croissance observé. Elle est obtenue en divisant D par Y , élément par élément. Cette matrice contient deux informations intéressantes.

D'une part, la colonne i de la matrice V indique comment la région i affecte toutes les autres régions de l'échantillon lorsqu'il y a un choc dans la région i . Chaque colonne contient donc les éléments permettant de mesurer l'ampleur de l'émission de la région i . Par conséquent, les colonnes de V permettent de détecter les régions qui sont les plus influentes dans l'échantillon. Cette influence est mesurée ici par la médiane de chaque colonne de V .

D'autre part, la ligne j de la matrice V indique comment la région j est affectée par un choc provenant de toutes les autres régions de l'échantillon. Chaque ligne mesure donc l'ampleur de la réception de la région j . Les lignes de V permettent ainsi de détecter les régions qui sont les plus influencées dans l'échantillon et cette influence est mesurée par la médiane de chaque ligne de V .

Etant donné que les chocs internes à une région sont les plus importants quantitativement, nous avons préféré ne pas les inclure dans les mesures d'influence, et nous avons donc fixé à 0 la première diagonale de la matrice V .

414.22 La géographie des chocs aléatoires dans le système régional européen

Les programmes nécessaires à l'application de cette procédure ont été rédigés à l'aide de GAUSS 3.2.26. Comme les estimations avec les matrices des plus proches voisins sont problématiques pour certains tests (cf. paragraphe 413, p. 298), nous avons uniquement effectué les calculs avec $W(1)$.

Considérons tout d'abord les résultats des simulations obtenus lorsqu'on s'intéresse à la façon dont chaque région subissant un choc aléatoire affecte toutes les autres régions de l'échantillon. Nous analysons, par conséquent, chaque colonne de la matrice V et calculons pour chaque région le premier quartile, la médiane et le troisième quartile de l'impact de l'émission à partir de cette région. Les boîtes à moustaches associées à chaque colonne de la matrice V sont présentées en Annexe 4.3 dans la figure 1 pour les données en Ecus et dans la figure 2 pour les données en SPA.

Les régions sont présentées dans un ordre décroissant relativement à leur médiane. Il apparaît que les régions les plus influentes par rapport au processus de diffusion spatiale sont les régions européennes riches appartenant surtout à la Belgique, l'Allemagne, les Pays-Bas et la France. Le Luxembourg est également une région très influente. Notons que ces régions appartiennent au cœur de l'Europe. Les régions pauvres éligibles à l'objectif 1 sont les moins influentes en raison de leur localisation périphérique.

Considérons maintenant les résultats de simulation obtenus lorsqu'on s'intéresse à la façon dont chaque région est affectée par un choc dans chacune des autres régions de l'échantillon. Nous analysons ainsi chaque ligne de la matrice V . Les figures 3 et 4 de

l'annexe 4.3 représentent les boîtes à moustaches associées à chaque ligne de la matrice V . A nouveau, il apparaît que les régions riches du Nord sont les plus influencées alors que les régions pauvres éligibles à l'objectif 1 sont les moins influencées. Ces résultats de simulation soulignent ainsi que le processus de diffusion spatiale est dominé par un schéma de type centre-périphérie et suggère une hétérogénéité potentielle dans le processus de croissance.

Finalement, nous comparons les quartiles de la distribution de l'impact médian d'émission et de l'impact médian de réception. Ceci est réalisé à l'aide des figures 3.4a et 3.4b, qui représentent les quartiles de la distribution des médianes de chaque colonne de la matrice V , et des figures 3.5a et 3.5b, qui représentent les quartiles de la distribution des médianes de chaque ligne de V . On note que la composition des quartiles les plus faibles reste globalement inchangée et concerne les mêmes régions pauvres éligibles à l'objectif 1. En revanche, la composition des quartiles les plus élevés diffère. Ainsi, les régions les plus influentes (les régions dans le dernier quartile) sont surtout situées en Belgique, aux Pays-Bas et en Allemagne du nord, alors que les régions les plus influencées sont situées en France, en Allemagne du sud et en Italie du nord. Deux éléments semblent jouer : la richesse et la localisation. Ainsi, la localisation périphérique des régions pauvres les maintient éloignées des processus de diffusion de richesse tandis que les régions plus riches et centrales bénéficient pleinement de cette diffusion.

Conclusion de section

Les résultats obtenus dans cette section tendent à confirmer l'hypothèse de β -convergence parmi les régions européennes sur la période 1980-1995. L'estimation du modèle de β -convergence absolue fait apparaître une vitesse de convergence très faible de l'ordre de 0,84%. Cependant, le modèle de β -convergence absolue souffre d'une mauvaise spécification due à l'omission d'une autocorrélation spatiale des erreurs. L'estimation de cette spécification révèle que le processus de convergence est légèrement plus rapide, avec des vitesses de convergence comprises entre 1% et 1,60%. En outre, cette estimation a fait apparaître un effet de débordement géographique sur les taux de croissance. Le processus de diffusion d'un choc aléatoire, impliqué par le modèle avec autocorrélation spatiale des erreurs, est souligné par des expériences de simulation. Ces dernières font apparaître une différence de comportement entre les régions du Nord et du Sud de l'Europe. Par conséquent, l'hétérogénéité spatiale, que nous n'avons pas traitée dans cette section, doit constituer l'objet de la prochaine section.

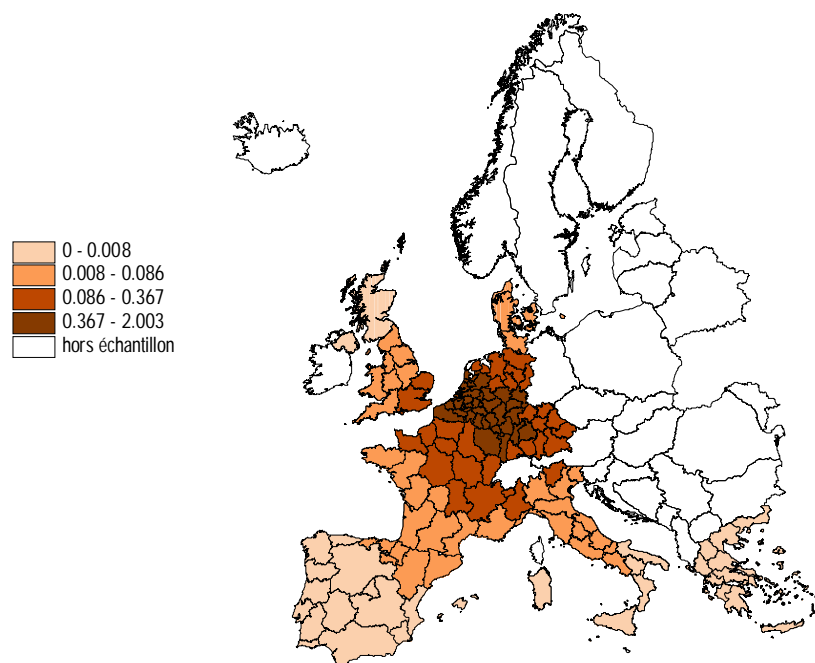


Figure 4.3a : Quartiles de la distribution des impacts médians d'émission pour $W(1)$ et PIB par tête en Ecus

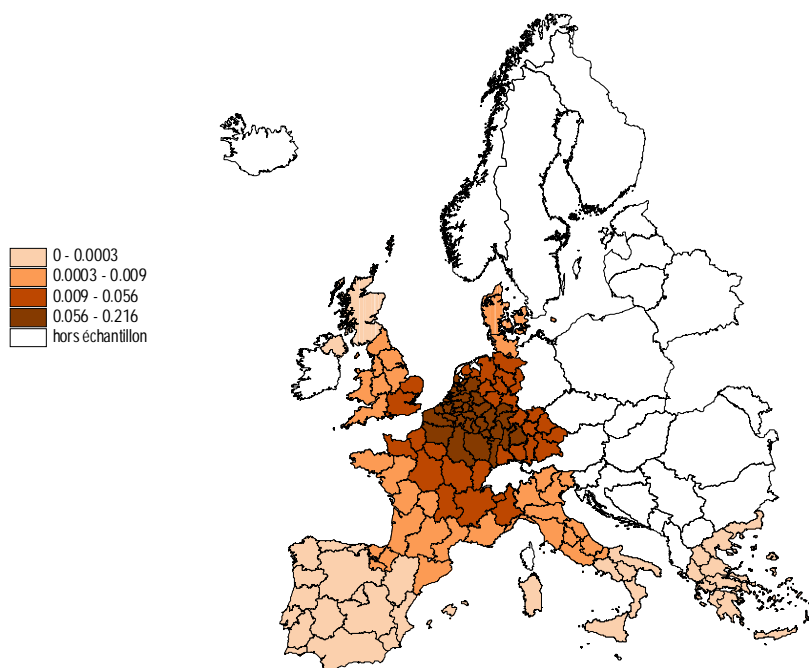


Figure 4.3b : Quartiles de la distribution des impacts médians d'émission pour $W(1)$ et PIB par tête en SPA

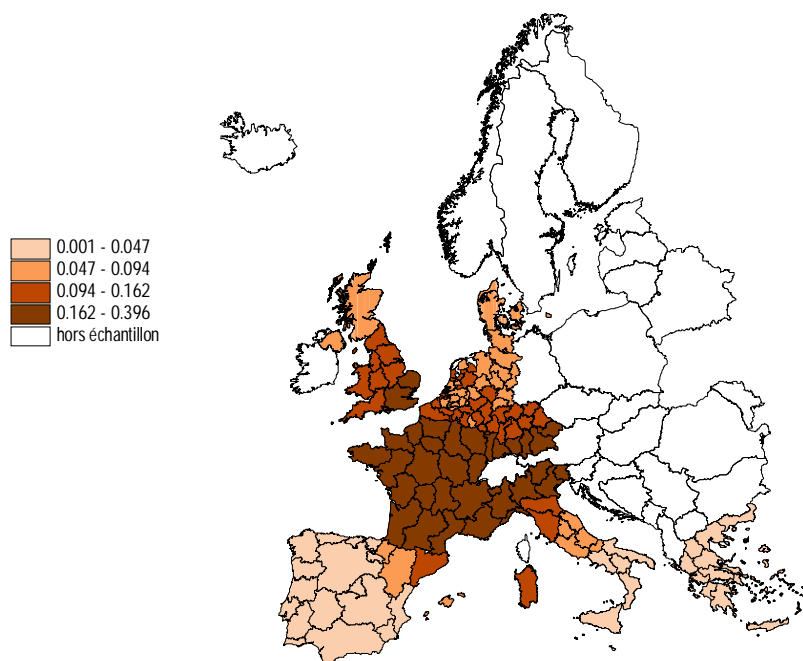


Figure 4.4a : Quartiles de la distribution des impacts médians de réception pour $W(1)$ et PIB par tête en Ecus

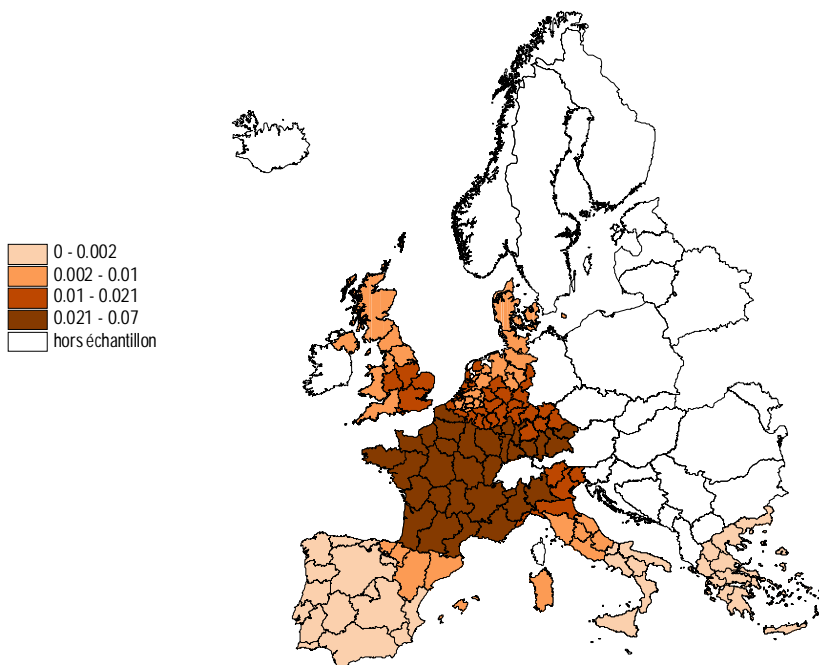


Figure 4.4b : Quartiles de la distribution des impacts médians de réception pour $W(1)$ et PIB par tête en SPA

SECTION 42 :

β -CONVERGENCE ET HETEROGENEITE SPATIALE

Les résultats obtenus dans le chapitre 3 indiquent que la répartition des PIB par tête régionaux est caractérisée par une hétérogénéité spatiale sous la forme d'une polarisation Nord-Sud. Cette hétérogénéité devrait donc être intégrée dans le modèle de β -convergence. Cependant, comme nous l'avons souligné dans le chapitre 2, des liens complexes unissent autocorrélation et hétérogénéité spatiales. Trois aspects sont alors étudiés dans cette section.

Premièrement, il faut identifier les différentes concentrations spatiales de richesse afin de déterminer un modèle de convergence à plusieurs régimes. Chaque regroupement spatial contient toutes les régions connectées par un critère d'association spatiale qui diffère entre les groupes. Deuxièmement, l'inférence statistique basée sur les MCO n'est pas fiable lorsqu'il y a hétérogénéité et/ou autocorrélation spatiale. Il faut donc tester si ces effets sont présents et il est nécessaire d'ajuster les tests d'hétéroscédasticité et d'instabilité structurelle en présence d'autocorrélation spatiale (Anselin, 1988b, 1990a, 1990b). Troisièmement, le rôle joué par les effets de débordement géographiques dans la convergence des régions européennes doit être reconsidéré. Dans la section précédente, nous avons montré que lorsque l'autocorrélation spatiale est détectée dans le modèle de β -convergence absolue, il est nécessaire d'estimer une spécification intégrant des effets de débordement géographiques dans le processus de convergence. Pourtant, puisque l'hétérogénéité spatiale est maintenant intégrée dans l'estimation du modèle de β -convergence, il faut à nouveau rechercher la spécification la plus appropriée et ajuster les différents tests de spécification afin d'obtenir des estimations fiables des effets de débordement géographiques sur la croissance régionale en Europe.

Dans cette section, nous définissons et appliquons notre méthodologie empirique qui étend celle développée par Durlauf et Johnson (1995). Elle prend en compte explicitement les effets spatiaux précédemment définis, dans le contexte du modèle de β -convergence standard. Dans la première étape, nous détectons les clubs spatiaux à l'aide de l'analyse exploratoire des données spatiales. Dans la seconde étape, nous estimons un modèle de β -convergence absolue par les MCO et appliquons divers tests afin de détecter la présence de l'autocorrélation et de l'hétérogénéité spatiales. Nous proposons enfin la spécification la plus appropriée pour ces deux problèmes⁶⁸.

421 Les clubs de convergence en Europe

Plusieurs études se sont intéressées à la question des clubs de convergence en Europe. Nous présentons d'abord les principaux résultats obtenus dans ces études (paragraphe 421.1)

⁶⁸ Une méthodologie similaire a également été utilisée dans un contexte différent par Baller *et al.* (2001).

et les principales limites rencontrées. Compte tenu de ces limites, nous exposons la façon dont nous avons choisi de construire les clubs (paragraphe 4212).

421.1 Les clubs de convergence en Europe, un tour d'horizon

Dans le cadre des régions européennes, peu d'auteurs se sont penchés sur l'existence de clubs de convergence. Les méthodologies appliquées sont diverses et conduisent à des résultats contrastés.

Neven et Gouyette (1994, 1995) définissent des régimes d'une façon arbitraire. Le Sud est constitué des régions de l'Espagne, du Portugal, de la Grèce et de l'Italie au sud de la région Lazio. Le Nord est constitué par les autres régions de l'échantillon. Sur la période 1980-1989, le coefficient associé au PIB par tête initial n'est significatif ni pour le régime Nord, ni pour le régime Sud. En revanche, les auteurs détectent une convergence entre les régions du Sud sur la période 1980-1985 et une convergence entre les régions du Nord sur la période 1985-1989. Cependant, les intervalles de temps étant très faibles, ces résultats devraient être reconsidérés pour une période plus longue.

Armstrong (1995b) et Dewhurst et Mutis-Gaitan (1995) utilisent la méthodologie suggérée par Chatterji (1992) pour détecter les clubs de convergence. Dewhurst et Mutis-Gaitan (1995) estiment des régressions, allant jusqu'à l'ordre 6, mesurant la convergence par rapport à la région *leader*, Hamburg. Deux problèmes doivent être soulignés dans cette analyse : d'une part, les biais éventuels dus à la multicolinéarité ne sont pas évoqués et d'autre part, le PIB par tête de la région Hamburg est fortement surestimé par les flux de migrants journaliers. Ainsi, Armstrong (1995b) reprend ces analyses avec une régression cubique et considère l'Ile-de-France comme la région *leader*. Il apparaît qu'il n'existe pas de clubs de convergence au sein des régions européennes qui convergent toutes vers un même point, situé à un niveau inférieur du PIB par tête de l'Ile de France. Cependant, Armstrong (1995b) admet que ces résultats pourraient être remis en cause avec un degré de désagrégation spatiale plus important et un plus grand nombre d'observations (son échantillon ne comprend que 85 régions).

Fagerberg et Verspagen (1996) utilisent la méthode des arbres de régression de Durlauf et Johnson (1995) avec quatre variables de classement : le PIB par tête initial, le R&D, le chômage et l'aide à l'investissement de l'Union Européenne. Ils détectent alors 3 clubs de convergence. Cependant, la taille de leur échantillon est relativement faible (72 régions) et les régions pauvres de Grèce, d'Espagne et du Portugal ne sont pas considérées.

Beine *et al.* (1999) divisent arbitrairement leur échantillon en distinguant trois groupes : l'échantillon total et deux groupes rassemblant respectivement 75% et 50% des régions initialement les plus riches. L'hypothèse de convergence en clubs est rejetée. Cependant, leurs résultats ont été obtenus avec des régions au niveau NUTS 1 et pourraient être remis en cause, comme ceux d'Armstrong (1995b), par l'utilisation d'un niveau de désagrégation spatiale plus élevé.

Jean-Pierre (1999) utilise la méthode des régressions réordonnées et distingue trois groupes différents. Le premier comporte les régions du Sud (objectif 1) et l'Irlande. Le second est plus hétérogène alors que le dernier comprend les régions riches d'Allemagne, de Belgique, des Pays-Bas et de France. Le premier groupe des régions les plus pauvres se caractérise par une absence de convergence absolue et conditionnelle. Les deux autres groupes convergent lentement. Ces résultats contrastent avec ceux trouvés par Beine et Jean-Pierre (2000) qui, en utilisant la même méthodologie mais un échantillon différent, distinguent deux groupes de convergence. Les régions pauvres convergent entre elles à un rythme lent alors que l'hypothèse de convergence des régions plus aisées est rejetée.

Face à ces différents résultats, nous proposons d'analyser la convergence des groupes en Europe avec un échantillon plus large que ceux précédemment utilisés. Par ailleurs, nous tenons compte explicitement de la dimension spatiale des données qui est négligée dans toutes ces études.

421.2 Détection des clubs spatiaux

Compte tenu de notre contexte de β -convergence, nous avons choisi de définir les groupes sur la base du niveau initial du PIB par tête en logarithmes. Comme le soulignent Durlauf et Johnson (1995), l'utilisation de variables de classement qui sont connues au début de la période est nécessaire pour éviter le problème du biais de sélection d'échantillon soulevé par De Long (1988). L'instabilité spatiale sous la forme de régimes spatiaux est étudiée à l'aide du diagramme de Moran pour les PIB par tête en Ecus (paragraphe 421.21) et des statistiques de Getis-Ord pour les PIB par tête en SPA (paragraphe 421.22).

421.21 Instabilité spatiale des PIB par tête en Ecus en 1980

Le diagramme de Moran permet de classer les régions européennes en quatre groupes selon les valeurs du PIB par tête initial et les valeurs de la variable spatiale décalée du PIB par tête initial. En outre, le diagramme permet de capter des interactions entre ces variables. Le tableau 4.4 donne, pour toutes les matrices de poids, le quadrant dans lequel se trouve une région. Deux schémas apparaissent, selon que l'analyse est menée avec les matrices de distance inverses ou les matrices des plus proches voisins.

Dans le premier cas, presque toutes les régions européennes sont caractérisées par une association spatiale positive puisque 90 régions sont du type HH et 45 régions sont du type BB. Pour $W(1)$, seules 3 régions anglaises relèvent d'une association spatiale de valeurs dissemblables : Wales et Northern Ireland se trouvent dans le quadrant BH, alors que Scotland est dans le quadrant HB. Ceci suggère une forme d'hétérogénéité spatiale dans les économies régionales européennes. Le processus de convergence, s'il existe, pourrait être différent entre les régimes. Nous allons donc considérer deux clubs spatiaux, que nous appelons régime

« Nord » et régime « Sud ». Puisque Wales, Scotland et Northern Ireland sont éliminées⁶⁹, notre nouvel échantillon contient 135 régions. Cet échantillon est le premier que nous considérons dans cette section. Il est défini de la façon suivante :

Echantillon 1 (135 régions) :

1/ *Nord* = {France, Allemagne, Pays-Bas, Belgique, Danemark, Luxembourg, Royaume-Uni (sauf Wales, Scotland et Northern Ireland) et Italie du nord (Piemonte, Valle d'Aosta, Liguria, Lombardia, Trentino, Veneto, Friuli, Emilia-Romagna, Toscana)}.

2/ *Sud* = {Portugal, Espagne, Grèce et Italie du sud (Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicate, Calabria, Sicilia et Sardegna)}.

Les diagrammes de Moran calculés avec les autres matrices de poids $W(2)$ et $W(3)$ conduisent sensiblement aux mêmes clubs : la seule différence est la présence de Scotland dans le régime du Nord. Ceci indique à nouveau la robustesse des résultats sur le choix de la matrice de poids. En outre, la polarisation observée semble être persistante sur toute la période puisque la composition des clubs définie par les diagrammes de Moran, calculés pour chaque année, reste globalement inchangée.

L'utilisation des matrices des plus proches voisins conduit à des résultats légèrement différents (cf. tableau 4.4). Ainsi, la région française Aquitaine est toujours dans le quadrant HB et Scotland passe dans le quadrant HH. Tous les autres résultats sont identiques. En conséquence, le deuxième échantillon de 135 régions, que nous utilisons avec les matrices des plus proches voisins, est défini de la façon suivante :

Echantillon 2 (135 régions) :

1/ *Nord* = {France (sauf Aquitaine), Allemagne, Pays-Bas, Belgique, Danemark, Luxembourg, Royaume-Uni (sauf Wales et Northern Ireland) et Italie du nord (Piemonte, Valle d'Aosta, Liguria, Lombardia, Trentino, Veneto, Friuli, Emilia-Romagna, Toscana)}.

2/ *Sud* = {Portugal, Espagne, Grèce et Italie du sud (Umbria, Marche, Lazio, Abruzzo, Molise, Campania, Puglia, Basilicate, Calabria, Sicilia et Sardegna)}.

421.22 Instabilité spatiale des PIB par tête en SPA en 1980

Pour les PIB par tête exprimés en SPA, l'approche précédente n'est plus applicable. En effet, comme indiqué dans le chapitre 3, les diagrammes de Moran contiennent alors beaucoup de régions atypiques (c'est-à-dire les régions qui se trouvent dans le quadrant HB ou dans le quadrant BH). Par conséquent, l'approche précédente nous conduirait à éliminer beaucoup trop de régions de l'échantillon, les régions atypiques n'étant pas assez nombreuses pour former de manière significative d'autres clubs de convergence.

⁶⁹ Les clubs spatiaux (BH) et (HB) ne contiennent respectivement que 2 et 1 régions. Ils sont donc omis à cause du faible nombre d'observations et de l'insuffisance de degrés de liberté pour la seconde étape de notre analyse.

	Diagramme de Moran (PIB par tête en Ecus)							Diagramme de Moran (PIB par tête en Ecus)					
	W(1)	W(2)	W(3)	W(10)	W(15)	W(20)		W(1)	W(2)	W(3)	W(10)	W(15)	W(20)
BE1	HH	HH	HH	HH	HH	HH	FR51	HH	HH	HH	HH	HH	HH
BE21	HH	HH	HH	HH	HH	HH	FR52	HH	HH	HH	HH	HH	HH
BE22	HH	HH	HH	HH	HH	HH	FR53	HH	HH	HH	HH	HH	HH
BE23	HH	HH	HH	HH	HH	HH	FR61	HH	HH	HH	HB	HB	HB
BE24	HH	HH	HH	HH	HH	HH	FR62	HH	HH	HH	HH	HH	HH
BE25	HH	HH	HH	HH	HH	HH	FR63	HH	HH	HH	HH	HH	HH
BE31	HH	HH	HH	HH	HH	HH	FR71	HH	HH	HH	HH	HH	HH
BE32	HH	HH	HH	HH	HH	HH	FR72	HH	HH	HH	HH	HH	HH
BE33	HH	HH	HH	HH	HH	HH	FR81	HH	HH	HH	HH	HH	HH
BE34	HH	HH	HH	HH	HH	HH	FR82	HH	HH	HH	HH	HH	HH
BE35	HH	HH	HH	HH	HH	HH	GR11	BB	BB	BB	BB	BB	BB
DE11	HH	HH	HH	HH	HH	HH	GR12	BB	BB	BB	BB	BB	BB
DE12	HH	HH	HH	HH	HH	HH	GR13	BB	BB	BB	BB	BB	BB
DE13	HH	HH	HH	HH	HH	HH	GR14	BB	BB	BB	BB	BB	BB
DE14	HH	HH	HH	HH	HH	HH	GR21	BB	BB	BB	BB	BB	BB
DE21	HH	HH	HH	HH	HH	HH	GR22	BB	BB	BB	BB	BB	BB
DE22	HH	HH	HH	HH	HH	HH	GR23	BB	BB	BB	BB	BB	BB
DE23	HH	HH	HH	HH	HH	HH	GR24	BB	BB	BB	BB	BB	BB
DE24	HH	HH	HH	HH	HH	HH	GR25	BB	BB	BB	BB	BB	BB
DE25	HH	HH	HH	HH	HH	HH	GR3	BB	BB	BB	BB	BB	BB
DE26	HH	HH	HH	HH	HH	HH	GR41	BB	BB	BB	BB	BB	BB
DE27	HH	HH	HH	HH	HH	HH	GR42	BB	BB	BB	BB	BB	BB
DE5	HH	HH	HH	HH	HH	HH	GR43	BB	BB	BB	BB	BB	BB
DE6	HH	HH	HH	HH	HH	HH	IT11	HH	HH	HH	HH	HH	HH
DE71	HH	HH	HH	HH	HH	HH	IT12	HH	HH	HH	HH	HH	HH
DE72	HH	HH	HH	HH	HH	HH	IT13	HH	HH	HH	HH	HH	HH
DE73	HH	HH	HH	HH	HH	HH	IT2	HH	HH	HH	HH	HH	HH
DE91	HH	HH	HH	HH	HH	HH	IT31	HH	HH	HH	HH	HH	HH
DE92	HH	HH	HH	HH	HH	HH	IT32	HH	HH	HH	HH	HH	HH
DE93	HH	HH	HH	HH	HH	HH	IT33	HH	HH	HH	HH	HH	HH
DE94	HH	HH	HH	HH	HH	HH	IT4	HH	HH	HH	HH	HH	HH
DEA1	HH	HH	HH	HH	HH	HH	IT51	HH	HH	HH	HH	HH	HH
DEA2	HH	HH	HH	HH	HH	HH	IT52	BB	BB	BB	BB	BB	BB
DEA3	HH	HH	HH	HH	HH	HH	IT53	BB	BB	BB	BB	BB	BB
DEA4	HH	HH	HH	HH	HH	HH	IT6	BB	BB	BB	BB	BB	BB
DEA5	HH	HH	HH	HH	HH	HH	IT71	BB	BB	BB	BB	BB	BB
DEB1	HH	HH	HH	HH	HH	HH	IT72	BB	BB	BB	BB	BB	BB
DEB2	HH	HH	HH	HH	HH	HH	IT8	BB	BB	BB	BB	BB	BB
DEB3	HH	HH	HH	HH	HH	HH	IT91	BB	BB	BB	BB	BB	BB
DEC	HH	HH	HH	HH	HH	HH	IT92	BB	BB	BB	BB	BB	BB
DEF	HH	HH	HH	HH	HH	HH	IT93	BB	BB	BB	BB	BB	BB
DK	HH	HH	HH	HH	HH	HH	ITA	BB	BB	BB	BB	BB	BB
ES11	BB	BB	BB	BB	BB	BB	ITB	BB	BB	BB	BB	BB	BB
ES12	BB	BB	BB	BB	BB	BB	LU	HH	HH	HH	HH	HH	HH
ES13	BB	BB	BB	BB	BB	BB	NL12	HH	HH	HH	HH	HH	HH
ES21	BB	BB	BB	BB	BB	BB	NL13	HH	HH	HH	HH	HH	HH
ES22	BB	BB	BB	BB	BB	BB	NL2	HH	HH	HH	HH	HH	HH
ES23	BB	BB	BB	BB	BB	BB	NL31	HH	HH	HH	HH	HH	HH
ES24	BB	BB	BB	BB	BB	BB	NL32	HH	HH	HH	HH	HH	HH
ES3	BB	BB	BB	BB	BB	BB	NL33	HH	HH	HH	HH	HH	HH
ES41	BB	BB	BB	BB	BB	BB	NL34	HH	HH	HH	HH	HH	HH
ES42	BB	BB	BB	BB	BB	BB	NL41	HH	HH	HH	HH	HH	HH
ES43	BB	BB	BB	BB	BB	BB	NL42	HH	HH	HH	HH	HH	HH
ES51	BB	BB	BB	BB	BB	BB	PT11	BB	BB	BB	BB	BB	BB
ES52	BB	BB	BB	BB	BB	BB	PT12	BB	BB	BB	BB	BB	BB
ES53	BB	BB	BB	BB	BB	BB	PT13	BB	BB	BB	BB	BB	BB
ES61	BB	BB	BB	BB	BB	BB	PT14	BB	BB	BB	BB	BB	BB
ES62	BB	BB	BB	BB	BB	BB	PT15	BB	BB	BB	BB	BB	BB
FR1	HH	HH	HH	HH	HH	HH	UK1	HH	HH	HH	HH	HH	HH
FR21	HH	HH	HH	HH	HH	HH	UK2	HH	HH	HH	HH	HH	HH
FR22	HH	HH	HH	HH	HH	HH	UK3	HH	HH	HH	HH	HH	HH
FR23	HH	HH	HH	HH	HH	HH	UK4	HH	HH	HH	HH	HH	HH
FR24	HH	HH	HH	HH	HH	HH	UK5	HH	HH	HH	HH	HH	HH
FR25	HH	HH	HH	HH	HH	HH	UK6	HH	HH	HH	HH	HH	HH
FR26	HH	HH	HH	HH	HH	HH	UK7	HH	HH	HH	HH	HH	HH
FR3	HH	HH	HH	HH	HH	HH	UK8	HH	HH	HH	HH	HH	HH
FR41	HH	HH	HH	HH	HH	HH	UK9	BH	BH	BH	BH	BH	BH
FR42	HH	HH	HH	HH	HH	HH	UKA	HB	HH	HH	HH	HH	HH
FR43	HH	HH	HH	HH	HH	HH	UKB	BH	BH	BH	BH	BH	BH

Tableau 4.4 : Détection des clubs spatiaux basée sur le diagramme de Moran

Nous avons choisi dans ce cas de baser la détermination des clubs spatiaux sur le signe des statistiques de Getis-Ord. Le tableau 4.5 reporte, pour toutes les matrices de poids, le signe de la statistique associée à une région. Plusieurs différences apparaissent selon la matrice de poids utilisée.

Tout d'abord, le signe de la statistique de Getis-Ord associée à Bruxelles (Belgique) est négatif pour $W(1)$, $W(2)$ et $W(3)$ mais positif pour $W(10)$, $W(15)$ et $W(20)$. Ensuite, les signes diffèrent quelque peu pour les régions Aquitaine (France), Abruzzo (Italie) et Northern Ireland (Royaume-Uni). Aucune de ces statistiques n'est significative. Dans ces conditions, nous avons choisi de placer Abruzzo dans le régime « Sud » car il s'agit d'une région éligible à objectif 1. Les trois autres régions sont dans le régime « Nord ». Le troisième échantillon de 138 régions que nous considérons est donc le suivant :

Echantillon 3 (138 régions) :

1/ *Nord* = {France, Allemagne, Pays-Bas, Belgique, Danemark, Luxembourg, Royaume-Uni et Italie du nord (Piemonte, Valle d'Aosta, Liguria, Lombardia, Trentino, Veneto, Friuli, Emilia-Romagna, Toscana)}.

2/ *Sud* = {Portugal, Espagne, Grèce et Italie du sud (Umbria, Marche, Lazio, Abbruzzo, Molise, Campania, Puglia, Basilicate, Calabria, Sicilia et Sardegna)}.

Idéalement, la détermination des différents régimes devrait être endogène. Cependant, à notre connaissance, une procédure permettant simultanément de déterminer les clubs de convergence et de prendre en compte l'autocorrélation spatiale n'a pas encore été proposée. En effet, ceci est beaucoup plus complexe à mettre en œuvre que dans un cadre non-spatial à cause de la matrice de poids et de l'ordonnement spatial des observations (Anselin et Cho, 2000).

La seconde étape de notre analyse concerne l'estimation du modèle de β -convergence standard intégrant simultanément l'hétérogénéité et l'autocorrélation spatiales. Si l'on suit les résultats de simulation présentés par Anselin (1990a) sur le seuil et la puissance des tests traditionnels d'instabilité structurelle en présence d'erreurs spatialement autocorrélées, nous sommes potentiellement dans le pire des cas : autocorrélation spatiale globale positive et deux régimes correspondant à des observations « compactes ». Les tests standards d'instabilité structurelle ne sont pas fiables dans ce cas.

En conséquence, nous définissons l'approche méthodologique à utiliser pour les résultats économétriques, en suivant la suggestion d'Anselin (1990b) et nous testons d'abord la présence d'autocorrélation spatiale. Cette approche empirique peut poser des problèmes de *pretest*, qui invalident la distribution asymptotique usuelle des différents tests. Cependant, les résultats de simulation présentés par Anselin (1990a) indiquent que ce problème n'est pas trop nuisible dans ce cas particulier.

	Statistiques Getis-Ord (PIB par tête en SPA)							Statistiques Getis-Ord (PIB par tête en SPA)					
	W(1)	W(2)	W(3)	W(10)	W(15)	W(20)		W(1)	W(2)	W(3)	W(10)	W(15)	W(20)
BE1	-	-	-	+	+	+	FR51	+	+	+	+	+	+
BE21	+	+	+	+	+	+	FR52	+	+	+	+	+	+
BE22	+	+	+	+	+	+	FR53	+	+	+	+	+	+
BE23	+	+	+	+	+	+	FR61	-	+	+	-	-	-
BE24	+	+	+	+	+	+	FR62	+	+	+	+	+	+
BE25	+	+	+	+	+	+	FR63	+	+	+	+	+	+
BE31	+	+	+	+	+	+	FR71	+	+	+	+	+	+
BE32	+	+	+	+	+	+	FR72	+	+	+	+	+	+
BE33	+	+	+	+	+	+	FR81	+	+	+	+	+	+
BE34	+	+	+	+	+	+	FR82	+	+	+	+	+	+
BE35	+	+	+	+	+	+	GR11	-	-	-	-	-	-
DE11	+	+	+	+	+	+	GR12	-	-	-	-	-	-
DE12	+	+	+	+	+	+	GR13	-	-	-	-	-	-
DE13	+	+	+	+	+	+	GR14	-	-	-	-	-	-
DE14	+	+	+	+	+	+	GR21	-	-	-	-	-	-
DE21	+	+	+	+	+	+	GR22	-	-	-	-	-	-
DE22	+	+	+	+	+	+	GR23	-	-	-	-	-	-
DE23	+	+	+	+	+	+	GR24	-	-	-	-	-	-
DE24	+	+	+	+	+	+	GR25	-	-	-	-	-	-
DE25	+	+	+	+	+	+	GR3	-	-	-	-	-	-
DE26	+	+	+	+	+	+	GR41	-	-	-	-	-	-
DE27	+	+	+	+	+	+	GR42	-	-	-	-	-	-
DE5	+	+	+	+	+	+	GR43	-	-	-	-	-	-
DE6	+	+	+	+	+	+	IT11	+	+	+	+	+	+
DE71	+	+	+	+	+	+	IT12	+	+	+	+	+	+
DE72	+	+	+	+	+	+	IT13	+	+	+	+	+	+
DE73	+	+	+	+	+	+	IT2	+	+	+	+	+	+
DE91	+	+	+	+	+	+	IT31	+	+	+	+	+	+
DE92	+	+	+	+	+	+	IT32	+	+	+	+	+	+
DE93	+	+	+	+	+	+	IT33	+	+	+	+	+	+
DE94	+	+	+	+	+	+	IT4	+	+	+	+	+	+
DEA1	+	+	+	+	+	+	IT51	+	+	+	+	+	+
DEA2	+	+	+	+	+	+	IT52	+	+	+	+	+	+
DEA3	+	+	+	+	+	+	IT53	+	+	+	+	+	+
DEA4	+	+	+	+	+	+	IT6	+	+	+	+	-	+
DEA5	+	+	+	+	+	+	IT71	-	-	+	+	+	-
DEB1	+	+	+	+	+	+	IT72	-	-	-	-	+	-
DEB2	+	+	+	+	+	+	IT8	-	-	-	-	-	-
DEB3	+	+	+	+	+	+	IT91	-	-	-	-	-	-
DEC	+	+	+	+	+	+	IT92	-	-	-	-	-	-
DEF	+	+	+	+	+	+	IT93	-	-	-	-	-	-
DK	+	+	+	+	+	+	ITA	-	-	-	-	-	-
ES11	-	-	-	-	-	-	ITB	+	+	+	+	+	+
ES12	-	-	-	-	-	-	LU	+	+	+	+	+	+
ES13	-	-	-	-	-	-	NL12	+	+	+	+	+	+
ES21	-	-	-	-	-	-	NL13	+	+	+	+	+	+
ES22	-	-	-	-	-	-	NL2	+	+	+	+	+	+
ES23	-	-	-	-	-	-	NL31	+	+	+	+	+	+
ES24	-	-	-	-	-	-	NL32	+	+	+	+	+	+
ES3	-	-	-	-	-	-	NL33	+	+	+	+	+	+
ES41	-	-	-	-	-	-	NL34	+	+	+	+	+	+
ES42	-	-	-	-	-	-	NL41	+	+	+	+	+	+
ES43	-	-	-	-	-	-	NL42	+	+	+	+	+	+
ES51	-	-	-	-	-	-	PT11	-	-	-	-	-	-
ES52	-	-	-	-	-	-	PT12	-	-	-	-	-	-
ES53	-	-	-	-	-	-	PT13	-	-	-	-	-	-
ES61	-	-	-	-	-	-	PT14	-	-	-	-	-	-
ES62	-	-	-	-	-	-	PT15	-	-	-	-	-	-
FR1	+	+	+	+	+	+	UK1	-	+	+	-	+	+
FR21	+	+	+	+	+	+	UK2	+	+	+	-	+	+
FR22	+	+	+	+	+	+	UK3	+	+	+	+	+	+
FR23	+	+	+	+	+	+	UK4	+	+	+	+	+	+
FR24	+	+	+	+	+	+	UK5	+	+	+	+	+	+
FR25	+	+	+	+	+	+	UK6	+	+	+	+	+	+
FR26	+	+	+	+	+	+	UK7	+	+	+	-	+	+
FR3	+	+	+	+	+	+	UK8	-	+	+	-	+	+
FR41	+	+	+	+	+	+	UK9	+	+	+	+	+	+
FR42	+	+	+	+	+	+	UKA	-	+	+	-	+	+
FR43	+	+	+	+	+	+	UKB	-	+	+	-	+	+

Tableau 4.5 : Détection des clubs spatiaux basée sur les statistiques Getis-Ord

422 Résultats des estimations

La procédure d'estimation est la suivante. Nous commençons par estimer le modèle de β -convergence absolue par les MCO et nous effectuons différents tests afin de détecter, d'une part, la présence d'autocorrélation spatiale et d'autre part, la présence d'hétérogénéité spatiale, sous la forme d'une hétéroscédasticité en groupe et/ou d'une instabilité structurelle entre les régimes spatiaux précédemment définis (paragraphe 422.1). Les résultats des tests d'autocorrélation spatiale nous orientent vers une spécification du modèle de β -convergence avec autocorrélation des erreurs que nous estimons (paragraphe 422.2). L'hétéroscédasticité n'étant cependant pas complètement éliminée de ce modèle, deux types de spécifications économétriques permettant d'intégrer l'hétérogénéité sont estimés : un modèle avec hétéroscédasticité en groupes et un modèle à deux régimes (paragraphe 422.3). Ces différents modèles ne permettant toujours pas de purger l'autocorrélation spatiale, nous estimons un modèle à deux régimes et erreurs autocorrélées (paragraphe 422.4). Enfin, comme les modèles précédents supposent que la forme de l'autocorrélation spatiale est identique dans les deux régimes, nous spécifions dans une dernière partie un modèle qui permet l'estimation d'effets spatiaux différenciés internes à chaque régime et entre les régimes (paragraphe 422.5).

422.1 Estimation par les MCO du modèle de β -convergence absolue et tests

Le point de départ est le modèle de β -convergence absolue [4.12]. Nous présentons les résultats pour la matrice de distance $W(1)$ et pour la matrice des 10 plus proches voisins $W(10)$. Les résultats de l'estimation du modèle [4.12] pour les PIB par tête en Ecus et en SPA sont reportés dans le tableau 4.6. Dans le second cas, ils correspondent aux estimations déjà fournies dans la section 41. Seul un test d'hétéroscédasticité supplémentaire est donné. Pour les PIB par tête en Ecus, les résultats diffèrent légèrement puisque les estimations ont été effectuées sur des échantillons différents de ceux de la section précédente.

Pour tous les échantillons considérés, le coefficient associé au PIB par tête initial est significatif et négatif. Cependant, la vitesse de convergence estimée varie entre 0,84% et 0,85%. Ces résultats confirment la faiblesse du processus de convergence en Europe.

Différents tests sur les erreurs ont été effectués. La normalité des erreurs est toujours rejetée. Concernant l'hétéroscédasticité, les résultats diffèrent entre les deux variables. Pour les PIB par tête en Ecus, l'hypothèse d'homoscédasticité est toujours rejetée quel que soit le test effectué : test de White, test de Breusch-Pagan contre la variable explicative y_0 ou test de Breusch-Pagan contre D_1 , la variable muette pour le régime « Nord ». Les tests d'autocorrélation spatiale pointent tous vers une spécification avec autocorrélation des erreurs. En outre, JLM_1 et JLM_2 , qui testent l'hypothèse jointe d'une autocorrélation spatiale des erreurs et d'une hétéroscédasticité (cf. paragraphe 233.31 du chapitre 2, p. 180), sont tous significatifs.

Pour les PIB par tête en SPA, l'hypothèse nulle d'homoscédasticité n'est jamais rejetée, même contre la variable de régime D_1 . Puisque la normalité est fortement rejetée, seul le résultat pour le test d'hétéroscédasticité de Koenker-Bassett (1982) est fourni. Les tests d'autocorrélation spatiale indiquent par ailleurs que le modèle de β -convergence absolue est mal spécifié dû à l'omission de l'autocorrélation spatiale des erreurs. Il n'a pas été possible d'effectuer les tests JLM_1 et JLM_2 . En effet, ces derniers sont calculés comme la somme de la statistique LM_{ERR} et de la statistique de Breusch-Pagan, qui n'a pas pu être obtenue.

Nous devons donc faire face à deux problèmes interconnectés : l'hétérogénéité et l'autocorrélation spatiales. En présence d'hétéroscédasticité, les résultats des tests d'autocorrélation spatiale ne sont pas fiables et vice-versa. Par conséquent, ils doivent être interprétés avec prudence. Par exemple, lorsque les tests indiquent de l'hétéroscédasticité, cette dernière peut être causée par la présence d'autocorrélation spatiale (McMillen, 1992).

Suivant la suggestion d'Anselin (1990a), nous considérons dans un premier temps le problème de l'autocorrélation spatiale.

422.2 Traitement de l'autocorrélation spatiale

Les résultats de l'estimation du modèle de β -convergence avec autocorrélation spatiale des erreurs (modèle [4.6]) sont fournis dans le tableau 4.7. Pour chaque régression, deux méthodes d'estimation ont été utilisées, la méthode du maximum de vraisemblance (MV) et la méthode des moments généralisés (GMM).

Concernant les deux échantillons de régions avec PIB par tête en Ecus, les résultats sont qualitativement similaires à ceux obtenus avec l'échantillon complet. En effet, les coefficients sont tous significatifs. Du point de vue de la convergence, le coefficient $\hat{\beta}$ est supérieur à celui obtenu dans le modèle de β -convergence absolue estimé par les MCO. En effet, la vitesse de convergence est de 1,2% et la demi-vie est réduite à 63 ans avec $W(1)$. Cette vitesse de convergence atteint 1,56% pour $W(10)$. La vitesse de convergence est donc plus importante tout en restant relativement faible. Il apparaît dans tous les cas une forte autocorrélation spatiale positive des erreurs : $\hat{\lambda}$ varie entre 0,788 et 0,828.

Différents tests ont été effectués à partir de cette spécification. Le test du ratio de vraisemblance du facteur commun indique que la restriction $\gamma + \lambda\beta = 0$ ne peut pas être rejetée. Ainsi, le modèle [4.6] peut être réécrit comme un modèle de Durbin contraint [4.9] :

$$g_T = \alpha(I - \lambda W)S + \beta y_0 + \lambda Wg_T + \gamma W y_0 + u$$

avec $\gamma = -\lambda\beta$. Cependant, ce coefficient n'est jamais significatif au seuil de 5%. Ainsi, le taux de croissance d'une région est influencé par le taux de croissance de ses régions voisines mais il n'est pas influencé par le PIB par tête initial de ses régions voisines.

Estimation (MCO – White)	Ecus (135 régions)		SPA (138 régions)	
	W(1) - ech.1	W(10) - ech.2	W(1) - ech.3	W(10) - ech.3
$\hat{\alpha}$	0,130 (0,000)	0,130 (0,000)	0,130 (0,000)	0,130 (0,000)
$\hat{\beta}$	-0,00797 (0,002)	-0,00797 (0,002)	-0,00790 (0,002)	-0,00790 (0,002)
Vitesse de convergence	0,85% (0,000)	0,84% (0,000)	0,85% (0,000)	0,85% (0,000)
Demi-vie	87	87	87	87
R^2	0,14	0,13	0,11	0,11
LIK	446,35	446,32	499,27	499,27
AIC	-888,69	-888,64	-994,54	-994,54
BIC	-882,88	-882,83	-988,69	-988,69
$\hat{\sigma}^2$	$7,984.10^{-5}$	$7,987.10^{-5}$	$4,279.10^{-5}$	$4,279.10^{-5}$
Tests				
JB	8,50 (0,014)	8,63 (0,0134)	12,47 (0,001)	12,47 (0,001)
BP ou KB* vs. y_0	5,85 (0,015)	5,88 (0,015)	1,20* (0,272)	1,20* (0,272)
BP ou KB* vs. D_1	14,57 (0,000)	14,68 (0,000)	0,297* (0,586)	0,297* (0,586)
White	28,39 (0,000)	28,606 (0,000)	0,89 (0,640)	0,89 (0,640)
I de Moran	12,94 (0,000)	13,32 (0,000)	5,62 (0,000)	5,81 (0,000)
LM_{ERR}	140,68 (0,000)	140,88 (0,000)	25,12 (0,000)	25,38 (0,000)
RLM_{ERR}	16,61 (0,000)	31,14 (0,000)	3,31 (0,068)	5,14 (0,023)
LM_{LAG}	124,58 (0,000)	113,57 (0,000)	21,841 (0,000)	20,510 (0,000)
RLM_{LAG}	0,509 (0,475)	3,83 (0,050)	0,029 (0,863)	0,275 (0,599)
JLM_1	155,25 (0,000)	155,56 (0,000)	-	-
JLM_2	146,53 (0,000)	146,76 (0,000)	-	-

Tableau 4.6 : Résultats d'estimation pour le modèle de β -convergence absolue

Notes : les valeurs critiques sont entre parenthèses. MCO-White indique l'utilisation de la matrice des variances-covariances convergente en présence d'hétéroscédasticité dans l'estimation par les MCO (White, 1980). LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). JB est le test de normalité des résidus de Jarque-Bera (1987). BP est le test de Breusch-Pagan (1979) pour l'hétéroscédasticité. KB est le test de Koenker-Bassett (1982) pour l'hétéroscédasticité en présence de non-normalité. White est le test de White (1980) d'hétéroscédasticité. I de Moran est le test de Moran adapté aux résidus (Cliff et Ord, 1981). LM_{ERR} est le test du multiplicateur de Lagrange pour une autocorrélation des erreurs et RLM_{ERR} est sa version robuste. LM_{LAG} est le test du multiplicateur de Lagrange pour une variable endogène décalée et RLM_{LAG} est sa version robuste (Anselin *et al.*, 1996). JLM_1 est le test du multiplicateur de Lagrange pour l'hypothèse nulle jointe d'absence d'hétéroscédasticité contre y_0 et d'autocorrélation spatiale. JLM_2 est le test du multiplicateur de Lagrange pour l'hypothèse nulle jointe d'absence d'hétéroscédasticité contre D_1 et d'autocorrélation spatiale (Anselin, 1988a, 1988b).

Concernant l'hétéroscédasticité, le test de Breusch-Pagan spatialement ajusté n'est plus significatif à 5% contre la variable initiale y_0 . Si ce test avait été le seul effectué pour détecter l'hétéroscédasticité dans le modèle avec autocorrélation des erreurs, nous aurions pu conclure que l'hétéroscédasticité trouvée dans le modèle précédent était entièrement due à la présence d'une autocorrélation spatiale. Pourtant, il apparaît que ce test reste significatif contre la variable de régime D_1 . Nous pouvons donc déduire de ces résultats que seule une partie de l'hétéroscédasticité trouvée dans le modèle précédent est due à l'autocorrélation spatiale du terme d'erreur. L'hétéroscédasticité en groupes demeure et doit toujours être prise en compte.

Estimation	Ecus (135 régions)				SPA (138 régions)			
	W(1) - ech.1		W(10) - ech.2		W(1) - ech.3		W(10) - ech.3	
	MV	GMM	MV	GMM	MV	GMM	MV	GMM
$\hat{\alpha}$	0,156 (0,000)	0,157 (0,000)	0,182 (0,000)	0,183 (0,000)	0,141 (0,000)	0,141 (0,000)	0,144 (0,000)	0,145 (0,000)
$\hat{\beta}$	-0,0110 (0,000)	-0,0110 (0,000)	-0,0139 (0,000)	-0,0140 (0,000)	-0,0091 (0,000)	-0,0092 (0,000)	-0,0095 (0,000)	-0,0096 (0,000)
Vitesse de convergence	1,20% (0,000)	1,20% (0,000)	1,56% (0,000)	1,58% (0,000)	0,98% (0,000)	0,99% (0,000)	1,02% (0,000)	1,04% (0,000)
Demi-vie	63	63	49	49	75	75	73	72
$\hat{\lambda}$	0,788 (0,000)	0,828	0,798 (0,000)	0,815	0,531 (0,000)	0,586	0,527 (0,000)	0,563
Sq.- Corr.	0,14	0,14	0,14	0,14	0,11	0,11	0,11	0,11
LIK	483,42	-	484,48	-	507,77	-	508,76	-
AIC	-962,85	-	-964,96	-	-1011,53	-	-1013,52	-
BIC	-957,03	-	-959,15	-	-1005,68	-	-1007,67	-
$\hat{\sigma}^2$	$4,078.10^{-5}$	-	$4,471.10^{-5}$	-	$3,578.10^{-5}$	-	$3,683.10^{-5}$	-
Tests								
BP-S vs. y_0	3,06 (0,080)	-	3,74 (0,053)	-	0,561 (0,454)	-	0,569 (0,450)	-
BP-S vs. D_1	4,27 (0,039)	-	5,36 (0,020)	-	3,179 (0,074)	-	4,124 (0,042)	-
LM_{LAG}^*	0,808 (0,369)	-	0,240 (0,624)	-	2,217 (0,136)	-	0,305 (0,580)	-
LR ou Wald*-facteur commun	0,177 (0,674)	-	0,046* (0,830)	-	0,536 (0,464)	-	0,014* (0,907)	-
$\hat{\gamma} = -\hat{\lambda}\hat{\beta}$	0,008 (0,871)	-	0,011 (0,834)	-	0,005 (0,938)	-	0,005 (0,938)	-

Tableau 4.7 : Résultats d'estimation pour le modèle de β -convergence avec autocorrélation des erreurs

Notes : Les valeurs critiques sont entre parenthèses. MV est la méthode du maximum de vraisemblance. GMM est la méthode des moments généralisés. Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). BP-S est la version ajustée spatialement du test de Breusch-Pagan (1979) pour l'hétéroscédasticité. LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). LR-facteur commun est le test du ratio de vraisemblance du facteur commun et Wald-facteur commun est le test de Wald du facteur commun (Burrige, 1981).

Pour les PIB par tête en SPA, les résultats sont les mêmes que ceux de la section 41. Nous avons rajouté la statistique du test de Breusch-Pagan spatialement ajusté contre la variable de régime D_1 . Pour $W(1)$, la valeur critique associée est de 0,074 alors que pour $W(10)$, elle diminue pour atteindre 0,042. Ces résultats, même moins tranchés que précédemment, pourraient également indiquer la présence d'une hétérogénéité sous la forme de deux régimes.

422.3 Hétérogénéité spatiale : hétéroscédasticité en groupes et/ou instabilité structurelle

L'hétérogénéité spatiale peut être considérée de deux points de vue. Le premier est lié au problème de l'hétéroscédasticité sous la forme d'une hétéroscédasticité en groupes entre les régimes précédemment définis. Le second est lié au problème de l'instabilité structurelle entre les deux régimes et peut en outre être associé à l'hétéroscédasticité en groupes.

Pour tenir compte de l'hétéroscédasticité en groupes, nous estimons le modèle suivant :

$$g_T = \alpha S + \beta y_0 + \varepsilon \quad [4.22a]$$

$$\varepsilon \sim N\left(0, \begin{bmatrix} \sigma_{\varepsilon,1}^2 I & 0 \\ 0 & \sigma_{\varepsilon,2}^2 I \end{bmatrix}\right) \quad [4.22b]$$

Les résultats d'estimation par les MCQG (Moindres carrés Quasi Généralisés) sont présentés dans le tableau 4.8. Quel que soit l'échantillon considéré, les coefficients sont tous significatifs. Les coefficients $\hat{\beta}$ sont plus faibles que dans le modèle précédent. Ainsi, la vitesse de convergence varie entre 0,70% et 0,79% et la demi-vie atteint jusqu'à 102 ans. Ces résultats indiquent que le processus de convergence est très faible.

Pour évaluer si les variances associées aux deux régimes sont significativement différentes, nous avons effectué un test de Wald. Les résultats indiquent que la différence ne semble pas être significative ($p = 0,051$ et $p = 0,052$ pour les PIB par tête en Ecus ; $p = 0,214$ pour les PIB par tête en SPA). Pourtant, ces résultats doivent être interprétés avec prudence à cause de la présence d'autocorrélation spatiale, détectée par les statistiques LM_{ERR} et LM_{LAG} . Dans tous les cas, ces deux tests indiquent qu'une autocorrélation spatiale des erreurs a été omise à tort. Prendre en compte l'hétéroscédasticité en groupes ne semble donc pas éliminer l'autocorrélation spatiale et les résultats obtenus sur ce modèle ne sont pas fiables.

Estimation (MCQG)	Ecus (135 régions)		SPA (138 régions)	
	W(1) - ech.1	W(10) - ech.2	W(1) - ech.3	W(10) - ech.3
$\hat{\alpha}$	0,120 (0,000)	0,119 (0,000)	0,126 (0,000)	0,126 (0,000)
$\hat{\beta}$	-0,00677 (0,000)	-0,00669 (0,000)	-0,0075 (0,000)	-0,0075 (0,000)
Vitesse de convergence	0,71% (0,000)	0,70% (0,000)	0,79% (0,000)	0,79% (0,000)
Demi-vie	102	103	92	92
Sq.- Corr.	0,14	0,14	0,11	0,11
$\hat{\sigma}_{\varepsilon,1}^2$	$6,228.10^{-5}$ (0,000)	$6,226.10^{-5}$ (0,000)	$3,734.10^{-5}$ (0,000)	$3,734.10^{-5}$ (0,000)
$\hat{\sigma}_{\varepsilon,2}^2$	$11,141.10^{-5}$ (0,000)	$11,153.10^{-5}$ (0,000)	$5,323.10^{-5}$ (0,000)	$5,323.10^{-5}$ (0,000)
Tests				
test de Wald d'hétéroscédasticité	3,78 (0,052)	3,80 (0,051)	1,541 (0,214)	1,541 (0,214)
LM_{ERR}	129,59 (0,000)	142,86 (0,000)	25,735 (0,000)	28,996 (0,000)
LM_{LAG}	119,20 (0,000)	125,65 (0,000)	22,974 (0,000)	24,860 (0,000)

Tableau 4.8 : Résultats d'estimation pour le modèle de β -convergence avec hétéroscédasticité en groupes

Notes : Les valeurs critiques sont entre parenthèses. MCQG est la méthode des Moindres Carrés Quasi-Généralisés. Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. $\hat{\sigma}_{\varepsilon,1}^2$ est la variance estimée pour le régime Nord et $\hat{\sigma}_{\varepsilon,2}^2$ est la variance estimée pour le régime Sud. Test de Wald d'hétéroscédasticité est le test de Wald pour des variances différentes dans les deux régimes. LM_{ERR} est le test du multiplicateur de Lagrange pour une autocorrélation des erreurs. LM_{LAG} est le test du multiplicateur de Lagrange pour une variable endogène décalée.

Nous considérons maintenant la possibilité d'une instabilité structurelle. Pour cela, nous estimons un modèle β -convergence absolue à 2 régimes, qui peut être spécifié de la façon suivante :

$$g_T = \alpha_1 D_1 + \alpha_2 D_2 + \beta_1 D_1 y_0 + \beta_2 D_2 y_0 + \varepsilon \quad [4.23a]$$

$$\varepsilon \sim N(0, \sigma_\varepsilon^2 I) \quad [4.23b]$$

où D_1 et D_2 sont les variables muettes qualifiant les deux régimes précédemment définis. Plus précisément, $D_{1,i}$ égal à 1 si la région i appartient au Nord et 0 si la région appartient au Sud. $D_{2,i}$ égal à 1 si la région i appartient au Sud et 0 si la région appartient au Nord.

Ce modèle peut aussi être formulé en forme matricielle de la façon suivante :

$$\begin{bmatrix} g_{T,1} \\ g_{T,2} \end{bmatrix} = \begin{bmatrix} S_1 & y_{0,1} & 0 & 0 \\ 0 & 0 & S_2 & y_{0,2} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad [4.24a]$$

$$\Rightarrow Z = X\delta + \varepsilon \quad [4.24b]$$

où $\varepsilon' = [\varepsilon_1' \quad \varepsilon_2']$ et $\varepsilon \sim N(0, \sigma_\varepsilon^2 I)$, l'indice 1 désigne le régime Nord et l'indice 2 désigne le régime Sud.

Ce type de spécification permet de prendre en compte le fait que le processus de convergence, s'il existe, est susceptible d'être différent entre les deux régimes. Ainsi, notre méthodologie peut être interprétée comme une approche en termes de clubs de convergence spatiaux, dans laquelle les clubs sont identifiés en utilisant un critère spatial. Notre approche étend la méthodologie empirique élaborée par Durlauf et Johnson (1995) pour prendre en compte explicitement la dimension spatiale des données.

Le tableau 4.9 présente les résultats de l'estimation du modèle [4.24] par les MCO avec la correction de White. On note que $\hat{\beta}_1$ n'est pas significatif pour le Nord. En revanche, $\hat{\beta}_2$ est significatif et négatif pour les régions du Sud. La vitesse de convergence varie alors entre 1,93% et 2,80% conduisant à une demi-vie égale respectivement à 41 ans et 30 ans. Le processus de convergence pour les régions du Sud semble être plus important que dans le modèle initial. Le test de Chow-White de stabilité globale rejette fortement l'hypothèse nulle jointe d'égalité des coefficients pour les PIB par tête en Ecus. Les tests de stabilité individuelle rejettent l'hypothèse nulle correspondante. Pour les PIB par tête en SPA, le test global de Chow-White n'est pas significatif à 5%, même si les tests individuels le sont. Le processus de convergence semble donc être différent entre les régimes. Plus précisément, deux processus opposés coexistent : à un groupe de régions pauvres qui convergent entre elles à un rythme moyen, s'oppose un groupe de régions riches pour lesquelles l'hypothèse de convergence absolue est rejetée.

Pour les PIB par tête en Ecus, le résultat du test de normalité de Jarque-Bera ne permet pas de rejeter la normalité. Le rejet d'une hétéroscédasticité en groupes est plus faible que dans le modèle initial avec une valeur critique égale à 0,045 pour $W(1)$ et égale à 0,042 pour $W(10)$. Pour les PIB par tête en SPA, la normalité est toujours fortement rejetée. Le test de Koenker-Bassett d'hétéroscédasticité contre D_1 ne permet pas de rejeter l'absence d'une hétéroscédasticité en groupes.

Concernant l'autocorrélation spatiale, les tests nous orientent toujours vers une autocorrélation des erreurs comme dans le modèle précédent. Ainsi, tous les tests précédemment effectués doivent être interprétés avec prudence à cause de la présence potentielle d'autocorrélation spatiale. Nous devons donc estimer un modèle à 2 régimes avec autocorrélation des erreurs pour vérifier l'ensemble des résultats obtenus.

Estimation (MCO – White)	Ecus (135 régions)		SPA (138 régions)	
	W(1) - ech.1	W(10) - ech.2	W(1) - ech.3	W(10) - ech.3
$\hat{\alpha}$ - Nord	-0,000825 (0,981)	-0,0033 (0,938)	0,083 (0,001)	0,083 (0,001)
$\hat{\alpha}$ - Sud	0,252 (0,000)	0,252 (0,000)	0,205 (0,000)	0,205 (0,000)
$\hat{\beta}$ - Nord	0,00663 (0,093)	0,0069 (0,093)	-0,0025 (0,376)	-0,0025 (0,376)
$\hat{\beta}$ - Sud	-0,0228 (0,000)	-0,0228 (0,000)	-0,0167 (0,004)	-0,0167 (0,004)
Vitesse de convergence dans le Sud	2,80% (0,000)	2,80% (0,000)	1,93% (0,000)	1,93% (0,000)
Demi-vie dans le Sud	30	30	41	41
R^2	0,25	0,26	0,13	0,13
LIK	457,81	458,05	502,11	502,11
AIC	-907,62	909,00	-996,21	-996,21
BIC	-896,00	-896,48	-984,50	-984,50
$\hat{\sigma}^2$	6,840.10 ⁻⁵	6,815.10 ⁻⁵	4,168.10 ⁻⁵	4,168.10 ⁻⁵
Tests				
JB	0,395 (0,821)	0,442 (0,802)	8,815 (0,009)	8,815 (0,012)
BP ou KB* vs. D_1	4,015 (0,045)	4,13 (0,042)	1,537* (0,215)	1,537* (0,422)
Test de stabilité individuelle sur α	18,64 (0,000)	23,17 (0,000)	4,030 (0,006)	4,030 (0,006)
Test de stabilité individuelle sur β	18,39 (0,000)	23,78 (0,000)	4,635 (0,031)	4,635 (0,031)
Test Chow-White de stabilité globale	18,86 (0,000)	12,43 (0,000)	5,149 (0,076)	5,149 (0,076)
I de Moran	11,95 (0,000)	13,744 (0,000)	5,603 (0,000)	6,394 (0,000)
LM_{ERR}	109,57 (0,000)	138,324 (0,000)	22,352 (0,000)	27,672 (0,000)
RLM_{ERR}	12,59 (0,000)	38,68 (0,000)	2,857 (0,000)	9,053 (0,002)
LM_{LAG}	97,48 (0,000)	101,44 (0,000)	19,656 (0,000)	20,680 (0,000)
RLM_{LAG}	0,512 (0,474)	101,44 (0,000)	0,161 (0,622)	2,061 (0,151)

Tableau 4.9 : Résultats d'estimation pour le modèle à deux régimes

Notes : les valeurs critiques sont entre parenthèses. MCO-White indique l'utilisation de la matrice des variances-covariances convergente en présence d'hétéroscédasticité dans l'estimation par les MCO (White, 1980). LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). JB est le test de normalité des résidus de Jarque-Bera (1987). BP est le test de Breusch-Pagan (1979) pour l'hétéroscédasticité. KB est le test de Koener-Bassett (1982) pour l'hétéroscédasticité en présence de non-normalité. Les tests de stabilité individuelle sont basés sur des statistiques asymptotiques de Wald utilisant la matrice des variances-covariances de White (1980). Ils convergent asymptotiquement vers un χ^2 à 1 degré de liberté. Le test de Chow-White de stabilité globale est également basé sur la statistique asymptotique de Wald utilisant la matrice des variances-covariances de White (1980). Il converge asymptotiquement vers un χ^2 à 2 degrés de liberté. I de Moran est le test de Moran adapté aux résidus (Cliff et Ord, 1981). LM_{ERR} est le test du multiplicateur de Lagrange pour une autocorrélation des erreurs et RLM_{ERR} est sa version robuste. LM_{LAG} est le test du multiplicateur de Lagrange pour une variable endogène décalée et RLM_{LAG} est sa version robuste (Anselin *et al.*, 1996).

422.4 Autocorrélation et hétérogénéité spatiales

Nous estimons un modèle avec autocorrélation spatiale des erreurs et régimes spatiaux, dans lequel on suppose que le même processus spatial autorégressif affecte toutes les erreurs :

$$g_T = \alpha_1 D_1 + \alpha_2 D_2 + \beta_1 D_1 y_0 + \beta_2 D_2 y_0 + \varepsilon \quad [4.25a]$$

$$\varepsilon = \lambda W \varepsilon + u \quad \text{et} \quad u \sim N(0, \sigma^2 I) \quad [4.25b]$$

avec les mêmes notations que précédemment.

Sous forme matricielle, ce modèle est reformulé de la façon suivante :

$$\begin{bmatrix} g_{T,1} \\ g_{T,2} \end{bmatrix} = \begin{bmatrix} S_1 & y_{0,1} & 0 & 0 \\ 0 & 0 & S_2 & y_{0,2} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix} \quad [4.26a]$$

$$\varepsilon = \lambda W \varepsilon + u \quad \text{et} \quad u \sim N(0, \sigma^2 I) \quad [4.26b]$$

où $\varepsilon' = [\varepsilon_1' \quad \varepsilon_2']$, l'indice 1 désigne le régime Nord et l'indice 2 désigne le régime Sud.

Cette spécification permet de capter, d'une part, l'autocorrélation spatiale des erreurs et d'autre part, un processus de convergence différent entre les régimes. Les effets spatiaux sont supposés être identiques dans les régions du Nord et les régions du Sud et toutes les régions de l'échantillon interagissent spatialement à travers la matrice de poids W . Nous préférons conserver cette procédure car si nous estimions séparément deux régressions afin de capter des effets spatiaux différents dans les régimes, les régions du Sud ne pourraient plus interagir spatialement avec les régions du Nord. Par ailleurs, il faudrait baser les estimations sur des matrices de poids propres à chaque régime, alors que W contient la distance basée sur la configuration géographique globale et reste donc complètement exogène.

Les résultats de l'estimation du modèle [4.26] par la méthode du maximum de vraisemblance et la méthode GMM sont donnés dans le tableau 4.10.

Dans tous les cas, $\hat{\beta}_1$ et $\hat{\beta}_2$ ont les signes attendus mais $\hat{\beta}_1$ n'est jamais significatif pour le Nord. Pour les régions du Sud, $\hat{\beta}_2$ est toujours fortement significatif et négatif. La vitesse de convergence et la demi-vie sont légèrement améliorées, en comparaison du modèle à 2 régimes estimé par les MCO, une fois que les effets spatiaux ont été contrôlés. Le test de Chow spatialement ajusté rejette fortement l'hypothèse nulle jointe d'une instabilité structurelle, sauf pour les PIB par tête en SPA avec $W(1)$: l'hypothèse nulle jointe n'est

rejetée qu'à 10% ($p = 0,088$). Cependant, les tests de stabilité des coefficients individuels conduisent tous au rejet des hypothèses nulles correspondantes.

Ces résultats confirment que le processus de convergence diffère entre les régimes. S'il y a un processus de convergence entre les régions européennes, il concerne principalement les régions du Sud et n'affecte pas les régions du Nord.

Le test de Breusch-Pagan spatial contre la variable de régime D_1 ne permet pas de rejeter l'hypothèse nulle d'absence d'hétéroscédasticité à 5%. Le seul problème subsiste pour la régression estimée avec les PIB par tête en Ecus et $W(10)$. Cependant, cette dernière régression pose de nombreux autres problèmes. Ainsi, le test du facteur commun ne peut être calculé avec la statistique de ratio de vraisemblance et le test LM_{LAG}^* ne permet pas de rejeter l'hypothèse nulle. Les résultats obtenus à l'aide des matrices des 10 plus proches voisins ne sont donc pas fiables et seuls les résultats obtenus à l'aide de la méthode GMM devraient être considérés.

Le second aspect des résultats que nous voulons souligner concerne les effets de débordement géographiques. Tout d'abord, on trouve, dans tous les cas, une forte autocorrélation spatiale positive des erreurs. Le test LM_{LAG}^* ne permet jamais de rejeter l'hypothèse nulle. En outre, les tests du facteur commun indiquent que le modèle [4.26] peut être réécrit sous la forme du modèle de Durbin spatial contraint :

$$g_T = \alpha_1(I - \lambda W)D_1 + \alpha_2(I - \lambda W)D_2 + \beta_1 D_1 y_0 + \beta_2 D_2 y_0 + \lambda W g_T + \gamma_1 W D_1 y_0 + \gamma_2 W D_2 y_0 + u \quad [4.27]$$

où $u \sim N(0, \sigma^2 I)$.

Deux restrictions non-linéaires doivent maintenant être respectées : $\gamma_1 = -\lambda\beta_1$ et $\gamma_2 = -\lambda\beta_2$. Cependant, ces deux coefficients ne sont jamais significatifs. Ainsi, il apparaît que, quel que soit le régime, le taux de croissance annuel moyen d'une région i est positivement influencé par le taux de croissance annuel moyen des régions voisines, à travers la variable endogène décalée Wg_T mais il ne semble pas être influencé par le PIB par tête initial des régions voisines.

Compte tenu des critères d'information, ce modèle est meilleur que les précédents alors que l'estimation par les GMM conduit pratiquement aux mêmes résultats sur les paramètres d'intérêt.

Pour les autres matrices de poids, on aboutit également toujours au modèle à 2 régimes et autocorrélation spatiale des erreurs. Les résultats complets sont fournis dans l'annexe 4.4 pour les PIB par tête en Ecus et dans l'annexe 4.5 pour les PIB par tête en SPA.

	Ecus (135 régions)				SPA (138 régions)			
	W(1) - ech.1		W(10) - ech.2		W(1) - ech.3		W(10) - ech.3	
Estimation	MV	GMM	MV	GMM	MV	GMM	MV	GMM
$\hat{\alpha}$ - Nord	0,0798 (0,014)	0,0837 (0,009)	0,0629 (0,069)	0,0651 (0,063)	0,0964 (0,001)	0,0984 (0,001)	0,0868 (0,007)	0,0998 (0,001)
$\hat{\alpha}$ - Sud	0,263 (0,000)	0,280 (0,000)	0,298 (0,000)	0,316 (0,000)	0,2105 (0,000)	0,2184 (0,000)	0,2286 (0,000)	0,2045 (0,000)
$\hat{\beta}$ - Nord	-0,0026 (0,438)	-0,0030 (0,405)	-0,00078 (0,841)	-0,00085 (0,827)	-0,0042 (0,231)	-0,0044 (0,203)	-0,0031 (0,397)	-0,0045 (0,189)
$\hat{\beta}$ - Sud	-0,0238 (0,000)	-0,0261 (0,000)	-0,0281 (0,000)	-0,0308 (0,000)	-0,0173 (0,000)	-0,0182 (0,001)	-0,0194 (0,000)	-0,0165 (0,001)
Vitesse de convergence dans le Sud	2,94% (0,000)	3,24% (0,000)	3,64% (0,000)	4,14% (0,000)	1,99% (0,000)	2,13% (0,000)	2,29% (0,000)	1,90% (0,000)
Demi-vie dans le Sud	29	26	24	22	40	37	35	42
$\hat{\lambda}$	0,788 (0,000)	0,793	0,811 (0,000)	0,793	0,534 (0,000)	0,575	0,556 (0,000)	0,581
Sq.- Corr.	0,22	0,25	0,26	0,26	0,14	0,14	0,14	0,14
LIK	489,65	-	489,72	-	512,08	-	512,54	-
AIC	-971,31	-	-971,43	-	-1017,71	-	-1017,07	-
BIC	-959,68	-	-959,81	-	-1006,01	-	-1005,36	-
$\hat{\sigma}^2$	3,719.10 ⁻⁵	-	3,778.10 ⁻⁵	-	3,468.10 ⁻⁵	-	3,485.10 ⁻⁵	-
Tests								
BP-S vs. D_1	3,396 (0,065)	-	7,14 (0,007)	-	2,765 (0,096)	-	3,455 (0,063)	-
Test de stabilité individuelle sur α	12,88 (0,000)	-	24,05 (0,000)	-	4,711 (0,030)	-	7,519 (0,006)	-
Test de stabilité individuelle sur β	12,57 (0,000)	-	24,15 (0,000)	-	4,618 (0,032)	-	7,463 (0,006)	-
Test Chow-Wald de stabilité globale	13,06 (0,001)	-	24,15 (0,000)	-	4,855 (0,088)	-	7,533 (0,023)	-
LM_{LAG}^*	0,032 (0,857)	-	41,20 (0,000)	-	0,521 (0,470)	-	0,123 (0,725)	-
LR ou Wald*-facteur commun	5,38 (0,068)	-	2,81* (0,245)	-	0,421 (0,810)	-	0,383* (0,825)	-
$\hat{\gamma}_1 = -\hat{\lambda}\hat{\beta}_1$	0,002 (0,970)	-	0,001 (0,991)	-	0,002 (0,971)	-	0,002 (0,979)	-
$\hat{\gamma}_2 = -\hat{\lambda}\hat{\beta}_2$	0,0187 (0,729)	-	0,023 (0,685)	-	0,009 (0,882)	-	0,011 (0,869)	-

Tableau 4.10 : Résultats d'estimation pour le modèle de β -convergence avec autocorrélation des erreurs et 2 régimes

Notes : Les valeurs critiques sont entre parenthèses. MV est la méthode du maximum de vraisemblance. GMM est la méthode des moments généralisés. Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). BP-S est la version ajustée spatialement du test de Breusch-Pagan (1979) pour l'hétéroscédasticité. Les tests de stabilité individuelle sont basés sur des statistiques asymptotiques de Wald utilisant la matrice des variances-covariances de White (1980). Ils convergent asymptotiquement vers un χ^2 à 1 degré de liberté. Le test de Chow-White de stabilité globale est également basé sur la statistique asymptotique de Wald utilisant la matrice des variances-covariances de White (1980). Il converge asymptotiquement vers un χ^2 à 2 degrés de liberté. LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). LR-facteur commun est le test du ratio de vraisemblance du facteur commun et Wald-facteur commun est le test de Wald du facteur commun (Burrige, 1981).

422.5 Effets spatiaux différenciés

Dans le modèle précédent, nous avons supposé que les effets spatiaux étaient identiques dans les deux régimes. Cette hypothèse est posée *a priori* alors qu'elle devrait être testée. En outre, effectuer deux régressions séparées, afin de capter des effets spatiaux différents, n'est pas satisfaisant puisque cela signifie que ces régions ne peuvent bénéficier d'aucune interaction avec les régions du Nord et vice versa. Par conséquent, il est plus intéressant d'étudier des effets spatiaux différenciés dans la modélisation des clubs de convergence. Ceci est effectué en étendant la méthodologie suggérée par Rietveld et Wintershoven (1998) dans un contexte différent. Un coefficient λ différent pour chaque régime est estimé et un coefficient d'interaction Nord-Sud est ajouté dans le modèle [4.25]. Formellement, nous considérons la spécification suivante :

$$g_T = \alpha_1 D_1 + \alpha_2 D_2 + \beta_1 D_1 y_0 + \beta_2 D_2 y_0 + \varepsilon \quad [4.28a]$$

$$\varepsilon = (\lambda_1 W_1 + \lambda_2 W_2 + \lambda_3 W_3) \varepsilon + u \text{ et } u \sim N(0, \sigma^2 I) \quad [4.28b]$$

Ce modèle permet de prendre en compte l'instabilité structurelle et des effets spatiaux différenciés à l'intérieur et entre les régimes.

Pour aboutir à cette spécification, la matrice de poids W est divisée en trois parties : W_1 inclut uniquement les interconnexions spatiales entre les régions appartenant au régime du Nord, W_2 inclut uniquement les interconnexions spatiales entre les régions appartenant au régime du Sud et W_3 inclut uniquement les interconnexions spatiales entre les régions appartenant au régime Nord et les régions appartenant au régime Sud.

Ces matrices peuvent être construites en utilisant deux approches différentes. La première est basée sur le découpage en trois matrices de la matrice W standardisée, ce qui conduit à des matrices non-standardisées W_j , $j=1,2,3$. L'avantage principal de cette première approche est que le test d'homogénéité des effets spatiaux peut être effectué sans problème puisque le modèle [4.25] est le modèle contraint sous l'hypothèse nulle de coefficient λ_j égaux dans le modèle [4.28]. L'inconvénient est l'utilisation de matrices non-standardisées dans l'estimation par la méthode du maximum de vraisemblance du modèle [4.28]. Les conditions de régularité usuelles peuvent alors ne pas être respectées. En outre, l'interprétation des coefficients λ_j comme des coefficients d'autocorrélation spatiale devient ambiguë. La seconde approche est basée sur le découpage en 3 matrices de la matrice W non standardisée, les matrices W_j étant ensuite standardisées. L'inconvénient principal de cette seconde approche est que le modèle [4.25] ne peut plus être considéré comme le modèle contraint pour le test d'homogénéité.

A la suite de Rietveld et Wintershoven (1998), nous utilisons la première approche en estimant le modèle [4.28] par la méthode du maximum de vraisemblance. Les codes

nécessaires pour ces estimations ont été rédigés sous GAUSS 3.2.26. L'annexe 4.6 présente en détail la procédure d'estimation adoptée pour rédiger ces codes.

Les résultats sont présentés dans le tableau 4.11. Concernant le paramètre β et les régimes de convergence, ils sont similaires à ceux obtenus précédemment. Pour les PIB par tête en Ecus, il apparaît que le paramètre $\hat{\lambda}_1$ pour les régions du Nord et le paramètre $\hat{\lambda}_2$ pour les régions du Sud sont significatifs et positifs alors que le paramètre $\hat{\lambda}_3$, représentant les interactions Nord-Sud, n'est pas significatif. Ce résultat surprenant peut s'expliquer par le fait que la matrice W_3 contient trop d'éléments nuls. Pour les PIB par tête en SPA, seul le coefficient $\hat{\lambda}_1$ est significatif à 5%.

Nous avons effectué le test du ratio de vraisemblance pour l'homogénéité des effets spatiaux sous l'hypothèse maintenue de deux régimes différents. Le modèle non-contraint est alors le modèle [4.28] alors que le modèle contraint est le modèle [4.25]. Il apparaît que l'hypothèse nulle d'égalité des effets spatiaux n'est jamais rejetée.

Estimation (MV)	Ecus (135 régions)		SPA (138 régions)	
	W(1) - ech.1	W(10) - ech.2	W(1) - ech.3	W(10) - ech.3
$\hat{\alpha}$ - Nord	0,085 (0,007)	0,067 (0,025)	0,086 (0,003)	0,085 (0,005)
$\hat{\alpha}$ - Sud	0,259 (0,000)	0,293 (0,000)	0,205 (0,000)	0,232 (0,000)
$\hat{\beta}$ - Nord	-0,0032 (0,350)	-0,0011 (0,388)	-0,0030 (0,190)	-0,0031 (0,197)
$\hat{\beta}$ - Sud	-0,0234 (0,000)	-0,0272 (0,000)	-0,0164 (0,002)	-0,0198 (0,000)
Vitesse de convergence dans le Sud	2,89% (0,000)	3,50% (0,000)	1,89% (0,000)	2,34% (0,000)
Demi-vie dans le Sud	29	25	42	35
$\hat{\lambda}_1$	0,871 (0,000)	0,822 (0,000)	0,741 (0,000)	0,818 (0,000)
$\hat{\lambda}_2$	0,704 (0,000)	0,816 (0,000)	0,566 (0,077)	0,590 (0,590)
$\hat{\lambda}_3$	-0,091 (0,924)	-0,152 (0,428)	-0,874 (0,081)	-0,877 (0,107)
LIK	489,89	489,88	510,25	514,452
AIC	-971,78	-971,76	-1012,49	-1020,90
BIC	-960,16	-960,14	-1000,78	-1009,20
$\hat{\sigma}^2$	3.653.10 ⁻⁵	3,769.10 ⁻⁵	3,493.10 ⁻⁵	3,434.10 ⁻⁵
Tests				
LR – régimes	11,84 (0,003)	20,861 (0,000)	8,800 (0,012)	11,518 (0,003)
LR – effets spatiaux	0,464 (0,793)	0,160 (0,923)	0,340 (0,844)	3,834 (0,147)

Tableau 4.11 : Résultats d'estimation avec régimes et effets spatiaux différenciés

Notes : Les valeurs critiques sont entre parenthèses. MV est la méthode du maximum de vraisemblance. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). LR - régimes est le test du ratio de vraisemblance pour les régimes. LR – effets spatiaux est le test du ratio de vraisemblance pour les effets spatiaux différenciés.

Nous avons également effectué le test du ratio de vraisemblance d'existence des régimes sous l'hypothèse maintenue des effets spatiaux différenciés. Le modèle non-contraint est toujours le modèle [4.28] alors que le modèle contraint s'écrit dans ce cas :

$$g_T = \alpha S + \beta y_0 + \varepsilon \quad [4.29a]$$

$$\varepsilon = (\lambda_1 W_1 + \lambda_2 W_2 + \lambda_3 W_3) \varepsilon + u \text{ et } u \sim N(0, \sigma^2 I) \quad [4.29b]$$

Les résultats de l'estimation du modèle [4.29] par le maximum de vraisemblance sont fournis dans le tableau 4.12. Ils permettent d'effectuer le test du ratio de vraisemblance pour les régimes. Il apparaît que l'hypothèse nulle d'absence de régimes est toujours fortement rejetée. Ces résultats confirment en fait que le modèle [4.25] à 2 régimes spatiaux mais effets spatiaux non-différenciés est la spécification la plus appropriée.

Estimation (MV)	Ecus (135 régions)		SPA (138 régions)	
	W(1) - ech.1	W(10) - ech.2	W(1) - ech.3	W(10) - ech.3
$\hat{\alpha}$	0,159 (0,000)	0,183 (0,000)	0,130 (0,000)	0,143 (0,000)
$\hat{\beta}$	-0,0114 (0,000)	-0,0141 (0,000)	-0,0078 (0,005)	-0,0094 (0,000)
Vitesse de convergence	1,25% (0,000)	1,58% (0,000)	0,83% (0,000)	1,01% (0,000)
Demi-vie	61	48	88	73
$\hat{\lambda}_1$	0,871 (0,000)	0,843 (0,000)	0,738 (0,000)	0,743 (0,000)
$\hat{\lambda}_2$	0,714 (0,000)	0,752 (0,000)	0,482 (0,007)	0,433 (0,095)
$\hat{\lambda}_3$	-0,488 (0,595)	-0,491 (0,594)	-0,745 (0,224)	-0,765 (0,220)
LIK	483,97	479,45	505,85	508,693
AIC	-965,94	-954,89	-1007,70	-1013,39
BIC	-958,13	-949,08	-1001,85	-1007,53
$\hat{\sigma}^2$	4,007.10 ⁻⁵	4,424.10 ⁻⁵	3,591.10 ⁻⁵	3,621.10 ⁻⁵

Tableau 4.12 : Résultats d'estimation avec effets spatiaux différenciés

Notes : Les valeurs critiques sont entre parenthèses. MV est la méthode du maximum de vraisemblance. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978).

Conclusion de section

Dans cette section, nous avons pris en compte l'hétérogénéité spatiale dans le processus de convergence sous la forme de 2 régimes spatiaux, l'un pour les régions du Nord et l'autre pour les régions du Sud. Les résultats indiquent que le processus de convergence global masque en fait deux processus opposés : les régions pauvres convergent à un rythme lent alors que l'hypothèse de β -convergence absolue est rejetée pour les régions riches. Par ailleurs, dans toutes les configurations testées, un effet de débordement géographique significatif apparaît sur les taux de croissance annuels moyens des régions.

Conclusion du chapitre 4

Dans ce chapitre, nous avons montré que l'autocorrélation et l'hétérogénéité spatiales comptent réellement dans l'estimation du processus de β -convergence. Face aux limites théoriques et aux problèmes économétriques rencontrés par le modèle de β -convergence standard, nous l'améliorons de deux façons différentes.

Premièrement, sur le plan économétrique, l'absence de fiabilité de l'inférence statistique basée sur l'estimation par les MCO en présence d'erreurs non-sphériques est bien connue. En utilisant les outils économétriques appropriés, nous avons détecté de l'autocorrélation spatiale et l'avons résolue en estimant un modèle avec autocorrélation spatiale des erreurs. *Ce dernier peut être interprété comme un modèle de β -convergence conditionnel minimal intégrant des variables d'environnement spatial.* Concernant l'hétérogénéité, il apparaît que le problème était essentiellement dû à l'instabilité structurelle sous la forme de régimes spatiaux. Ces derniers, interprétés comme des clubs de convergence spatiaux, ont été définis à l'aide de l'analyse exploratoire des données spatiales, plus précisément, à l'aide du diagramme de Moran et des statistiques de Getis-Ord. Nous avons donc pris en compte simultanément l'autocorrélation spatiale et l'instabilité structurelle. *L'estimation du modèle approprié à 2 régimes et autocorrélation spatiale indique que le processus de convergence est différent selon les régimes.* Il apparaît ainsi qu'un tel processus n'existe pas pour les régions du Nord, et que la convergence est faible pour les régions du Sud. Ce résultat est cohérent avec celui obtenu par Beine et Jean-Pierre (2000) sur des données au niveau NUTS 2 et celui obtenu par Durlauf et Johnson (1995) sur données internationales. Il peut être dû à une hétérogénéité résiduelle intra-régimes qui n'a pas été prise en compte. L'inclusion de variables additionnelles dans un cadre de β -convergence conditionnelle pourrait mener à un résultat de convergence pour le régime du Nord. *La convergence globale trouvée dans l'estimation du modèle de β -convergence standard apparaît donc être un artéfact.*

Deuxièmement, du point de vue économique, l'utilisation des outils de l'économétrie spatiale nous a permis d'estimer un effet de débordement géographique dans le cadre du modèle de β -convergence standard. Cet effet apparaît être très significatif dans toutes les configurations et indique que le taux de croissance annuel moyen d'une région donnée est positivement influencé par le taux de croissance annuel moyen de ses régions voisines. L'environnement géographique joue donc un rôle important dans l'étude des processus de croissance et de convergence.

Nous affinons maintenant ce rôle dans le cadre d'un autre concept de convergence où l'évolution, tout au long de la période, des disparités économiques est explicitement analysée et mise en relation avec l'évolution de l'environnement économique des économies régionales.

Annexe 4.1

Estimation du modèle avec autocorrélation spatiale des erreurs avec $W(2)$, $W(3)$, $W(15)$ et $W(20)$; PIB par tête en Ecus

Estimation	W(2)		W(3)		W(15)		W(20)	
	MV	GMM	MV	GMM*	MV	GMM	MV	GMM*
$\hat{\alpha}$	0,156 (0,000)	0,153 (0,000)	0,153 (0,000)	0,184 (0,000)	0,199 (0,000)	0,200 (0,000)	0,192 (0,000)	-0,016 (0,000)
$\hat{\beta}$	-0,0110 (0,000)	-0,0108 (0,000)	-0,0110 (0,000)	-0,0104 (0,000)	-0,0160 (0,000)	-0,0163 (0,000)	-0,0159 (0,000)	-0,0165 (0,000)
Vitesse de convergence	1,20% (0,000)	1,19% (0,000)	1,18% (0,000)	1,14% (0,000)	1,83% (0,000)	1,86% (0,000)	1,81% (0,000)	1,90% (0,000)
Demi-vie	63	64	63	66	42	42	43	41
$\hat{\lambda}$	0,838 (0,000)	0,933	0,877 (0,000)	1	0,869 (0,000)	0,894	0,921 (0,000)	0,997
Sq.- Corr.	0,13	0,13	0,13	0,13	0,13	0,13	0,13*	0,13
LIK	491,22	-	490,63	-	492,38	-	491,22	-
AIC	-978,44	-	-977,26	-	-980,76	-	-978,44	-
BIC	-972,59	-	-971,41	-	-974,90	-	-972,59	-
$\hat{\sigma}^2$	$4,309 \cdot 10^{-5}$	-	$4,341 \cdot 10^{-5}$	-	$4,660 \cdot 10^{-5}$	-	$4,740 \cdot 10^{-5}$	-
Tests								
BP-S	3,441 (0,063)	-	3,861 (0,049)	-	3,151 (0,076)	-	3,285 (0,070)	-
LM_{LAG}^*	2,436 (0,118)	-	2,845 (0,092)	-	0,146 (0,703)	-	0,173 (0,678)	-
LR ou Wald*- facteur commun	0,710 (0,399)	-	0,967 (0,325)	-	0,609* (0,435)	-	0,774* (0,379)	-
$\hat{\gamma} = -\hat{\lambda}\hat{\beta}$	0,009 (0,865)	-	0,010 (0,851)	-	0,014 (0,773)	-	0,015 (0,691)	-

Tableau 1 : Résultats d'estimation du modèle avec autocorrélation des erreurs avec diverses matrices de poids ; PIB par tête en Ecus

Notes : Les valeurs critiques sont entre parenthèses. MV est la méthode du maximum de vraisemblance, GMM est la méthode GMM itérée, GMM* est la méthode GMM en deux étapes. Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). BP-S est la version ajustée spatialement du test de Breusch-Pagan (1979) pour l'hétéroscédasticité. LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). LR-facteur commun est le test du ratio de vraisemblance du facteur commun et Wald-facteur commun est le test de Wald du facteur commun (Burrige, 1981).

Annexe 4.2

**Estimation du modèle avec autocorrélation spatiale des
erreurs avec $W(2)$, $W(3)$, $W(15)$ et $W(20)$;
PIB par tête en SPA**

Estimation	W(2)		W(3)		W(15)		W(20)	
	MV	GMM	MV	GMM	MV	GMM	MV	GMM
$\hat{\alpha}$	0,141 (0,000)	0,141 (0,000)	0,141 (0,000)	0,141 (0,000)	0,149 (0,000)	0,151 (0,000)	0,151 (0,000)	0,154 (0,000)
$\hat{\beta}$	-0,0090 (0,000)	-0,0090 (0,000)	-0,0090 (0,000)	-0,0090 (0,000)	-0,0100 (0,000)	-0,0102 (0,000)	-0,0102 (0,000)	-0,0106 (0,000)
Vitesse de convergence	0,98% (0,000)	0,98% (0,000)	0,98% (0,000)	0,99% (0,000)	1,09% (0,000)	1,11% (0,000)	1,11% (0,000)	1,16% (0,000)
Demi-vie	76	76	76	75	68	67	67	65
$\hat{\lambda}$	0,553 (0,000)	0,634	0,575 (0,000)	0,659	0,561 (0,000)	0,606	0,652 (0,000)	0,758
Sq.- Corr.	0,11	0,11	0,11	0,11	0,11	0,11	0,11	0,13
LIK	506,29	-	506,11	-	506,62	-	507,39	-
AIC	-1008,59	-	-1008,23	-	-1009,24	-	-1010,78	-
BIC	-1002,74	-	-1002,37	-	-1003,38	-	-1004,92	-
$\hat{\sigma}^2$	$3,687 \cdot 10^{-5}$	-	$3,698 \cdot 10^{-5}$	-	$3,803 \cdot 10^{-5}$	-	$3,759 \cdot 10^{-5}$	-
Tests								
BP-S	0,489 (0,484)	-	0,511 (0,474)	-	0,667 (0,414)	-	0,540 (0,462)	-
LM_{LAG}^*	2,789 (0,095)	-	2,512 (0,113)	-	0,0006 (0,980)	-	0,446 (0,504)	-
LR- ou Wald*- facteur commun	0,822 (0,365)	-	0,779 (0,377)	-	0,085* (0,770)	-	0,125* (0,723)	-
$\hat{\gamma} = -\hat{\lambda}\hat{\beta}$	0,005 (0,945)	-	0,005 (0,947)	-	0,006 (0,943)	-	0,007 (0,939)	-

Tableau 1 : Résultats d'estimation du modèle avec autocorrélation des erreurs avec diverses matrices de poids ; PIB par tête en SPA

Notes : Les valeurs critiques sont entre parenthèses. MV est la méthode du maximum de vraisemblance, GMM est la méthode GMM itérée, GMM* est la méthode GMM en deux étapes. Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). BP-S est la version ajustée spatialement du test de Breusch-Pagan (1979) pour l'hétéroscédasticité. LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). LR-facteur commun est le test du ratio de vraisemblance du facteur commun et Wald-facteur commun est le test de Wald du facteur commun (Burrige, 1981).

Annexe 4.3

Boîtes à moustaches pour les impacts d'émission et de réception

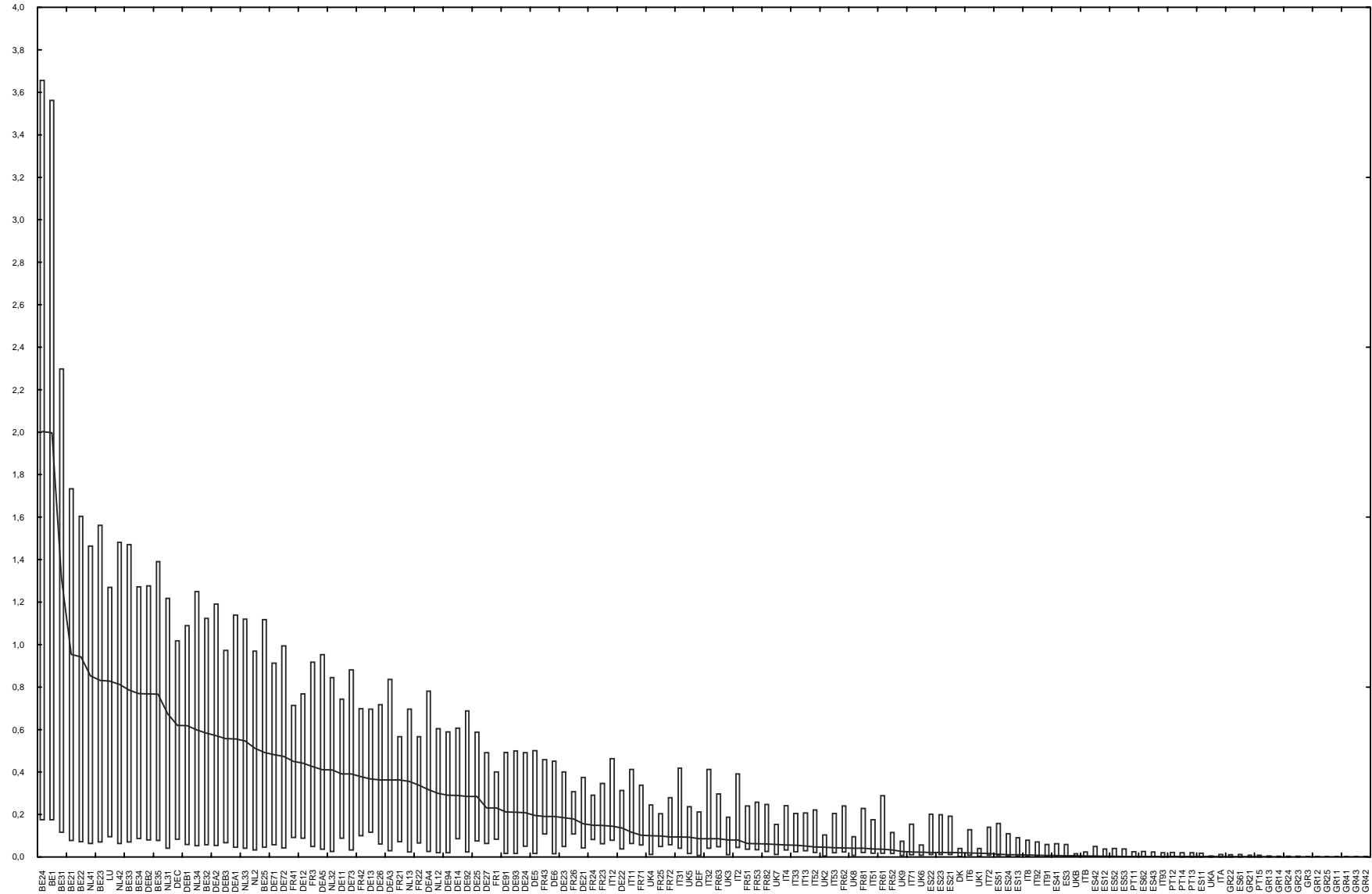


Figure 1 : Boîte à moustaches pour l'impact de l'émission avec $W(1)$ et PIB par tête en Ecus (Q_1 - médiane - Q_3)

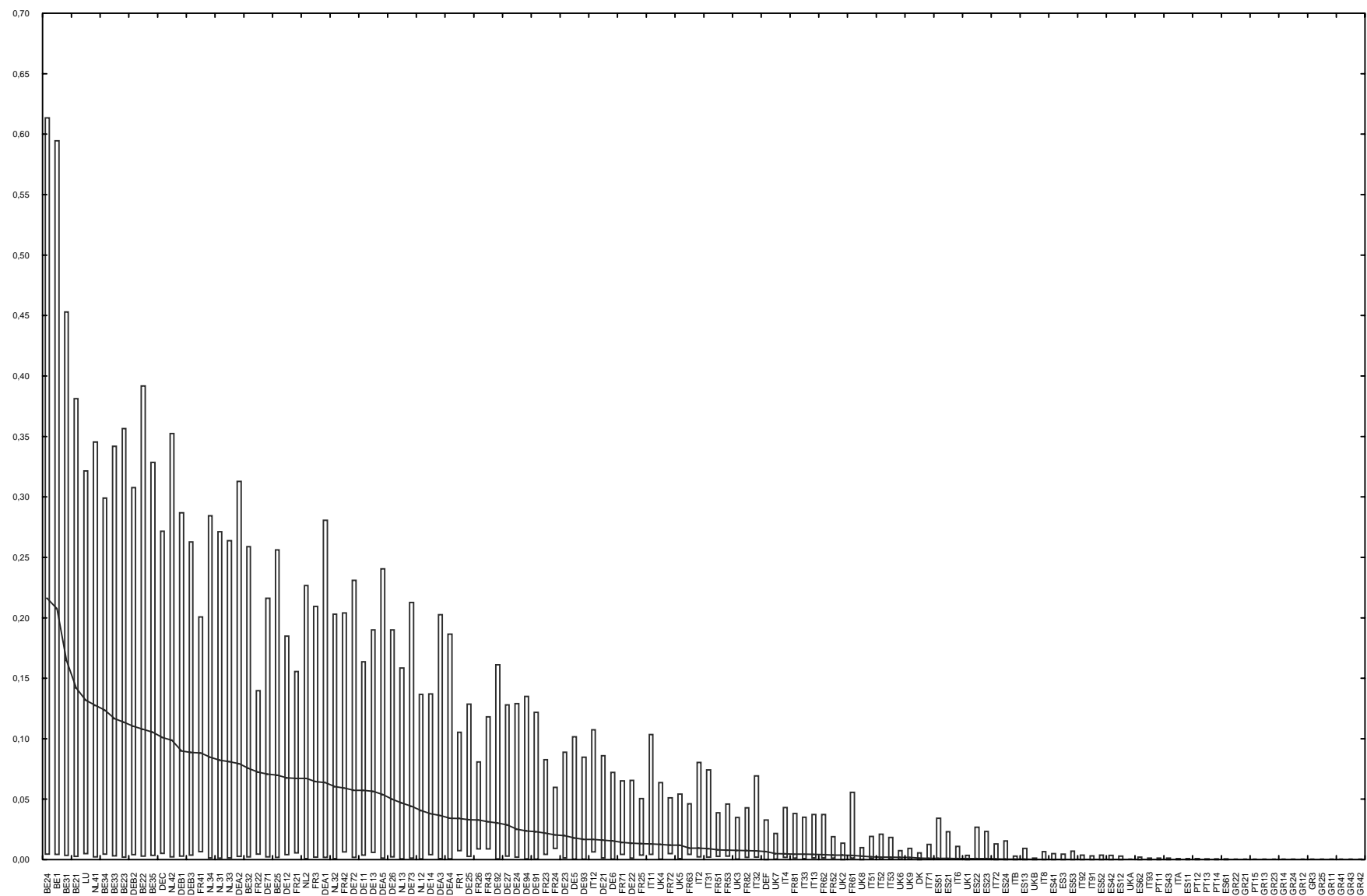


Figure 2 : Boîte à moustaches pour l'impact de l'émission avec $W(1)$ et PIB par tête en SPA (Q_1 - médiane - Q_3)

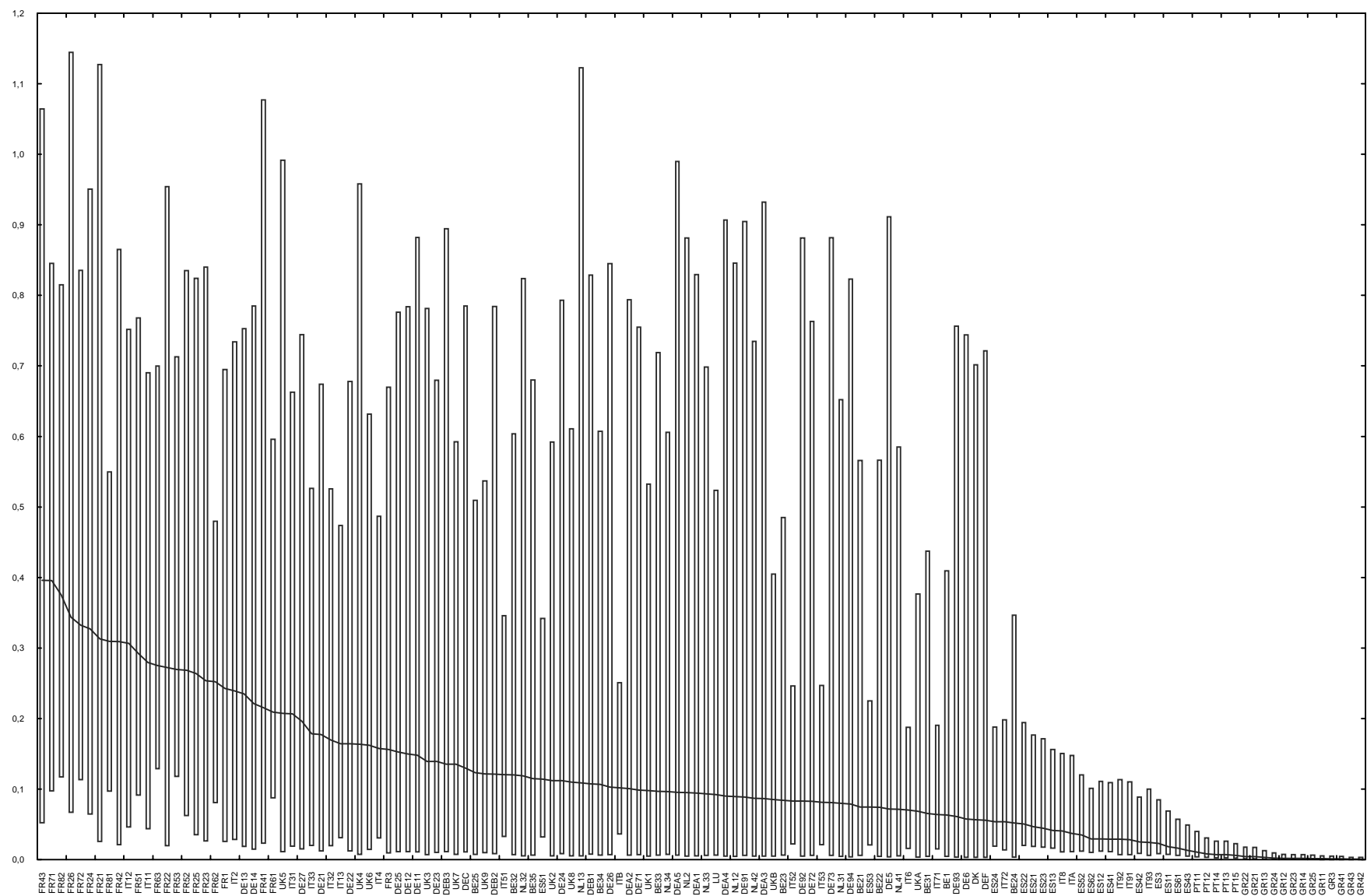


Figure 3 : Boîte à moustaches pour l'impact de la réception avec $W(1)$ et PIB par tête en Ecus (Q_1 - médiane - Q_3)

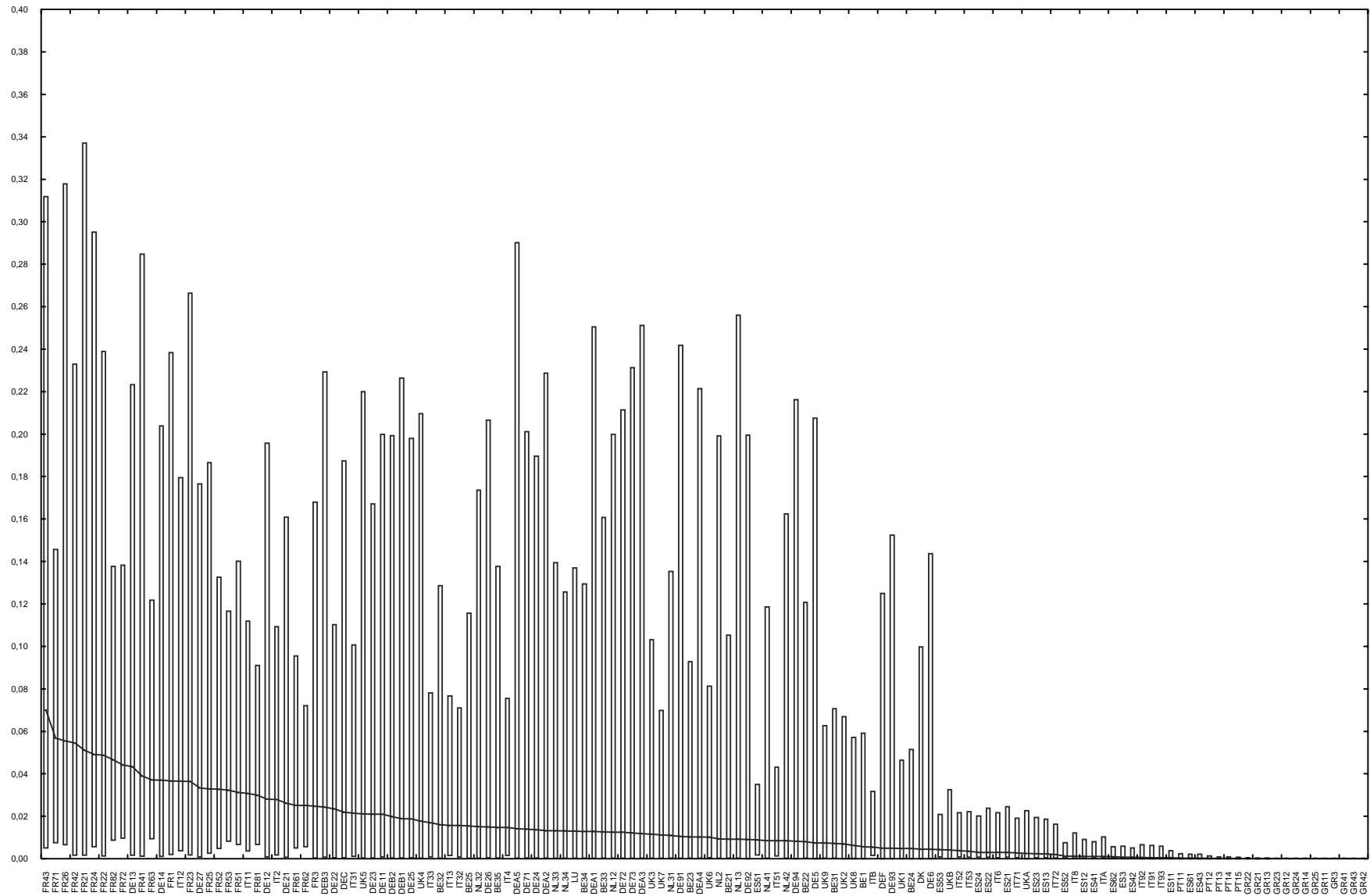


Figure 4 : Boîte à moustaches pour l'impact de la réception avec $W(1)$ et PIB par tête en SPA (Q_1 - médiane - Q_3)

Annexe 4.4

Estimation du modèle avec autocorrélation spatiale des erreurs et régimes spatiaux avec $W(2)$, $W(3)$, $W(15)$ et $W(20)$; PIB par tête en Ecus

	$W(2)$ - ech.1		$W(3)$ - ech.1		$W(15)$ - ech.2		$W(20)$ - ech.2	
	MV	GMM	MV	GMM	MV	GMM	MV	GMM
$\hat{\alpha}$ - Nord	0,072 (0,024)	0,077 (0,018)	0,070 (0,028)	0,076 (0,027)	0,062 (0,071)	0,062 (0,088)	0,054 (0,122)	0,060 (0,080)
$\hat{\alpha}$ - Sud	0,266 (0,000)	0,277 (0,000)	0,264 (0,000)	0,266 (0,000)	0,318 (0,000)	0,326 (0,000)	0,285 (0,000)	0,312 (0,000)
$\hat{\beta}$ - Nord	-0,0019 (0,590)	-0,0026 (0,473)	-0,0018 (0,614)	-0,0024 (0,529)	-0,0010 (0,787)	-0,0007 (0,864)	-0,0011 (0,766)	-0,0007 (0,858)
$\hat{\beta}$ - Sud	-0,0241 (0,000)	-0,0261 (0,000)	-0,0241 (0,000)	-0,0254 (0,000)	-0,0304 (0,000)	-0,0326 (0,000)	-0,0276 (0,000)	-0,0297 (0,000)
Vitesse de convergence dans le Sud	2,99% (0,000)	2,99% (0,000)	2,99% (0,000)	3,19% (0,000)	4,05% (0,000)	4,481% (0,000)	3,56% (0,000)	3,933% (0,000)
Demi-vie dans le Sud	28	28	28	27	22	21	25	23
$\hat{\lambda}$	0,833 (0,000)	0,866	0,861 (0,000)	0,842	0,898 (0,000)	0,859	0,944 (0,000)	0,859
Sq.- Corr.	0,21	0,24	0,22	0,24	0,22	0,22	0,22	0,23
LIK	486,21	-	485,94	-	493,50	-	491,38	-
AIC	-964,43	-	-963,88	-	-978,99	-	-974,77	-
BIC	-952,82	-	-952,26	-	-967,37	-	-963,15	-
$\hat{\sigma}^2$	$3,960 \cdot 10^{-5}$	-	$3,988 \cdot 10^{-5}$	-	$3,912 \cdot 10^{-5}$	-	$4,036 \cdot 10^{-5}$	-
Tests								
BP-S vs. D_1	2,819 (0,093)	-	2,614 (0,105)	-	3,957 (0,047)	-	4,020 (0,045)	-
Test de stabilité individuelle sur α	15,383 (0,000)	-	14,342 (0,000)	-	31,929 (0,000)	-	27,044 (0,000)	-
Test de stabilité individuelle sur β	14,998 (0,000)	-	14,043 (0,000)	-	33,847 (0,000)	-	31,556 (0,000)	-
Test Chow-Wald de stabilité globale	15,400 (0,000)	-	14,343 (0,000)	-	33,943 (0,000)	-	31,631 (0,000)	-
LM_{LAG}^*	0,009 (0,925)	-	0,098 (0,754)	-	2,037 (0,153)	-	0,584 (0,444)	-
LR ou Wald*-facteur commun	6,743 (0,031)	-	7,776 (0,020)	-	5,346* (0,021)	-	6,591* (0,037)	-
$\hat{\gamma}_1 = -\hat{\lambda}\hat{\beta}_1$	0,002 (0,977)	-	0,001 (0,978)	-	0,001 (0,981)	-	0,001 (0,968)	-
$\hat{\gamma}_2 = -\hat{\lambda}\hat{\beta}_2$	0,020 (0,717)	-	0,021 (0,709)	-	0,027 (0,502)	-	0,0276 (0,339)	-

Tableau 1 : Résultats d'estimation du modèle avec autocorrélation des erreurs et régimes spatiaux avec diverses matrices de poids ; PIB par tête en Ecus

Notes : Les valeurs critiques sont entre parenthèses. MV est la méthode du maximum de vraisemblance. GMM est la méthode des moments généralisés. Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). BP-S est la version ajustée spatialement du test de Breusch-Pagan (1979) pour l'hétéroscédasticité. Les tests de stabilité individuelle sont basés sur des statistiques asymptotiques de Wald utilisant la matrice des variances-covariances de White (1980). Ils convergent asymptotiquement vers un χ^2 à 1 degré de liberté. Le test de Chow-White de stabilité globale est également basé sur la statistique asymptotique de Wald utilisant la matrice des variances-covariances de White (1980). Il converge asymptotiquement vers un χ^2 à 2 degrés de liberté. LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). LR-facteur commun est le test du ratio de vraisemblance du facteur commun et Wald-facteur commun est le test de Wald du facteur commun (Burrige, 1981).

Annexe 4.5

Estimation du modèle avec autocorrélation spatiale des erreurs et régimes spatiaux avec $W(2)$, $W(3)$, $W(15)$ et $W(20)$; PIB par tête en SPA

Estimation	$W(2)$ - ech.3		$W(3)$ - ech.3		$W(15)$ - ech.3		$W(20)$ - ech.3	
	MV	GMM	MV	GMM	MV	GMM	MV	GMM
$\hat{\alpha}$ - Nord	0,093 (0,025)	0,096 (0,001)	0,093 (0,025)	0,097 (0,001)	0,092 (0,005)	0,096 (0,003)	0,095 (0,003)	0,101 (0,002)
$\hat{\alpha}$ - Sud	0,204 (0,000)	0,213 (0,000)	0,208 (0,000)	0,213 (0,000)	0,236 (0,000)	0,248 (0,000)	0,222 (0,000)	0,230 (0,000)
$\hat{\beta}$ - Nord	-0,0039 (0,266)	-0,0042 (0,223)	-0,0038 (0,268)	-0,0043 (0,211)	-0,0037 (0,314)	-0,0042 (0,223)	-0,0041 (0,260)	-0,0048 (0,194)
$\hat{\beta}$ - Sud	-0,0164 (0,001)	-0,0176 (0,002)	-0,0170 (0,001)	-0,0178 (0,002)	-0,0202 (0,001)	-0,0219 (0,002)	-0,0186 (0,000)	-0,0199 (0,000)
Vitesse de convergence dans le Sud	1,88% (0,000)	2,03% (0,000)	1,96% (0,000)	2,06% (0,000)	2,41% (0,000)	2,65% (0,000)	2,18% (0,000)	2,36% (0,000)
Demi-vie dans le Sud	42	39	40	39	34	31	37	35
$\hat{\lambda}$	0,565 (0,000)	0,616 (0,000)	0,588 (0,000)	0,637 (0,000)	0,609 (0,000)	0,639 (0,000)	0,685 (0,000)	0,761 (0,000)
Sq.- Corr.	0,14	0,14	0,14	0,14	0,14	0,15	0,14	0,15
LIK	508,68	-	508,67	-	510,52	-	510,63	-
AIC	-1009,37	-	-1009,34	-	-1013,05	-	-1013,27	-
BIC	-997,66	-	-997,63	-	-1001,34	-	-1001,56	-
$\hat{\sigma}^2$	$3,555 \cdot 10^{-5}$	-	$3,558 \cdot 10^{-5}$	-	$3,591 \cdot 10^{-5}$	-	$3,584 \cdot 10^{-5}$	-
Tests								
BP-S vs. D_1	2,588 (0,108)	-	2,587 (0,107)	-	3,491 (0,062)	-	3,394 (0,065)	-
Test de stabilité individuelle sur α	4,418 (0,035)	-	4,540 (0,033)	-	7,554 (0,023)	-	5,972 (0,014)	-
Test de stabilité individuelle sur β	4,273 (0,039)	-	4,390 (0,036)	-	7,554 (0,006)	-	5,914 (0,015)	-
Test Chow-Wald de stabilité globale	4,796 (0,090)	-	4,873 (0,087)	-	7,534 (0,006)	-	5,979 (0,05)	-
LM_{LAG}^*	0,384 (0,535)	-	0,311 (0,577)	-	0,948 (0,330)	-	0,125 (0,723)	-
LR ou Wald*-facteur commun	1,189 (0,552)	-	1,879 (0,391)	-	1,529* (0,465)	-	2,967* (0,227)	-
$\hat{\gamma}_1 = -\hat{\lambda}\hat{\beta}_1$	0,002 (0,976)	-	0,002 (0,977)	-	0,002 (0,977)	-	0,003 (0,974)	-
$\hat{\gamma}_2 = -\hat{\lambda}\hat{\beta}_2$	0,009 (0,899)	-	0,009 (0,900)	-	0,012 (0,876)	-	0,012 (0,882)	-

Tableau 1 : Résultats d'estimation du modèle avec autocorrélation des erreurs et régimes spatiaux avec diverses matrices de poids ; PIB par tête en SPA

Notes : Les valeurs critiques sont entre parenthèses. MV est la méthode du maximum de vraisemblance. GMM est la méthode des moments généralisés. Sq.- Corr. est la corrélation élevée au carré entre les valeurs observées et les valeurs prévues. LIK est la valeur de la fonction de vraisemblance. AIC est le critère d'information d'Akaike (1974). BIC est le critère d'information de Schwarz (1978). BP-S est la version ajustée spatialement du test de Breusch-Pagan (1979) pour l'hétéroscédasticité. Les tests de stabilité individuelle sont basés sur des statistiques asymptotiques de Wald utilisant la matrice des variances-covariances de White (1980). Ils convergent asymptotiquement vers un χ^2 à 1 degré de liberté. Le test de Chow-White de stabilité globale est également basé sur la statistique asymptotique de Wald utilisant la matrice des variances-covariances de White (1980). Il converge asymptotiquement vers un χ^2 à 2 degrés de liberté. LM_{LAG}^* est le test du multiplicateur de Lagrange pour une variable endogène décalée dans le modèle avec autocorrélation spatiale des erreurs (Anselin, 1988a). LR-facteur commun est le test du ratio de vraisemblance du facteur commun et Wald-facteur commun est le test de Wald du facteur commun (Burrige, 1981).

Annexe 4.6

Estimation du modèle avec effets spatiaux différenciés par le maximum de vraisemblance

Le but de cette section est de détailler la procédure d'estimation adoptée pour rédiger les codes GAUSS 3.2.26 des modèles avec effets spatiaux différenciés.

Le modèle

Le modèle avec autocorrélation spatiale des erreurs à 3 décalages est défini de la façon suivante :

$$y = X\beta + B^{-1}u \quad [1]$$

où y est le vecteur $(N,1)$ des observations de la variable expliquée, β est le vecteur $(K,1)$ des K paramètres inconnus, X est la matrice (N,K) des observations des K variables explicatives, u est un vecteur d'erreur tel que $u \sim N(0, \sigma^2 I)$ et :

$$B = I - \lambda_1 W_1 - \lambda_2 W_2 - \lambda_3 W_3 \quad [2]$$

où W_1 est la matrice comportant uniquement les connexions spatiales entre les régions appartenant au régime Nord, W_2 est la matrice comportant uniquement les connexions spatiales entre les régions appartenant au régime Sud et W_3 est la matrice comportant uniquement les connexions spatiales entre les régions appartenant au régime Nord et celles appartenant au régime Sud. Ces matrices sont construites telles que $W_1 + W_2 + W_3 = W$, où W est la matrice de poids *standardisée*.

Fonction de vraisemblance

La fonction de vraisemblance de y par rapport aux coefficients inconnus β , λ_1 , λ_2 , λ_3 et σ^2 , est donnée par :

$$\ln L = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) + \ln|B| - \frac{1}{2\sigma^2} v'v \quad [3]$$

$$v'v = (y - X\beta)' B' B (y - X\beta)$$

Score

Soit $\Psi = [\beta \lambda_1 \lambda_2 \lambda_3 \sigma^2]'$. Le vecteur score $S(\Psi) = \frac{\partial \ln L}{\partial \Psi}$ est donné par :

$$\frac{\partial \ln L}{\partial \beta} = \frac{1}{\sigma} X' B' v \quad [4]$$

$$\frac{\partial \ln L}{\partial \lambda_i} = -\text{tr}(B^{-1} W_i) + \frac{1}{\sigma} v' W_i (y - X\beta) \quad i = 1, 2 \text{ et } 3 \quad [5]$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} v' v \quad [6]$$

Les estimateurs du maximum de vraisemblance sont obtenus en résolvant le système d'équation $S(\Psi) = 0$. Ce système n'est pas linéaire et n'admet pas de solutions analytiques. Pour le modèle [1], les solutions issues des conditions du premier ordre [4] et [6] permettent de construire une fonction de log-vraisemblance concentrée.

Fonction de vraisemblance concentrée

Des conditions de premier ordre [4] et [6], on tire les estimateurs de β et de σ^2 , notés $\hat{\beta}$ et $\hat{\sigma}^2$, qui dépendent de $\hat{\lambda}_i, i = 1, 2 \text{ et } 3$.

$$\hat{\beta} = (X' \hat{B}' \hat{B} X)^{-1} X' \hat{B}' \hat{B} y \quad [7]$$

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})' \hat{B}' \hat{B} (y - X\hat{\beta})}{N} \quad [8]$$

La substitution de [7] et [8] dans la fonction de log-vraisemblance [3] aboutit à la fonction de vraisemblance concentrée, qui ne dépend plus que des paramètres spatiaux, λ_1, λ_2 et λ_3 :

$$\ln L^* = \text{cste} - \frac{N}{2} \ln \left(\frac{1}{N} \cdot \hat{\varepsilon}' B' B \hat{\varepsilon} \right) + \ln |B| \quad [9]$$

avec $\hat{\varepsilon} = y - X\hat{\beta}$.

Matrice d'information

La matrice d'information du vecteur de paramètres Ψ , notée $I(\Psi)$, est donnée par :

$$I(\Psi) = -E \left[\frac{\partial L}{\partial \Psi} \frac{\partial L}{\partial \Psi'} \right].$$

On obtient :

$$I_{\beta\beta'} = \frac{1}{\sigma^2} X' B' B X \quad [10]$$

$$I_{\beta\lambda_i} = 0 \quad i = 1, 2 \text{ et } 3 \quad [11]$$

$$I_{\beta\sigma^2} = 0 \quad [12]$$

$$I_{\lambda_i\lambda_j} = \text{tr}(B^{-1}W_j B^{-1}W_i) + \text{tr}[(W_j B^{-1})W_i B^{-1}] \quad i, j = 1, 2 \text{ et } 3 \quad [13]$$

$$I_{\lambda_i\sigma^2} = \frac{1}{\sigma^2} W_i B^{-1} \quad i = 1, 2 \text{ et } 3 \quad [14]$$

$$I_{\sigma^2\sigma^2} = \frac{N}{2\sigma^4} \quad [15]$$

Matrice des variances-covariances

Nous supposons que les conditions habituelles de régularité sont respectées. Dans ces conditions, les estimateurs du maximum de vraisemblance, solutions du système $S(\Psi) = 0$, sont asymptotiquement normaux et efficaces. Cela implique que la matrice asymptotique des variances-covariances de $\hat{\Psi}$ atteint la borne inférieure de Cramer-Rao, donnée par l'inverse de la matrice d'information :

$$V_{as}(\hat{\Psi}) = [I(\hat{\Psi})]^{-1} \quad [16]$$

La matrice estimée $\hat{V}_{as}(\hat{\Psi})$ des variances-covariances est obtenue en remplaçant les paramètres par leurs estimations du maximum de vraisemblance.

Estimation du modèle [1]

La structure du programme GAUSS 3.2.26, associé à la librairie MAXLIK 4.0, destiné à estimer le modèle [1] est le suivant :

1/ Estimation du modèle : $y = X\beta + \varepsilon$, $\varepsilon \sim N(0, \sigma^2 I)$ par les MCO. Les estimations $\hat{\beta}_{MCO}$ et $\hat{\sigma}_{MCO}^2$ fournissent les valeurs initiales. $\lambda_1 = \lambda_2 = \lambda_3 = 0$ fournissent les valeurs initiales pour les paramètres spatiaux.

2/ Calcul des erreurs estimées initiales : $\hat{\varepsilon}_{MCO} = y - X\hat{\beta}_{MCO}$.

3/ Compte tenu de $\hat{\varepsilon}_{MCO}$ et $\hat{\sigma}_{MCO}^2$, maximisation par rapport à λ_1 , λ_2 et λ_3 de la fonction de vraisemblance concentrée [9].

4/ Compte tenu des estimations de λ_1 , λ_2 et λ_3 , calcul de $\hat{\beta}$ et de $\hat{\sigma}^2$ grâce à [7] et [8].

5/ Calcul des nouveaux résidus $\hat{\varepsilon} = y - X\hat{\beta}$.

6/ Si le critère de convergence est atteint ($< 10^{-6}$), aller à l'étape suivante, sinon retour à l'étape 3 jusqu'à convergence de tous les paramètres du modèle [1].

7/ Une fois la convergence atteinte, calcul analytique (et non pas par la procédure numérique du logiciel) de la matrice des variances-covariances de $\hat{\beta}_{ML}^2$, $\hat{\sigma}_{ML}^2$, $\hat{\lambda}_1$, $\hat{\lambda}_2$ et $\hat{\lambda}_3$ grâce à [16].

CHAPITRE 5

Convergence des régions européennes et chaînes de Markov

Une partie des résultats empiriques présentés dans ce chapitre est tirée du texte suivant :

Le Gallo J. (2001b) Space-time analysis of GDP disparities among European regions : a Markov chains approach, *Document de travail du LATEC*, Dijon, n°2001-06 (Mars 2001).

Chapitre 5

Convergence des régions européennes et chaînes de Markov

Les chapitres précédents ont mis en évidence la persistance des disparités de PIB par tête parmi les régions européennes malgré la forte croissance de certaines régions pauvres. Ce résultat a été obtenu, d'une part, à l'aide d'une analyse exploratoire des données spatiales et d'autre part, grâce à l'analyse de la β -convergence au sein de notre échantillon. Nous avons ainsi souligné la faible vitesse de convergence parmi les régions européennes ainsi que l'existence de deux clubs de convergence.

Cependant, le concept de β -convergence pose de nombreux problèmes méthodologiques tels que l'hétérogénéité, l'endogénéité et les problèmes de mesure (Temple, 1999). Un autre concept est parfois utilisé, celui de σ -convergence. Il consiste à analyser l'évolution d'un indicateur de dispersion de la distribution des PIB par tête dans le temps, l'indicateur le plus souvent utilisé étant l'écart-type de la distribution des PIB par tête en logarithmes sur un échantillon de régions (cf. chapitre 1, paragraphe 122.1, p. 55).

Quah (1993a, 1993b) souligne cependant que les indicateurs de dispersion ne fournissent aucune information sur le comportement de la distribution en coupe transversale dans sa totalité. Il suggère, par conséquent, un concept alternatif de convergence, basé sur l'étude de la forme des distributions de PIB par tête et sur l'analyse de matrices de transition. Cette méthodologie a été appliquée à différents groupes de régions ou de pays (Bianchi, 1997 ; Paap et Van Dijk, 1998 ; Desdoigts, 1997, 1999 ; Johnson, 2000). Elle permet l'identification de clubs de convergence qui se traduisent alors par une augmentation de l'homogénéité entre régions ou pays d'un même groupe, ainsi qu'une augmentation de la différence entre groupes.

Dans ce chapitre, nous analysons l'évolution des disparités entre les régions européennes en étudiant la distribution en coupe transversale des PIB par tête sur la période 1980-1995. L'organisation de ce chapitre est la suivante.

Dans la section 51, nous étudions le processus de convergence des régions européennes à l'aide de la méthodologie développée par Quah (1993, 1993b, 1996c). La convergence est alors mesurée à partir de l'évolution de la forme de la distribution des PIB par tête en coupe transversale et par les changements de positions relatives des régions dans cette distribution. L'étude de la forme de la distribution des PIB par tête et de son évolution sur la période 1980-1995 est principalement effectuée à l'aide de l'estimation non-paramétrique de fonctions de densité des distributions des PIB par tête. Pour étudier les

trajectoires de croissance des régions et évaluer leur mobilité au sein de cette distribution, nous avons modélisé le processus de croissance comme une chaîne de Markov de premier ordre stationnaire.

Cependant, à l'instar des concepts « traditionnels » de convergence, l'approche suggérée par Quah ne prend pas en compte l'interdépendance spatiale entre les régions qui sont traitées comme si elles étaient « isolées ». En conséquence, dans la section 52, nous analysons la façon dont l'espace doit être introduit dans ce cadre particulier et comment les résultats obtenus précédemment sont modifiés. Nous utilisons en particulier la procédure de conditionnement suggérée par Quah (1996b) ainsi que les chaînes de Markov spatiales proposées par Rey (2001). Ces outils permettent d'étudier comment les performances économiques d'une région peuvent être expliquées par son environnement géographique, d'évaluer l'ampleur de l'influence de l'environnement d'une région sur sa position dans la distribution des PIB par tête et de mettre en évidence le rôle de l'espace dans la constitution de clubs de convergence⁷⁰.

⁷⁰ Les résultats statistiques obtenus dans ce chapitre ont été obtenus à l'aide du logiciel SpaceStat 1.90 (Anselin, 1999). Les cartes et les figures ont été réalisées avec Arcview 3.2 (Esri), Gauss 3.2.26 (Aptech) et Statistica (StatSoft).

SECTION 51 :

LA DYNAMIQUE DE LA DISTRIBUTION DES PIB REGIONAUX EN EUROPE

L'étude de l'évolution de la distribution des PIB par tête de 1980 à 1995 s'effectue à partir des *PIB par tête relatifs à la moyenne européenne*⁷¹, encore appelé *conditionnement européen*. En effet, il est préférable de travailler sur les PIB relatifs plutôt que sur les PIB absolus afin d'éliminer les mouvements communs dus aux cycles d'affaire européens.

L'analyse de la distribution des PIB par tête s'effectue autour de deux axes principaux. Dans un premier temps, nous caractérisons la forme de la distribution des PIB régionaux pour chaque année, ainsi que les changements de cette forme au cours de la période considérée. Ceci est effectué à l'aide de trois méthodologies complémentaires. Tout d'abord, nous étudions l'évolution de la dispersion de la distribution des PIB par tête. Ensuite, nous mobilisons des méthodes d'estimation non-paramétriques appliquées à la distribution des PIB par tête. Enfin, nous analysons les caractéristiques de polarisation de cette distribution (paragraphe 511).

Dans un deuxième temps, l'estimation de matrices de transition entre différentes classes de PIB permet de préciser la dynamique temporelle dans la distribution des PIB régionaux. Nous présentons d'abord rapidement les concepts principaux nécessaires à l'utilisation des chaînes de Markov (paragraphe 512) et nous appliquons ensuite cette méthodologie à notre échantillon de 138 régions (paragraphe 513).

511 Evolution de la forme de la distribution des PIB par tête

Au sein des régions européennes, les schémas de convergence simples (polarisation en deux clubs de convergence) ou plus complexes (stratification impliquant plusieurs clubs de convergence), ainsi que l'évolution potentielle de ces schémas au cours de la période d'étude 1980-1995, peuvent être détectés par une étude de la forme et de l'évolution de la distribution des PIB par tête. Cette évolution au cours de la période est appréhendée à l'aide de trois méthodes complémentaires. Nous étudions d'abord l'évolution d'un indicateur synthétique, l'écart-type, dans le but d'appliquer le test de la σ -convergence (paragraphe 511.1). Ensuite, nous analysons la forme de la distribution des PIB par tête en estimant cette dernière à l'aide de la méthode non-paramétrique du noyau (paragraphe 511.2). Nous calculons également l'indice de polarisation proposé par Esteban et Ray (1994) dans le but d'affiner les résultats trouvés dans les paragraphes précédents (paragraphe 512.3). Enfin, nous soulignons les limites de ces différentes analyses pour l'étude de la convergence (paragraphe 512.4).

⁷¹ Ils sont définis comme le ratio des PIB par tête sur la moyenne européenne des PIB par tête.

511.1 Le test de la σ -convergence

Si l'on souhaite étudier globalement l'évolution de la distribution des PIB par tête, une procédure simple consiste à appliquer le test de la σ -convergence qui permet de comparer l'écart-type de la distribution des PIB par tête en logarithmes calculé pour l'année finale avec l'écart-type calculé pour l'année initiale. Si la dispersion est plus faible à la fin de la période, il y a σ -convergence.

Nous avons calculé l'écart-type de la distribution des PIB par tête en logarithmes pour toutes les années de la période 1980-1995. Les résultats sont présentés dans le graphique 5.1a pour les PIB par tête en Ecus et dans le graphique 5.1b pour les PIB par tête en SPA.

Si l'on s'intéresse à l'évolution de la dispersion des PIB par tête pour l'ensemble de l'échantillon, il apparaît que la dispersion est légèrement inférieure en 1995 comparée à celle de 1980 ($\sigma_{1980} = 0,443$ et $\sigma_{1995} = 0,412$ pour les PIB par tête en Ecus ; $\sigma_{1980} = 0,291$ et $\sigma_{1995} = 0,275$ pour les PIB par tête en SPA). Ces résultats semblent donc indiquer la présence d'une faible σ -convergence pour les 138 régions européennes de notre échantillon. Cependant, nous constatons que cette diminution n'est pas uniforme. Les 8 premières années sont marquées par des perturbations durant lesquelles aucune tendance claire ne se dégage. Après une baisse nette pendant la période 1987-1992, la dispersion augmente à nouveau légèrement dans les dernières années.

Dans le chapitre précédent, l'existence de deux régimes de convergence a été mise en évidence, l'un concernant les régions du Sud de l'Europe et l'autre concernant les régions du Nord. On peut alors souhaiter vérifier si le processus de σ -convergence est différent dans ces deux régimes. Nous procédons à une partition de l'échantillon global en deux sous-échantillons selon les mêmes critères que dans le chapitre 4 (paragraphe 421.2, p. 323). Ensuite, nous calculons l'écart-type de la distribution en logarithmes pour chacun des deux sous-échantillons. Les résultats sont présentés dans le graphique 5.1a pour les PIB par tête en Ecus et dans le graphique 5.1b pour les PIB par tête en SPA.

La conclusion précédente de σ -convergence pour l'échantillon global ne tient plus dès lors que l'on considère les deux sous-échantillons. En effet, pour les régions du Sud de l'Europe, il y a σ -convergence ($\sigma_{1980} = 0,317$ et $\sigma_{1995} = 0,253$ pour les PIB par tête en Ecus ; $\sigma_{1980} = 0,244$ et $\sigma_{1995} = 0,175$ pour les PIB par tête en SPA). Après une période d'augmentation modérée jusqu'en 1989, l'écart-type de la distribution des PIB par tête en logarithmes diminue clairement après cette date. En revanche, on n'observe pas de phénomène de σ -convergence pour les régions du Nord ($\sigma_{1980} = 0,193$ et $\sigma_{1995} = 0,242$ pour les PIB par tête en Ecus ; $\sigma_{1980} = 0,177$ et $\sigma_{1995} = 0,193$ pour les PIB par tête en SPA). Ces résultats renforcent ainsi les conclusions trouvées à la suite de l'étude de la β -convergence.

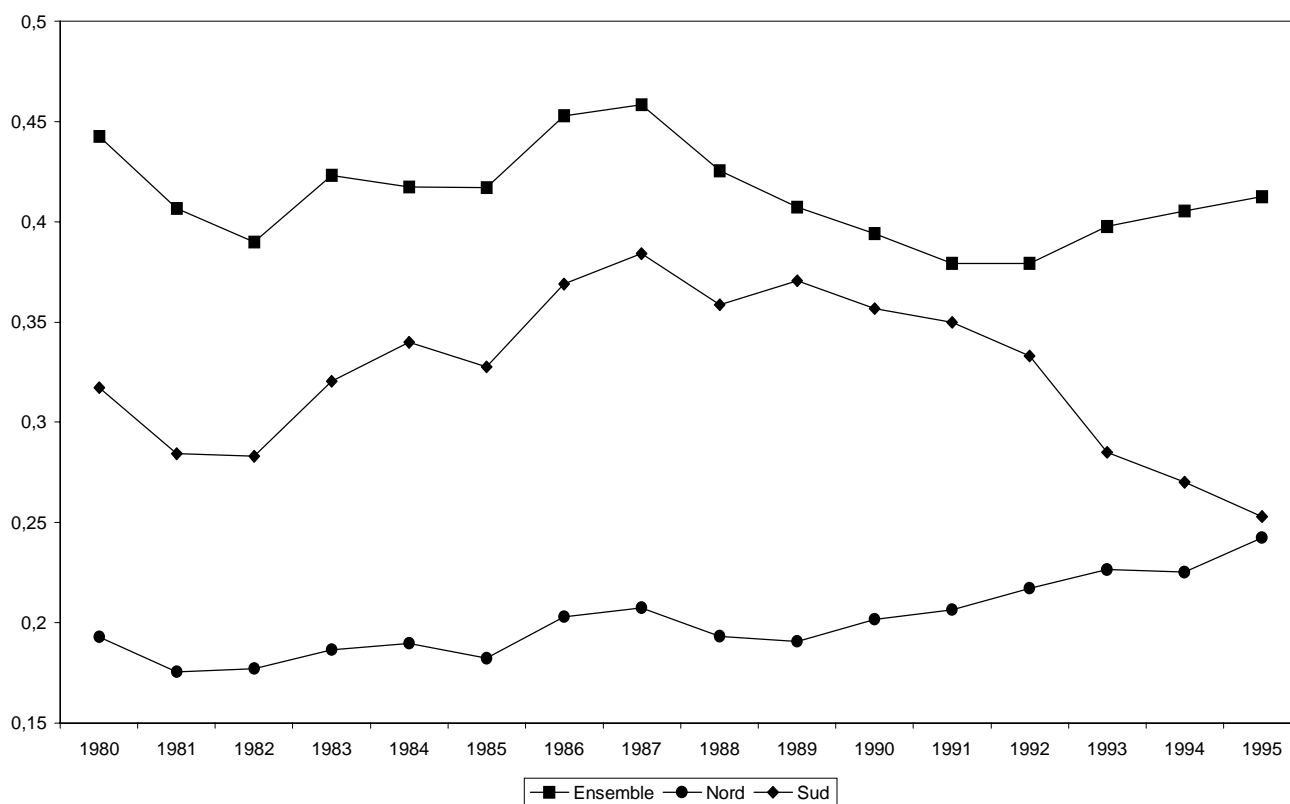


Figure 5.1a : σ -convergence pour la période 1980 - 1995 ; PIB par tête en Ecus

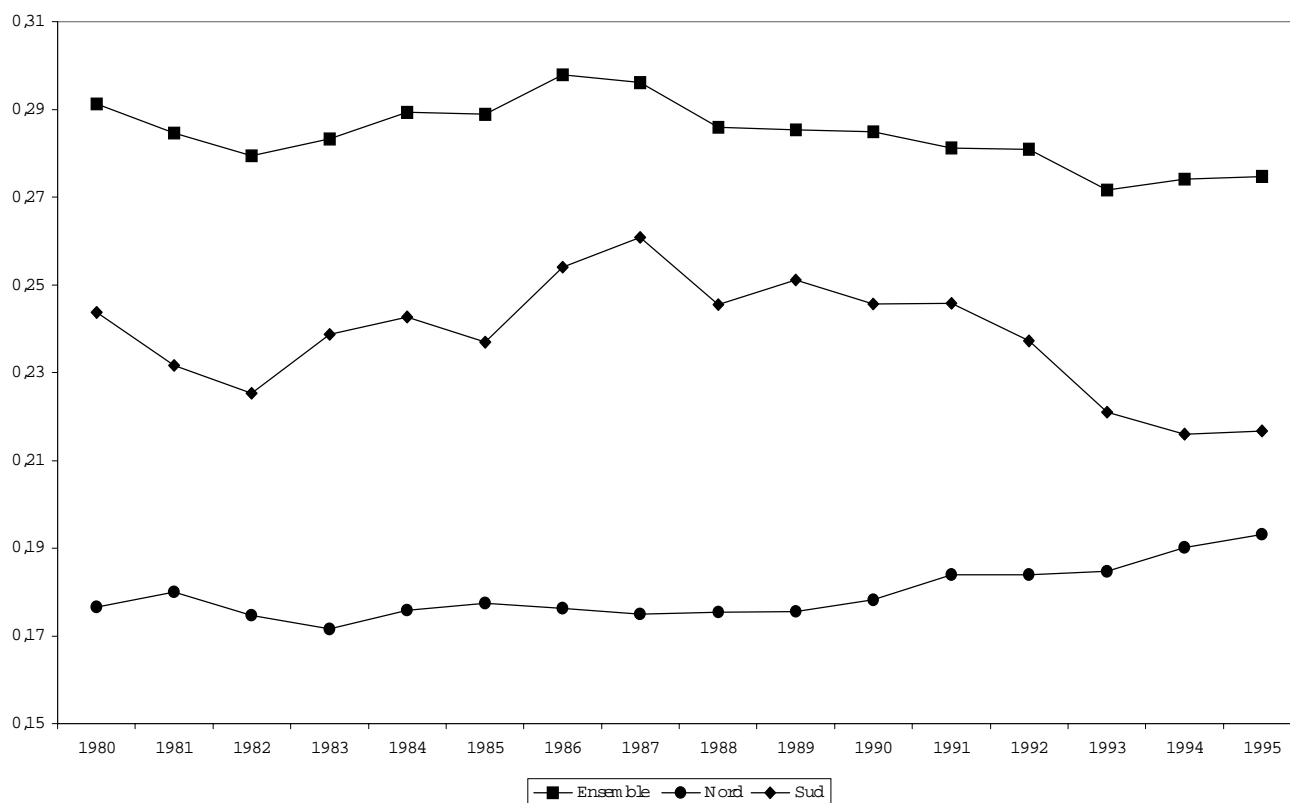


Figure 5.1b : σ -convergence pour la période 1980 - 1995 ; PIB par tête en SPA

Ces résultats soulignent la possibilité d'existence de deux clubs de convergence. Cependant, compte tenu des limites rencontrées par l'approche en termes de σ -convergence, il est important de compléter notre étude par une analyse plus individuelle de la place occupée par chaque région dans la distribution des PIB par tête, ainsi que par l'évolution de cette place au cours du temps.

511.2 Estimation non paramétrique des fonctions de densité : principes et résultats

La caractérisation des clubs de convergence peut être effectuée en estimant les fonctions de densité pour la distribution des PIB par tête des pays et en analysant les caractéristiques de mono- ou de multimodalité de ces fonctions. En utilisant cette méthodologie, Quah (1993b, 1996f), Bianchi (1997) et Jones (1997) concluent à la bimodalité pour les revenus internationaux, c'est-à-dire à l'existence de 2 clubs de convergence. En revanche, Quah (1996b) n'obtient pas ce résultat pour un échantillon des régions européennes.

Pour caractériser l'évolution des PIB régionaux, nous examinons la distribution des PIB par tête relatifs à la moyenne européenne en 1980 et la façon dont cette distribution évolue dans le temps jusqu'à 1995. Dans ce but, nous utilisons l'estimateur du noyau. Formellement, l'estimateur du noyau d'une série X au point x est donné par :

$$f(x) = \frac{1}{Nh} \sum_{i=1}^N K\left(\frac{x - X_i}{h}\right) \quad [5.1]$$

où N est le nombre d'observations, h est la fenêtre qui détermine le degré de lissage de la série, $K(\cdot)$ est une fonction qui détermine la forme du noyau. Cette fonction doit posséder les propriétés suivantes :

$$1/ K(u) \geq 0 \quad \forall u$$

$$2/ \int_{-\infty}^{\infty} K(u).du = 1$$

$$3/ K(u) = K(-u)$$

Différentes formes de la fonction $K(\cdot)$ sont possibles possédant diverses propriétés théoriques d'optimalité et de vitesse de convergence (Silverman, 1986 ; Izenman, 1991).

Deux fonctions souvent utilisées sont la fonction Epanechnikov (équation [5.2]) et la fonction normale (équation [5.3]) :

$$K(u) = \frac{3}{4} (1 - u^2) I(|u| \leq 1) \quad [5.2]$$

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} u^2\right) \quad [5.3]$$

où u est l'argument de la fonction et $I(\cdot)$ est une fonction indicatrice qui prend la valeur 1 si l'argument est vrai et 0 sinon.

Pour déterminer la fenêtre, Silverman (1986) propose d'utiliser la formule suivante :

$$h = 0,9kN^{-1/5} \min\{s, R/134\} \quad [5.4]$$

où N est le nombre d'observations, s est l'écart-type, R est l'écart interquartile de la série. Le facteur k est une transformation qui diffère selon la fonction choisie.

Dans ce paragraphe, toutes les densités sont calculées avec un noyau gaussien et la fenêtre fixe optimale [5.4].

La figure 5.2a comprend deux graphiques indiquant les fonctions de densité des PIB par tête relatifs à la moyenne européenne en Ecus respectivement pour l'année initiale 1980 et pour l'année finale 1995. La figure 5.2b présente ces fonctions de densité pour les PIB par tête en SPA. Ces fonctions de densité peuvent être interprétées comme l'équivalent continu d'un histogramme, dans lequel le nombre d'intervalles tend vers l'infini. Par définition des données, le nombre 1 sur l'axe des abscisses indique le PIB moyen européen, le nombre 2 indique le double du PIB moyen, etc.

Lorsque les PIB par tête sont mesurés en Ecus, la comparaison de l'année initiale avec l'année finale fait apparaître qu'en 1995, plus de régions possèdent des PIB par tête inférieurs à la moyenne européenne ou supérieurs à deux fois la moyenne européenne. En plus du mode principal situé aux alentours de la moyenne européenne, on observe également la persistance sur la période étudiée d'un second mode. Ce dernier se situe autour de 50% de la moyenne européenne. Ce mode secondaire peut refléter l'existence d'un groupe important de régions avec des niveaux de PIB par tête inférieurs à la moyenne européenne et qui ont tendance à converger vers un niveau inférieur de celui du reste des régions. Notons cependant que la distance entre les deux modes pour notre échantillon est bien inférieure à celle détectée dans les échantillons comprenant les différents pays du monde (Bianchi, 1997 ; Desdoigts, 1999).

Concernant la distribution des PIB par tête en SPA, la bimodalité ne semble apparaître qu'en fin de période. Cette différence entre les deux variables confirme que les disparités régionales sont lissées lorsqu'on utilise les PIB avec la correction des parités de pouvoir

d'achat. Pour les deux variables, on observe également l'existence d'un petit pic situé à environ 80% de la moyenne européenne sur les fonctions de densité de 1995. Ce résultat semble indiquer une tendance à la divergence des régions très riches.

Ces résultats contredisent ceux obtenus par Quah (1996b) mais ils sont similaires à ceux de López-Bazo *et al.* (1999). L'explication tient surtout aux échantillons utilisés dans notre thèse et dans celle de López-Bazo *et al.* (1999) : nous avons inclus les régions pauvres du Portugal et de la Grèce. Au contraire, l'échantillon utilisé par Quah (1996b) comporte 78 régions de 1980 à 1989 et ces régions pauvres ne sont pas incluses.

511.3 Etude de la polarisation

Les différentes fonctions de densité calculées dans le paragraphe précédent semblent suggérer une polarisation progressive des PIB par tête régionaux, c'est-à-dire le regroupement des régions autour de deux pôles locaux. Le premier pôle comprend les régions qui possèdent un PIB par tête inférieur à 75% de la moyenne européenne (rappelons que le seuil de 75% de la moyenne européenne est la définition utilisée par la Commission Européenne pour déterminer l'éligibilité des fonds structurels à l'Objectif 1). Le second pôle comprend toutes les autres régions.

Différents indices peuvent être calculés pour apprécier les plus ou moins grandes inégalités de répartition. Dans notre cas, la détection d'un schéma de polarisation nécessite d'apprécier les inégalités, non plus seulement au niveau global, mais également au niveau local. Comme le montrent Esteban et Ray (1994) et Wolfson (1994), la polarisation ne peut pas être mesurée par l'indice de Lorenz ou ses dérivés. En effet, ces derniers ne permettent pas de distinguer entre la convergence vers la moyenne globale de la distribution et la convergence vers des pôles locaux. Nous utilisons par conséquent l'indice de polarisation défini par Esteban et Ray (1994) :

$$ER = \sum_{i=1}^n \sum_{j=1}^n f_i^{1+\alpha} f_j |y_i - y_j| \quad [5.5]$$

où f_i est la fréquence de la population de la classe i , y_i est le logarithme de la moyenne conditionnelle de la classe i , n est le nombre de classes et α est « l'effet d'opposition » qui doit être restreint à l'intervalle $[0, 1.6]$ pour vérifier un certain nombre d'axiomes, comme le prouvent Esteban et Ray (1994).

Dans [5.5], l'indice ER est proportionnel au coefficient de Gini lorsque $\alpha = 0$. En revanche, plus α est important, plus l'indice ER diffère du coefficient de Gini.

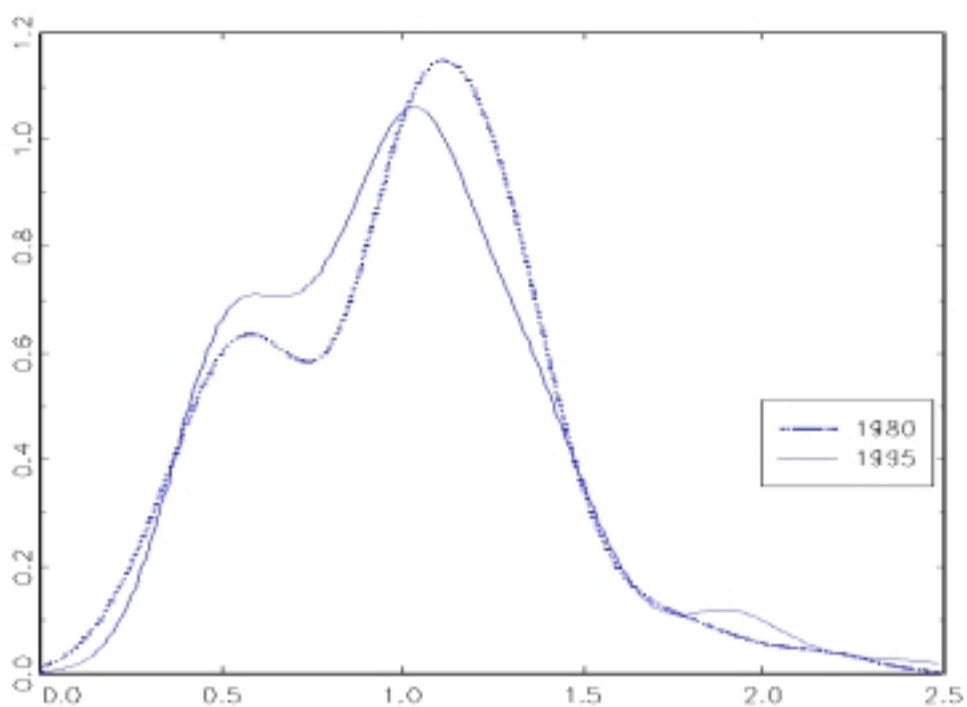


Figure 5.2a : Densités des distributions de PIB par tête en Ecus ;
conditionnement européen

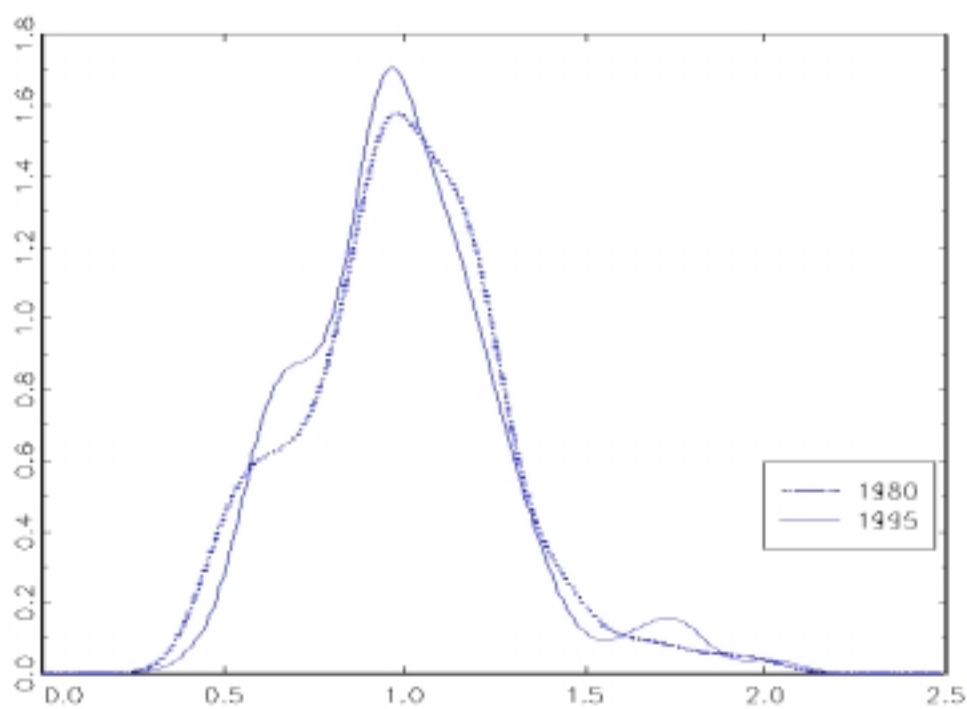


Figure 5.2b : Densités des distributions de PIB par tête en SPA ;
conditionnement européen

La figure 5.3 représente l'évolution de l'indice de polarisation pour les deux groupes de PIB par tête définis précédemment sur la période 1980-1995. Le calcul est effectué à la fois pour le PIB par tête en Ecus et en SPA. Le paramètre α est fixé à 1,5 pour que la polarisation soit fortement pondérée.

Concernant le PIB par tête en Ecus, on constate que l'indice ER a légèrement augmenté sur la période 1980-1995. Il y a eu une forte décroissance entre 1987 et 1991, mais la polarisation augmente à nouveau à la fin de la période. Globalement, ces résultats indiquent que la polarisation persiste entre le groupe à haut PIB et le groupe à bas PIB. Lorsque le PIB par tête est mesuré en SPA, l'indice de polarisation est inférieur à celui obtenu avec le PIB en Ecus : il est approximativement égal à 0,03 sur toute la période. Il existe cependant une légère tendance à l'augmentation de l'indice ER pour les 3 dernières années de la période. La tendance à la polarisation apparaît donc moins forte lorsque la correction des pouvoirs d'achat est effectuée.

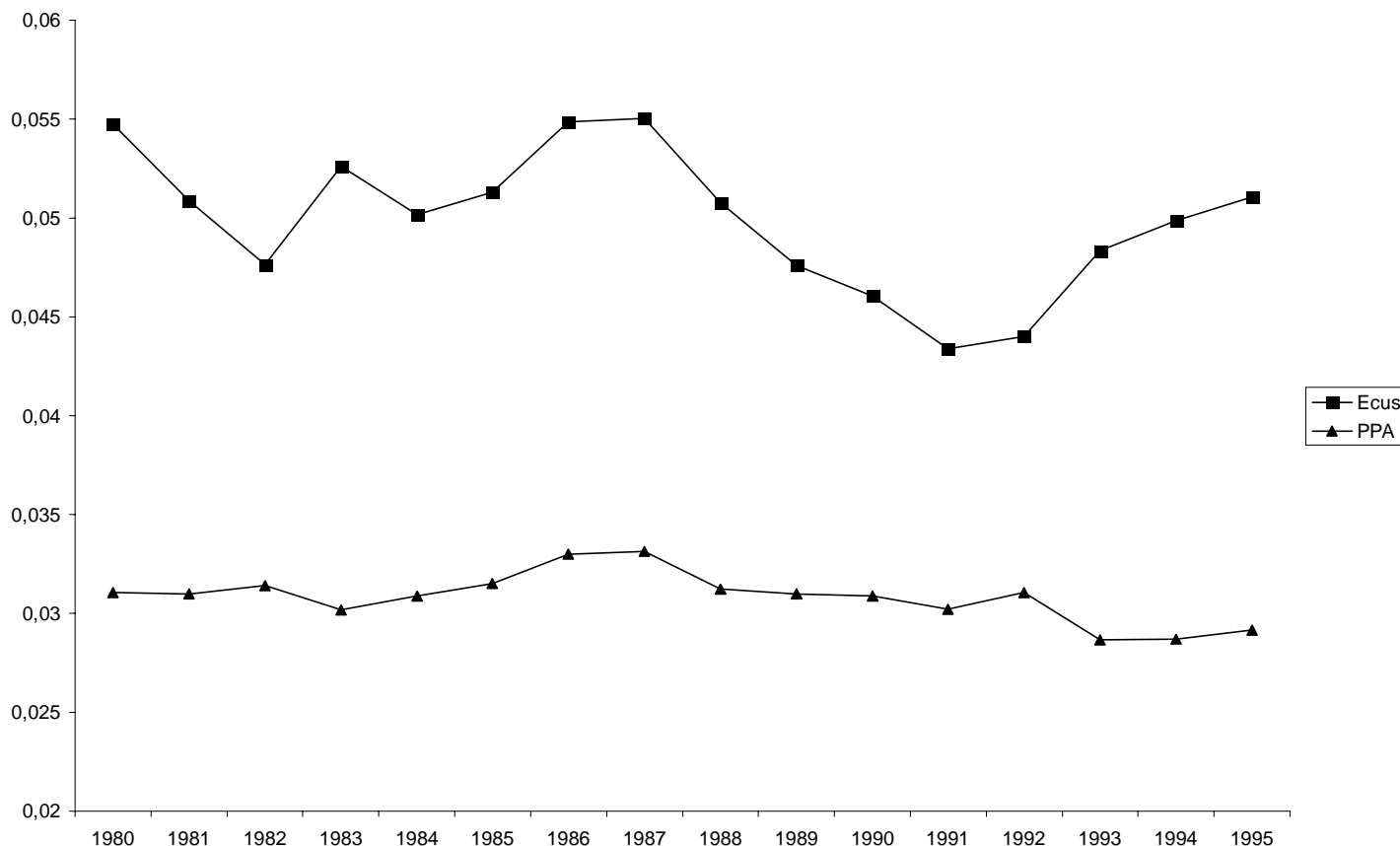


Figure 5.3 : Indice de polarisation ($\alpha = 1.5$) ; 1980-1995

511.4 Limites de ces approches

Les fonctions de densité indiquent qu'il y a plus de régions possédant des PIB par tête très élevés ou très faibles en 1995 qu'en 1980. Cependant, ces fonctions de densité ne permettent pas de révéler si la queue de droite de la distribution initiale (1980) contient les mêmes régions que la queue de droite de la distribution finale (1995). En d'autres termes, si ces fonctions de densité permettent de caractériser l'évolution de la distribution globale, elles ne fournissent pas d'information sur les mouvements des régions à l'intérieur de cette distribution. La mesure de σ -convergence et l'indice de polarisation ne permettent pas non plus de répondre à cette interrogation.

Considérons par exemple la figure 5.4 qui indique une évolution possible de la distribution des PIB par tête. Initialement, cette distribution est unimodale alors qu'à la fin de la période, la distribution est bimodale. Cependant, cette évolution globale ne permet pas de capturer la trajectoire de croissance des différentes régions : certaines régions sont montées dans la hiérarchie, d'autres ont vu leur situation se dégrader alors que quelques régions sont restées à leur niveau initial (Quah, 1996b, 1996e).

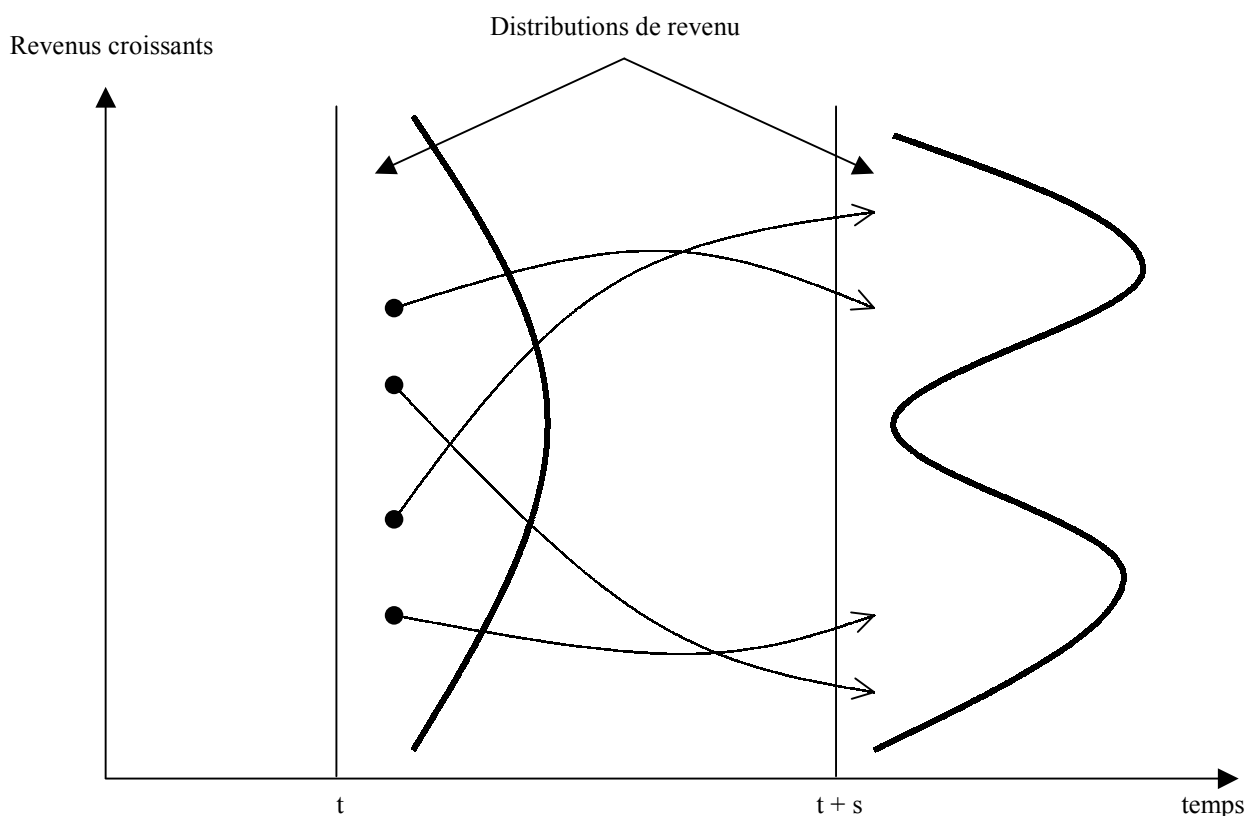


Figure 5.4 : Stratification, polarisation et clubs de convergence

Source : Adaptation de Quah (1996b), p. 1369

Pour répondre aux interrogations précédentes, il faut analyser l'évolution de la trajectoire de croissance de chaque région dans le temps. Pour cela, nous modélisons le processus de croissance des régions européennes comme une chaîne de Markov.

512 Chaînes de Markov : principes

Avant d'aborder les résultats empiriques obtenus à l'aide de la méthodologie des chaînes de Markov, nous présentons dans ce paragraphe les outils fondamentaux nécessaires à l'analyse en termes de chaîne de Markov (Chung, 1960 ; Kemeny et Snell, 1976 ; Berchtold, 1998). Tout d'abord, nous définissons un processus de chaîne de Markov stationnaire de premier ordre ainsi que la matrice de transition associée à ce processus dans le paragraphe 512.1. Une typologie des processus de Markov est proposée dans le paragraphe 512.2. Enfin, le comportement de long terme des processus de Markov est analysé dans le paragraphe 512.3.

512.1 Définitions

Une chaîne de Markov est un processus stochastique dont les probabilités de transition sont des probabilités conditionnelles au passé.

Formellement, soit X_t une variable aléatoire discrète ($X_t ; t \in \mathbb{N}$) prenant ses valeurs dans l'ensemble fini $V = \{1, \dots, m\}$. Le but est de calculer la probabilité que la variable aléatoire X_t prenne chacune des m différentes modalités possibles au temps t , connaissant la modalité prise par la variable aux périodes précédentes $t-1$, $t-2$, etc :

$$P[X_t = j | X_{t-1} = i, X_{t-2} = i_2, \dots] \quad [5.6]$$

1) Processus de Markov d'ordre 1

Si le processus de Markov est d'ordre 1, l'ensemble du passé de l'époque t est résumé par l'époque $t-1$. Dans ce cas, la probabilité que la variable aléatoire X_t prenne la valeur j au temps t ne dépend que de l'état dans lequel se trouvait la variable aléatoire au temps $t-1$. Cette probabilité est notée $p_{ij}(t)$ et est définie de la façon suivante :

$$p_{ij}(t) = P[X_t = j | X_{t-1} = i] \quad [5.7]$$

On définit de manière similaire un processus d'ordre 2 ou d'ordre supérieur.

2) *Matrice de transition*

Les probabilités $p_{ij}(t)$ correspondant aux différentes valeurs possibles de i et de j sont résumées par la matrice de transition de dimension (m, m) , appelée matrice de transition de l'époque $t - 1$ à l'époque t :

$$M_{(t-1,t)} = [p_{ij}(t)]$$

[5.8a]

soit
$$M_{(t-1,t)} = \begin{pmatrix} p_{1,1}(t) & \cdots & p_{1,m}(t) \\ \vdots & \ddots & \vdots \\ p_{m,1}(t) & \cdots & p_{m,m}(t) \end{pmatrix}$$
 [5.8b]

Chacune des lignes de la matrice de transition est une loi de probabilité, ce qui implique que la somme des éléments de chaque ligne est égale à 1.

3) *Processus de Markov stationnaire*

La matrice $M_{(t-1,t)}$ dépend du temps t . Il existe donc une matrice pour chaque époque. Cependant, s'il apparaît que les probabilités de transition ne varient pas d'une période à l'autre, il est possible de calculer une matrice de transition résumant l'ensemble de la période. Dans ce cas, les probabilités de transition sont stationnaires. En d'autres termes, la probabilité de passer d'un état à un autre est constante dans le temps :

$$p_{ij}(t) = p_{ij} \quad \forall t$$

[5.9]

Si les probabilités de transition sont stationnaires, alors le processus de Markov est complètement déterminé par la matrice de transition suivante :

$$M = [p_{ij}] = \begin{pmatrix} p_{1,1} & \cdots & p_{1,m} \\ \vdots & \ddots & \vdots \\ p_{m,1} & \cdots & p_{m,m} \end{pmatrix}$$

[5.10]

M résume l'ensemble des probabilités de passage d'un état d'une époque à l'état de l'époque suivante.

4) Probabilités des états futurs

Le concept de chaîne de Markov permet de déterminer la probabilité que la variable aléatoire prenne une valeur donnée après un nombre déterminé de périodes.

On démontre, en effet, que la probabilité que la chaîne prenne la valeur j à la période $n + r$, sachant qu'elle a pris la valeur i à la période n , se détermine de la façon suivante :

$$P(X_{n+r} = j | X_n = i) = p_{ij}^r \quad [5.11]$$

où p_{ij} est la probabilité que la variable aléatoire X_t prenne la valeur j à la période t sachant qu'elle a pris la valeur i à la période $t - 1$. Les équations de Chapman-Kolmogorov donnent la façon dont ces probabilités sont calculées :

$$p_{ij}^{n+r} = \sum_{k=0}^{\infty} p_{i,k}^n p_{k,j}^r \quad [5.12]$$

Sous forme matricielle, définissons F_t le vecteur-ligne de taille m comprenant les probabilités que la variable X_t prenne chacune des valeurs possibles à l'époque t :

$$F_t = (P[X_t = 1] \quad \dots \quad P[X_t = m]) \quad [5.13]$$

Soient F_0 le vecteur de probabilités à l'époque 0 et M la matrice de transition contenant les transitions de probabilités stationnaires, alors les vecteurs de probabilités des périodes successives se calculent comme suit :

$$F_1 = F_0 M \quad [5.14]$$

$$F_2 = F_1 M = F_0 M^2 \quad [5.15]$$

et ainsi de suite jusqu'à :

$$F_t = F_{t-1} M = F_0 M^t \quad [5.16]$$

Par conséquent, la détermination des probabilités des états futurs passe par le calcul des puissances successives de la matrice de transition M . Dans le cas des matrices de transition régulières, que nous définissons dans le paragraphe suivant, la structure de M^t lorsque t tend vers l'infini prend une forme très particulière. La chaîne de Markov est donc entièrement définie dès lors que le vecteur F_0 et la matrice M sont connus.

512.2 Typologie des chaînes de Markov

Nous présentons d'abord une classification des différents états que peut prendre le système (paragraphe 512.21) et nous en déduisons une typologie des chaînes de Markov (paragraphe 512.22).

512.21 Classification des états

Nous reprenons les notations précédentes : X_t est une variable aléatoire discrète ($X_t ; t \in \mathbb{N}$) prenant ses valeurs dans l'ensemble fini $V = \{1, \dots, m\}$, M est une matrice de transition homogène. Enfin, soit $E = \{e_1, \dots, e_m\}$ l'ensemble des états du système.

La classification des états d'une chaîne de Markov est effectuée selon le critère suivant : est-il possible d'aller d'un état vers un autre état donné ? Les définitions qui suivent sont tirées de Kemeny et Snell (1976) et Berchtold (1998).

Définition 1 : Probabilité de premier retour

Supposons qu'à l'époque initiale le système est défini par l'état e , $f_e(t)$ est la probabilité que le système retourne en e pour la première fois à l'époque t :

$$f_e(t) = P[X_t = e, X_{t-1} \neq e, X_{t-2} \neq e, \dots, X_1 \neq e | X_0 = e] \quad [5.17]$$

Définition 2 : Probabilité de retour

Par extension, la probabilité de retour à l'état e , notée f_e , est la somme pour t allant de 1 à l'infini des probabilités de premier retour à l'état e :

$$f_e = \sum_{t=1}^{\infty} f_e(t) \quad [5.18]$$

Définition 3 : Espérance de retour

L'espérance (ou temps moyen) de retour à l'état e , notée μ_e , est l'espérance des probabilités de premier retour.

Ces trois concepts permettent d'obtenir une typologie des états d'une chaîne de Markov. La probabilité de retour peut ainsi être inférieure ou égale à 1, l'espérance de retour peut être finie ou non, etc.

Définition 4 : Etat persistant

Un état e est dit persistant si sa probabilité de retour est égale à 1. Si e est un état persistant, alors l'espérance de retour se calcule de la façon suivante :

$$\mu_e = \sum_{t=1}^{\infty} t \cdot f_e(t) \quad [5.19]$$

Si $\mu_e \rightarrow \infty$, alors l'état e est dit nul. En revanche, si l'espérance de retour de e est finie, l'état e est dit non-nul.

Définition 5 : Etat de transition

Un état e est appelé état de transition si sa probabilité de retour est strictement inférieure à 1. Si e est un état de transition, alors son espérance de retour est définie comme étant infinie.

Définition 6 : Etat périodique

Un état e est dit périodique d'ordre λ , si, sachant que $X_0 = e$, il n'est possible de retourner en e que dans une période s'écrivant $k\lambda$, $k \in \mathbb{N} \setminus \{0,1\}$. En d'autres termes, un état e est périodique, si sachant que $X_0 = e$, le plus grand commun diviseur des périodes lors desquelles il est possible de retourner en e est strictement supérieur à 1.

Par exemple, considérons la chaîne de Markov d'ordre 1 avec 4 états, associée à la matrice de transition M suivante :

$$M = \begin{matrix} a \\ b \\ c \\ d \end{matrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0,5 & 0 & 0,5 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad [5.20]$$

Dans cet exemple, les 4 états sont tous périodiques d'ordre 2. En effet, si le système démarre dans l'état a ou dans l'état b , il ne peut retourner dans cet état que 2 périodes plus tard. Si le système démarre dans l'état c ou dans l'état d , il ne peut retourner dans cet état que 4 périodes plus tard.

Définition 7 : Etat ergodique

Un état e est dit ergodique s'il est non-nul (son espérance de retour est finie) et non-périodique.

Définition 8 : Etat absorbant

Un état est dit absorbant si une fois atteint cet état, il est impossible de le quitter :

$$P[X_{t+1} = e | X_t = e] = 1 \quad [5.21]$$

Par conséquent, une matrice de transition comporte autant d'états absorbants qu'il existe de probabilités égales à 1 sur sa diagonale principale.

512.22 Classification des chaînes de Markov

Définissons une classe d'états comme un sous-ensemble non-vidé de l'ensemble des états E de la chaîne de Markov. Une typologie des chaînes de Markov se déduit des différents types de classes d'états qui peuvent la composer.

Définition 9 : Classe ergodique

Une classe est dite ergodique si tous les états la composant sont ergodiques.

Définition 10 : Classe irréductible

Une classe est dite irréductible s'il est possible de passer en un nombre fini d'étapes de n'importe quel état de cette classe à n'importe quel autre état de cette même classe. Dans une classe irréductible, tous les états sont de même nature (ergodiques, non-périodiques, etc., ...).

Définition 11 : Chaîne de Markov ergodique

Une chaîne de Markov à nombre fini d'états est dite « ergodique » si elle est irréductible et si tous ses états sont ergodiques.

Définition 12 : Chaîne de Markov régulière

Une chaîne de Markov à nombre fini d'états est dite « régulière » si elle est formée d'une seule classe ergodique et irréductible. En d'autres termes, une chaîne est régulière si elle est formée d'une seule classe irréductible et non-périodique. Nous verrons dans le paragraphe 513.2 (p. 376) que c'est ce type de chaîne de Markov qui correspond à notre étude empirique.

Les chaînes de Markov régulières possèdent plusieurs propriétés importantes (Kemeny et Snell, 1976, chap. 4). Soit une chaîne de Markov régulière d'ordre 1, et soit M sa matrice de transition. M est alors appelée une matrice de transition régulière.

On démontre alors qu'une matrice de transition est régulière si et seulement si, il existe un entier k tel que la matrice M^k ne contient pas d'éléments nuls.

512.3 Comportement limite des chaînes de Markov

Considérons une chaîne de Markov régie par l'égalité matricielle [5.16] :

$$F_t = F_0 M^t$$

Nous nous interrogeons maintenant sur la possibilité d'existence d'un vecteur d'état limite F^* lorsque le nombre de transitions devient très élevé. Lorsque la chaîne de Markov est régulière, Kemeny et Snell démontrent le théorème suivant (p. 70) :

Théorème 1 :

Si M est une matrice de transition régulière alors :

1/ Les puissances M^t approchent une matrice limite, notée M^* :

$$\lim_{t \rightarrow \infty} M^t = M^* \quad [5.22]$$

2/ Toutes les lignes de M^* sont égales entre elles et égales au même vecteur de probabilité limite, noté F^* .

3/ Les composantes de F^* sont positives.

Ainsi, une matrice de transition régulière converge vers une matrice limite M^* de rang 1. Soit $F^* = [F^{*1} \dots F^{*m}]$, alors la matrice limite M^* prend la forme suivante :

$$M^* = \begin{pmatrix} F^{*1} & \dots & F^{*m} \\ \vdots & \ddots & \vdots \\ F^{*1} & \dots & F^{*m} \end{pmatrix} \quad [5.23]$$

avec $F^{*i} > 0, i = 1, \dots, m$.

Le théorème suivant permet de déterminer la distribution limite, ou distribution ergodique F^* de la chaîne de Markov (Kemeny et Snell, 1976, p. 71).

Théorème 2 :

Si M est une matrice de transition régulière alors :

1/ Pour chaque vecteur de probabilité F , le vecteur FM^t tend vers F^* lorsque t tend vers l'infini :

$$\lim_{t \rightarrow \infty} FM^t = F^* \quad [5.24]$$

2/ Le vecteur F^* est l'unique vecteur tel que :

$$F^*M = F^* \quad [5.25]$$

3/ Les égalités suivantes sont vérifiées :

$$M.M^* = M^*.M = M^* \quad [5.26]$$

De l'égalité [5.25], on déduit que la distribution ergodique est donnée par le vecteur propre associé à la valeur propre unitaire de M .

Après cette présentation rapide des principaux concepts liés à l'analyse en termes des chaînes de Markov, nous appliquons maintenant cette méthodologie à l'analyse de la convergence des régions européennes.

513 Convergence et processus de Markov : résultats empiriques

Quah (1996a) propose d'analyser la dynamique des disparités régionales et montre comment le processus de croissance peut être modélisé à l'aide d'une chaîne de Markov. Nous appliquons cette méthodologie et nous indiquons quels sont les principaux résultats obtenus sur divers échantillons de régions européennes (paragraphe 513.1). Ensuite, nous abordons la façon dont les probabilités de transition sont calculées dans ce cadre particulier et nous présentons la matrice de transition estimée pour notre échantillon de régions européennes (paragraphe 513.2). En outre, nous présentons le test de stabilité des probabilités de transition entre deux sous-périodes (paragraphe 513.3). Enfin, nous analysons les principaux résultats obtenus en termes de convergence des régions européennes (paragraphe 513.4).

513.1 Principes de modélisation

Notons F_t la distribution en coupe transversale des PIB par tête relatifs à la moyenne européenne au temps t . On considère K classes différentes de PIB par tête, qui fournissent une approximation discrète de la distribution des PIB par tête. Nous supposons dans ce paragraphe que la fréquence de la distribution dans le temps suit un processus de Markov de premier ordre et stationnaire. Dans ce cas, l'évolution de la distribution des PIB par tête est représentée par une matrice de transition, notée M , dans laquelle chaque élément (i, j) indique la probabilité qu'une région qui se trouve dans la classe i à la période $t - 1$ se retrouve dans la classe j à la période t .

Dans ces conditions, le vecteur F_t de dimension $(K,1)$, indiquant la fréquence des régions dans chaque classe au temps t , est supposé évoluer de la façon suivante :

$$F_{t+1} = MF_t \quad [5.27]$$

où M est la matrice (K,K) de transition représentant la transition d'une distribution à l'autre.

La matrice M possède un certain nombre de propriétés qui peuvent être exploitées pour l'analyse de l'évolution de la distribution des PIB.

Premièrement, il est intéressant de connaître la propension des régions dans chaque classe à se déplacer dans les autres classes ainsi que le temps moyen nécessaire à une région pour se déplacer entre chaque paire de classes i et j . Ces informations seront fournies par l'estimation des probabilités de transition pour notre échantillon et par la détermination d'une matrice des temps moyens de premier passage (cf. paragraphe 513.42, p. 382).

Deuxièmement, si la matrice de transition estimée est régulière, il est possible de déterminer sa distribution ergodique. Cette distribution limite permet de caractériser le type de convergence ayant lieu entre les régions européennes.

Troisièmement, l'estimation d'une matrice de transition aboutit au calcul de différents indices permettant de caractériser la mobilité des régions entre les différentes classes.

Dans le cas des régions européennes, différentes études basées sur cette méthodologie ont été réalisées. Plusieurs bases de données sont utilisées dans les travaux empiriques.

Neven et Gouyette (1995), López-Bazo *et al.* (1999) et Maurseth (2001) utilisent la base de données Eurostat-Regio. Neven et Gouyette utilisent un échantillon de 142 régions au niveau NUTS II sur la période 1980-1989. Ils estiment une matrice de transition avec 7 classes de PIB par tête. Les résultats indiquent une faible mobilité entre les classes et une concentration des fréquences dans les classes de PIB les plus faibles. López-Bazo *et al.* (1999) utilisent un échantillon de 143 régions au niveau NUTS II pour tous les pays sauf pour le Royaume-Uni. La période étudiée est 1980-1992. Ils estiment une matrice de transition avec 5 classes de PIB par tête. Les résultats obtenus par ces auteurs sont très similaires à ceux

obtenus par Neven et Gouyette (1995) : il y a une très grande persistance des régions dans chaque classe de PIB et la distribution de long terme semble indiquer que les régions les plus pauvres ne convergent pas vers le même niveau de PIB par tête que les autres régions de l'échantillon. Enfin, Maurseth (2001) utilise également un échantillon de 143 régions au niveau NUTS II (sauf pour le Royaume-Uni) sur la période 1980-1994. Il estime une matrice de transition avec 4 classes de PIB par tête et trouve qu'il existe davantage de mobilité pour les régions les plus riches de l'échantillon.

D'autres études ont été effectuées avec d'autres bases de données que la base utilisée dans le cadre de cette thèse. En conséquence, les résultats trouvés dans ces différents articles ne peuvent pas être directement comparés aux nôtres.

Par exemple, Fingleton (1997, 1999) utilise la base de données de Cambridge sur un échantillon de 169 régions sur la période 1975-1993 dans l'article de 1997 et sur un échantillon de 178 régions sur la période 1975-1995 dans l'article de 1999. Il se place dans le cadre particulier des modèles log-linéaires (Bishop *et al.*, 1975 ; Amemiya, 1985). Ses résultats semblent indiquer la persistance d'importantes disparités de revenus entre les régions européennes qui convergent vers une distribution limite multimodale.

Magrini (1999) et Cheshire et Magrini (2000) travaillent sur les régions urbaines fonctionnelles (FUR). Cheshire et Magrini (2000) utilisent les FUR sur les 12 pays membres de l'Union Européenne sur la période 1978-1994. L'analyse des distributions ergodiques indique que 6 régions urbaines fonctionnelles les plus riches – Hamburg, Düsseldorf, Munich, Stuttgart, Paris et Francfort – divergent des autres régions de l'échantillon. Ce résultat corrobore ceux trouvés par Magrini (1999) sur la période 1979-1990.

513.2 Estimation des matrices de transition

Pour modéliser le processus de croissance des régions européennes, nous devons tout d'abord discrétiser la distribution des PIB par tête relatifs à la moyenne européenne. Cette discrétisation est problématique et deux méthodes s'affrontent. Soit on applique une forme arbitraire de discrétisation, soit on estime des noyaux stochastiques continus à l'aide de méthodes non-paramétriques. Comme le soulignent Quah (1996a), Magrini (1999) et Bulli (2001), une discrétisation arbitraire de la distribution sous-jacente (qui est continue) peut conduire à d'importantes distorsions. Pour cette raison, Quah (1996c, 1997) préfère employer une estimation non paramétrique des matrices de transition conduisant à la représentation de graphiques en 3 dimensions. Cette approche réduit en revanche les caractéristiques des processus de croissance et de convergence qui peuvent être analysées, comme par exemple les distributions de long terme et les indices de mobilité.

Compte tenu des avantages et inconvénients propres à chaque méthode, nous préférons l'approche discrétionnaire : nous avons choisi de discrétiser les distributions des PIB par tête

de telle sorte que les classes initiales comportent un nombre similaire d'individus, comme le suggèrent Quah (1993a), López-Bazo *et al.* (1999) ou Kawagoe (1999).

Pour la distribution des PIB par tête relatifs en Ecus, nous avons classé les régions en cinq classes différentes : 1/ moins de 65% de la moyenne européenne 2/ entre 65% et 95% de la moyenne européenne 3/ entre 95% et 110% de la moyenne européenne 4/ entre 110% et 125% de la moyenne européenne 5/ plus de 125% de la moyenne européenne.

Pour la distribution des PIB par tête relatifs en SPA, les classes sont définies de la façon suivante : 1/ moins de 75% de la moyenne européenne 2/ entre 75% et 95% de la moyenne européenne 3/ entre 95% et 105% de la moyenne européenne 4/ entre 105% et 120% de la moyenne européenne 5/ plus de 120% de la moyenne européenne.

Notons que des chaînes de Markov comportant d'autres seuils de classes et d'autres nombres de classes ont été estimées mais les principaux résultats obtenus dans ce chapitre ne sont pas remis en cause par une discrétisation différente.

Compte tenu de ces différentes discrétisations, les deux matrices de transition peuvent être estimées à l'aide de la méthode du maximum de vraisemblance.

Soit p_{ij} la probabilité d'être dans la classe i à la période $t - 1$ et dans la classe j à période t . Si les probabilités de transition sont stationnaires, Anderson et Goodman (1957, p. 92) et Basawa et Prakasa Rao (1980, p. 54) montrent que l'estimateur de p_{ij} par le maximum de vraisemblance est défini de la façon suivante ⁷² :

$$\hat{p}_{ij} = \frac{n_{ij}}{n_i} \quad [5.28]$$

où n_{ij} est le nombre total de régions qui passent de la classe i à la période $t - 1$ à la classe j à la période t et n_i est le nombre total de régions dans la classe i sur les T transitions, soit : $n_i = \sum_j n_{ij}$. Dans notre échantillon, il y a 16 années et $T = 15$ transitions annuelles.

Le tableau 5.1a donne la matrice de transition entre 1980 et 1995. Elle contient les estimations par le maximum de vraisemblance des transitions pour la distribution des PIB par tête relatifs à la moyenne européenne en Ecus. Le tableau 5.1b représente la même matrice de transition entre 1980 et 1995 pour la distribution des PIB par tête relatifs à la moyenne européenne en SPA.

⁷² Il s'agit en fait de la fréquence empirique.

		Conditionnement européen					Nombre d'observations
		1	2	3	4	5	
<i>t_i</i>	<i>t_{i+1}</i>	<65%	<95%	<110%	<125%	>125%	
		Conditionnement européen	1	0,959	0,041	0,000	0,000
2	0,047		0,865	0,087	0,000	0,000	401
3	0,000		0,095	0,822	0,083	0,000	433
4	0,000		0,000	0,123	0,796	0,081	382
5	0,000		0,000	0,000	0,073	0,927	438

Tableau 5.1a : Matrice de transition ;
PIB par tête en Ecus ; conditionnement européen

		Conditionnement européen					Nombre d'observations
		1	2	3	4	5	
<i>t_i</i>	<i>t_{i+1}</i>	<75%	<95%	<105%	<120%	>120%	
		Conditionnement européen	1	0,962	0,038	0,000	0,000
2	0,042		0,884	0,074	0,000	0,000	431
3	0,000		0,079	0,858	0,063	0,000	416
4	0,000		0,002	0,070	0,874	0,054	445
5	0,000		0,000	0,000	0,067	0,933	387

Tableau 5.1b : Matrice de transition ;
PIB par tête en SPA ; conditionnement européen

Considérons le tableau 5.1a pour la distribution des PIB par tête relatifs à la moyenne européenne en Ecus. Par exemple, durant les 15 transitions, il y a eu 416 occurrences d'une région qui avait un PIB par tête inférieur à 65% de la moyenne européenne (première ligne de la matrice). La majorité de ces régions (95.9%) sont restées dans cette classe à la fin de l'année (cellule (1,1) de la matrice) alors que 4.1% sont montées dans la classe supérieure des régions possédant un PIB par tête entre 65% et 95% de la moyenne européenne (cellule (1,2) de la matrice).

Ces matrices de transition sont des matrices régulières. En effet, il suffit d'élever la matrice de transition des PIB par tête en Ecus à la puissance 4 pour obtenir une matrice dont tous les éléments sont non-nuls :

$$M_{Ecus}^4 = \begin{bmatrix} 0,856 & 0,126 & 0,017 & 0,001 & 2,4 \cdot 10^{-5} \\ 0,146 & 0,606 & 0,216 & 0,030 & 0,002 \\ 0,021 & 0,234 & 0,533 & 0,183 & 0,029 \\ 0,002 & 0,048 & 0,270 & 0,466 & 0,213 \\ 4 \cdot 10^{-5} & 0,003 & 0,039 & 0,192 & 0,766 \end{bmatrix} \quad [5.29]$$

De même, il suffit d'élever la matrice de transition des PIB par tête en SPA à la puissance 4 pour obtenir une matrice dont tous les éléments sont non-nuls :

$$M_{SPA}^4 = \begin{bmatrix} 0,863 & 0,122 & 0,014 & 6.10^{-4} & 1.10^{-5} \\ 0,133 & 0,646 & 0,199 & 0,021 & 8.10^{-4} \\ 0,016 & 0,214 & 0,588 & 0,165 & 0,016 \\ 0,001 & 0,031 & 0,185 & 0,621 & 0,161 \\ 4.10^{-5} & 0,002 & 0,022 & 0,200 & 0,775 \end{bmatrix} \quad [5.30]$$

Nous pourrions donc utiliser cette propriété fondamentale de régularité des matrices de transition estimées afin de déterminer des distributions ergodiques.

A l'aide de ces matrices de transition, nous pouvons étudier différentes caractéristiques de la dynamique de transition des PIB régionaux. Cependant, avant de présenter nos résultats, nous étudions le problème de la stabilité du processus entre deux sous-périodes.

513.3 Test de stabilité

López-Bazo *et al.* (1999) et Neven et Gouyette (1995) estiment des matrices de transition pour deux sous-périodes en raison des changements dans le processus de convergence détectés dans d'autres travaux. Les sous-périodes choisies par López-Bazo *et al.* sont : 1980-1985 et 1985-1992 alors que Neven et Gouyette choisissent les sous-périodes 1980-1985 et 1985-1989.

Ici, nous testons si les deux matrices de transition calculées pour les périodes 1980-1989 (sous-échantillon 1) et 1989-1995 (sous-échantillon 2) diffèrent l'une de l'autre et de la matrice calculée pour la période entière. Nous choisissons 1989 comme année charnière car cette date correspond à la réforme des fonds structurels en Europe. Le tableau 5.2a présente les deux matrices de transition pour la distribution des PIB par tête en Ecus et le tableau 5.2b présente ces deux matrices pour la distribution des PIB par tête en SPA.

Soit $\hat{p}_{ij}(l)$ la probabilité de se trouver dans la classe i à la période $t - 1$ et dans la classe j à la période t dans le sous-échantillon $l = 1, 2$. Notre but est de tester si les probabilités de transition des deux sous-périodes sont égales. En d'autres termes, nous considérons l'hypothèse nulle :

$$p_{ij}(1) = p_{ij}(2) = p_{ij} \quad \forall i = 1, \dots, 5 \quad j = 1, \dots, 5 \quad [5.31]$$

		$t_i \backslash t_{i+1}$	Conditionnement européen					Nombre d'observations
			1 <65%	2 <95%	3 <110%	4 <125%	5 >125%	
1980-1989	Conditionnement européen	1	0,947	0,053	0,000	0,000	0,000	262
		2	0,034	0,879	0,087	0,000	0,000	206
		3	0,000	0,084	0,816	0,100	0,000	261
		4	0,000	0,000	0,120	0,814	0,066	242
		5	0,000	0,000	0,000	0,081	0,919	271
1989-1995	Conditionnement européen	1	0,981	0,019	0,000	0,000	0,000	154
		2	0,062	0,851	0,087	0,000	0,000	195
		3	0,000	0,110	0,831	0,058	0,000	172
		4	0,000	0,000	0,129	0,764	0,107	140
		5	0,000	0,000	0,000	0,060	0,940	167

Tableau 5.2a : Matrices de transition par sous-période ; PIB par tête en Ecus ; conditionnement européen

		$t_i \backslash t_{i+1}$	Conditionnement européen					Nombre d'observations
			1 <75%	2 <95%	3 <105%	4 <120%	5 >120%	
1980-1989	Conditionnement européen	1	0,970	0,030	0,000	0,000	0,000	233
		2	0,036	0,873	0,092	0,000	0,000	251
		3	0,000	0,069	0,879	0,053	0,000	247
		4	0,000	0,004	0,063	0,890	0,044	272
		5	0,000	0,000	0,000	0,075	0,925	239
1989-1995	Conditionnement européen	1	0,949	0,051	0,000	0,000	0,000	158
		2	0,050	0,900	0,050	0,000	0,000	180
		3	0,000	0,095	0,828	0,077	0,000	169
		4	0,000	0,000	0,081	0,850	0,069	173
		5	0,000	0,000	0,000	0,054	0,946	148

Tableau 5.2b : Matrices de transition par sous-période ; PIB par tête en SPA ; conditionnement européen

Sur la base des différentes estimations des matrices de transition, Anderson et Goodman (1957) et Kullback *et al.* (1962) proposent un test du ratio de vraisemblance. Cette statistique s'écrit de la façon suivante :

$$Q_1 = -2 \log \left\{ \prod_{l=1}^2 \prod_{i=1}^K \prod_{j=1}^K \left[\frac{\hat{p}_{ij}}{\hat{p}_{ij}(l)} \right]^{n_{ij}(l)} \right\} \quad [5.32]$$

où K est le nombre de classes dans la distribution : $K = 5$, \hat{p}_{ij} est l'estimation de la probabilité de transition pour l'ensemble des périodes (équation [5.28]), $\hat{p}_{ij}(l)$ la probabilité de se trouver dans la classe i à la période $t - 1$ et dans la classe j à la période t dans le sous-échantillon $l = 1, 2$. Enfin, $n_{ij}(l)$ est le nombre correspondant de régions.

La statistique Q_1 converge asymptotiquement vers une loi du χ^2 à $K(K-1)$ degrés de liberté. Cependant, compte tenu de la présence d'éléments nuls dans la matrice de transition pour l'ensemble de la période et dans les deux sous-matrices, il faut corriger le nombre de degrés de liberté. Les éléments nuls dans ces différentes matrices ne doivent pas être pris en compte. Pour la distribution des PIB par tête en Ecus, le nombre de degrés de liberté se réduit ainsi à 8 alors qu'en SPA, le nombre de degrés de liberté se réduit à 9.

Nous trouvons les résultats suivants : $Q_1 = 10,722$ pour les PIB par tête en Ecus et $Q_1 = 9,845$ pour les PIB par tête en SPA. Avec 8 et 9 degrés de liberté, les probabilités critiques sont respectivement de $p = 0,218$ et $p = 0,363$. Par conséquent, on ne peut pas rejeter l'hypothèse que les probabilités de transition pour les deux sous-périodes et la période totale sont égales. Dans le reste de ce chapitre, nous analysons donc le processus de convergence pour la période totale.

513.4 Caractéristiques de la dynamique de transition

Nous pouvons tirer plusieurs enseignements de l'analyse des matrices de transition estimées dans le paragraphe précédent. Ainsi, la diagonale principale des matrices de transition permet de caractériser la persistance des régions dans les différentes classes de PIB par tête (paragraphe 513.41). La matrice des temps moyens de premier passage est utilisée pour analyser le temps mis pour passer d'une classe à n'importe quelle autre classe de la matrice de transition (paragraphe 513.42). La détermination du type de convergence est effectuée à l'aide de l'étude des distributions ergodiques (paragraphe 513.43). Enfin, les indices de mobilité permettent de synthétiser les informations fournies par les matrices de transition (paragraphe 513.44).

513.41 Une forte persistance des situations régionales

Considérons tout d'abord la matrice de transition estimée avec la distribution des PIB par tête relatifs à la moyenne européenne en Ecus (tableau 5.1a). On remarque que les probabilités de transition sur la diagonale principale sont relativement élevées. Ainsi, si une région se trouve dans une certaine classe, la probabilité d'être dans cette classe l'année suivante est au moins égale à 79,6% (cellule (4,4) de la matrice de transition). Pour la distribution des PIB par tête en SPA, cette probabilité est au moins égale à 87,4%.

Dans les deux cas, les éléments de la diagonale principale dominant, surtout parmi les classes extrêmes. Ces résultats indiquent que les régions les plus pauvres et les plus riches ne modifient pas leurs standings relatifs dans le temps. En termes des fonctions de densité, ces résultats indiquent que les queues de distribution de l'année finale sont composées des mêmes régions que celles de l'année initiale. En outre, on ne détecte pas de changements spectaculaires d'année en année. En effet, des probabilités strictement positives ne sont observées que pour les cellules autour de la diagonale. Seule une très faible proportion des régions (0,2%) qui se trouvaient dans la 4^{ème} classe de PIB au début d'une année a rétrogradé de deux classes à la fin de l'année (cellule (4,2) de la matrice de transition pour les PIB par tête en SPA).

513.42 Une fluidité des transitions assez faible

Pour analyser la vitesse à laquelle les régions se déplacent dans la distribution, nous considérons la matrice des temps moyens de premier passage, notée M_p , où chaque élément $M_{p,ij}$ indique le temps espéré pour une région de passer de la classe i à la classe j pour la première fois.

Nous avons montré dans le paragraphe 513.2 (p. 377) que les matrices de transition estimées pour notre échantillon de régions européennes sont régulières. Dans ce cas, Kemeny et Snell (1976, chap. 4, p. 79) démontrent que la matrice des temps moyens de premier passage est définie de la façon suivante :

$$M_p = (I_K - Z + SS'Z_{dg})D \quad [5.33]$$

où I_K est la matrice identité d'ordre K , Z est la matrice fondamentale : $Z = (I_K - M + M^*)^{-1}$, M^* est la matrice limite, S est le vecteur somme d'ordre K , Z_{dg} dérive de Z en posant les éléments hors-diagonale principale à 0. Enfin, D est une matrice diagonale de dimension (K, K) où les éléments de la diagonale principale sont égaux à $1/m_j^*$.

La matrice des temps moyens de premier passage permet d'examiner la fluidité des transitions. Elle est représentée dans le tableau 5.3a pour la distribution des PIB par tête relatifs en Ecus et dans le tableau 5.3b pour la distribution des PIB par tête relatifs en SPA. Notons que les éléments de la diagonale sont des temps moyens de premier retour, où le premier retour implique de rester dans la cellule une année ou de retourner la première fois dans cette cellule après l'avoir quittée la première année. Par exemple, dans le premier cas (tableau 5.3a), considérons une région se trouvant dans la classe de PIB la plus pauvre. Le temps moyen lui permettant de passer dans la classe immédiatement supérieure est de 24 ans environ. Ce temps passe à 49 ans pour se retrouver dans la troisième classe, 89 ans pour aboutir à la quatrième classe et près de 163 ans pour atteindre la classe de PIB la plus riche. Enfin, il ne faut que 4 ans pour retourner dans la première classe après l'avoir quittée la première année.

Lorsqu'on se concentre sur les éléments hors de la diagonale principale, on observe que les transitions sont globalement assez élevées puisque le temps moyen de premier passage le plus bas est de 13,7 années et celui le plus haut est de 162,7 années pour la distribution des PIB par tête en Ecus. Concernant la distribution des PIB par tête en SPA, ces temps moyens sont respectivement de 14,9 années et 194,5 années. Globalement, on observe que les mouvements à la hausse sont plus lents que les mouvements à la baisse, surtout pour les classes de PIB les plus élevés.

	1	2	3	4	5
1	(3,791)	24,471	49,211	89,416	162,693
2	68,309	(4,396)	24,741	64,946	138,222
3	93,934	25,624	(4,769)	40,205	113,482
4	111,090	42,780	17,155	(7,057)	73,277
5	124,778	56,468	30,843	13,687	(6,353)

Tableau 5.3a : Matrice des temps moyens de premier passage ; PIB par tête en Ecus

	1	2	3	4	5
1	(4,072)	26,067	54,198	105,902	194,553
2	80,089	(4,433)	28,131	79,835	168,487
3	111,670	31,581	(4,853)	51,704	140,356
4	135,753	55,664	25,949	(5,584)	88,652
5	150,638	70,548	40,833	14,884	(6,956)

Tableau 5.3b : Matrice des temps moyens de premier passage ; PIB par tête en SPA

513.43 L'équilibre régional de long terme : un désavantage pour les régions pauvres

Puisque les matrices de transition sont régulières pour les deux variables que nous considérons, elles convergent vers une matrice limite M^* de rang 1, dans laquelle toutes les lignes sont égales entre elles et égales à la distribution ergodique. Cette dernière peut s'interpréter comme l'équilibre régional de long terme de la distribution des PIB dans le système régional. Elle sert donc à détecter les différentes formes de convergence présentes entre les régions étudiées sur la période considérée.

Trois cas peuvent se présenter. Si la distribution de long terme se réduit à une seule classe, alors il y a convergence des régions vers un seul niveau de PIB par tête. Interprété en termes de fonctions de densité, ce cas de figure correspond à une distribution unimodale des PIB par tête. En revanche, une concentration des régions dans certaines classes, c'est-à-dire une distribution limite multimodale, peut être interprétée comme une tendance à la stratification vers différents clubs de convergence. Enfin, une dispersion de cette distribution entre les différentes classes, sans concentration particulière des régions dans une seule de ces classes, s'interprète comme de la divergence.

Nous avons déterminé les distributions ergodiques pour les PIB par tête en Ecus et en SPA en calculant les puissances successives des matrices de transition jusqu'à ce qu'il y ait convergence. Les résultats sont représentés dans le tableau 5.4.

Même si les résultats ne sont pas directement comparables puisque la discrétisation n'est pas exactement similaire, on note une certaine similitude dans les deux cas. Les distributions ergodiques sont davantage concentrées autour de la moyenne, ce qui indique une certaine convergence vers la moyenne. Cette concentration est en particulier due au fait que les classes de PIB les plus riches sont réduites par rapport à la distribution initiale.

Le tableau indique également une certaine *trappe à pauvreté*. En effet, les fréquences dans les deux classes les plus pauvres augmentent relativement à leurs distributions initiales. Ce résultat indique que les régions les plus pauvres de notre échantillon convergent vers un niveau de PIB par tête inférieur à celui des autres régions.

		<65%	<95%	<110%	<125%	>125%
Ecus	Dist. initiale	0,201	0,194	0,209	0,185	0,212
	Dist. ergodique	0,264	0,227	0,210	0,142	0,157
PPA		<75%	<95%	<105%	<120%	>120%
	Dist. initiale	0,189	0,208	0,201	0,215	0,187
	Dist. ergodique	0,246	0,226	0,206	0,179	0,144

Tableau 5.4 : Distribution ergodique et distribution initiale ; PIB par tête en Ecus et en SPA ; conditionnement européen

513.44 *Indice de mobilité en convergence et indice de mobilité à l'équilibre*

Les résultats précédents semblent indiquer une très faible mobilité inter-classes. Pour vérifier ce résultat, nous analysons dans ce paragraphe les indices de mobilité, qui visent à réduire l'information sur la mobilité d'une matrice à une seule statistique (Shorrocks, 1978 ; Geweke *et al.*, 1986 ; Maasoumi, 1998). Il existe deux catégories d'indices de mobilité, d'une part les indices mesurant la mobilité durant la convergence et d'autre part, les indices mesurant la mobilité une fois que la distribution ergodique est atteinte (tableau 5.5).

Les indices de mobilité en convergence mesurent la mobilité associée à la convergence vers l'équilibre. En d'autres termes, ils caractérisent de différentes façons la vitesse à laquelle l'état régulier est approché. Ces indices sont basés sur les valeurs propres de la matrice de transition. Pour les deux variables que nous considérons, PIB par tête en Ecus et en SPA, toutes les mesures sont très basses, ce qui implique une mobilité très faible entre les classes dans le temps. En outre, remarquons que la seconde valeur propre en valeur absolue est reliée à la mesure de la demi-vie. Cette dernière est définie comme le temps nécessaire à une région pour remplir la moitié de la distance qui la sépare de la distribution stationnaire (Shorrocks, 1978). Elle s'écrit de la façon suivante :

$$dm = -\frac{\log 2}{\log |\lambda_2|} \quad [5.34]$$

où λ_2 est la seconde valeur propre la plus élevée en valeur absolue. Cet indicateur tend vers l'infini (si la seconde valeur propre est égale à 1) auquel cas le système n'a pas de distribution stationnaire. Il tend vers 0 (si λ_2 tend vers 0) lorsque le système a déjà atteint son équilibre stationnaire. Les résultats indiquent que la demi-vie est égale à 24 ans pour la matrice de transition estimée avec les PIB par tête en Ecus et de 27 ans pour la matrice de transition estimée pour les PIB par tête en SPA.

Les indices de mobilité en convergence ne captent, cependant, qu'un seul aspect de la mobilité. En effet, il y aura toujours des mouvements d'une classe à une autre lorsque la distribution aura atteint son équilibre. Pour déterminer ce type de mobilité, nous étudions une seconde catégorie d'indices, les indices de mobilité à l'équilibre : la probabilité inconditionnelle de quitter la classe actuelle, l'indice de Bartholomew, qui est le nombre espéré de passages de classes d'une année sur l'autre à l'état stationnaire. Finalement, pour deux régions choisies au hasard, le temps moyen de premier passage est le nombre espéré d'années avant que le premier individu atteigne la classe du second individu. Pour les PIB par tête en Ecus et en SPA, les deux premières mesures sont très basses alors que la troisième mesure est supérieure à un demi-siècle, ce qui indique que, même dans le long terme, les sauts entre les classes seront très rares.

Ainsi, les résultats obtenus grâce à l'étude des distributions ergodiques, combinés avec la très faible mobilité observée sur la période, impliquent que les régions les plus pauvres vont probablement rester pauvres, comparées à la moyenne européenne. Ces résultats confirment les résultats précédents et indiquent une polarisation dans la distribution des PIB par tête relatifs régionaux : il y a une faible convergence entre les régions avec le PIB par tête inférieur et près de la moyenne européenne et les régions au-dessus de la moyenne européenne. Pourtant, les régions les plus pauvres ne semblent pas converger vers la moyenne.

	INDICE DE MOBILITE – CONVERGENCE			INDICES DE MOBILITE – EQUILIBRE		
	Prais ⁷³	Deuxième plus grande valeur propre	Déterminant	Bartholomew	Probabilité inconditionnelle de quitter la classe courante	Temps moyen de premier passage
	$\frac{K - tr(M)}{K - 1}$	$1 - \lambda_2 $	$1 - \det(M) $	$\sum_{i=1}^K F_i^* \sum_{j=1}^K M_{ij} i - j $	$\frac{K}{K - 1} \sum_{i=1}^K F_i^* (1 - M_{ii})$	$F^{*'} M_p F^*$
Ecus	0,158	0,029	0,515	1,701	0,149	55
SPA	0,122	0,024	0,417	1,129	0,129	65

Notes : M est la matrice de transition associée à la distribution ergodique F^* , K est le nombre de classes, λ_2 est la deuxième valeur propre la plus élevée en valeur absolue, M_p est la matrice des temps moyens de premier passage.

Sources : Shorrocks (1978), Geweke *et al.* (1986), Dardanoni (1993), Quah (1996d), Maasoumi (1998).

Tableau 5.5 : Indices de mobilité ; PIB par tête en Ecus et en SPA

Conclusion de section

Dans cette section, nous avons analysé le processus de convergence entre régions européennes en appliquant la méthodologie de Quah (1993a, 1993b, 1996a) qui consiste à étudier l'évolution de la distribution des PIB par tête et les changements de position relative des régions à l'intérieur de cette distribution.

⁷³ Un autre indice de mobilité parfois utilisé est $I = (K - \sum_i |\lambda_i|) / (n - 1)$, où $\lambda_i, i = 1 \dots K$, sont les K valeurs propres de M . Dans notre cas, il est égal à l'indice de Prais car toutes les valeurs propres sont réelles et non-négatives.

L'analyse de la forme des fonctions de densité et de l'indice de polarisation ont fait apparaître une persistance des disparités entre les régions européennes sur la période étudiée. Pour étudier les trajectoires de développement de chaque région de notre échantillon, nous avons ensuite estimé une matrice de transition associée à une chaîne de Markov stationnaire d'ordre 1. Les résultats indiquent une très faible mobilité inter-classes et une forte persistance des régions à l'intérieur de chaque classe. Finalement, si les tendances observées se poursuivent, la distribution des PIB sera progressivement biaisée vers les régions relativement pauvres. Globalement, la situation est remarquablement stable et persistante : il n'y a pas de changements marquants pour la forme extérieure des distributions, ni de mobilité intra-distribution importante.

Tous ces résultats font abstraction de la dimension spatiale des données utilisées. Nous étudions donc la façon dont les résultats précédents sont modifiés et affinés par la prise en compte de cette dimension dans la prochaine section.

SECTION 52 :

L'INTEGRATION DE LA DIMENSION SPATIALE DANS LES PROCESSUS DE MARKOV

Dans le cadre de cette thèse, nous avons souvent souligné le fait que la dimension spatiale des données devait être prise en compte pour permettre une meilleure appréhension des phénomènes de croissance et de convergence. Plusieurs travaux, portant sur l'analyse de l'évolution des distributions de PIB par tête pour les régions européennes ou pour d'autres échantillons de régions, ont également mis en évidence la nécessité de tenir compte de l'autocorrélation spatiale dans l'analyse des processus de croissance et de convergence (Armstrong, 1995b ; Fingleton, 1999 ; López-Bazo *et al.*, 1999 ; Rey et Montouri, 1999).

Dans le cadre particulier de l'analyse des dynamiques transitionnelles des PIB des régions par les chaînes de Markov, seules quelques études récentes ont intégré la dimension spatiale des données. Ainsi, Quah (1996b) utilise le « conditionnement régional » des séries de PIB par tête pour apprécier dans quelle mesure les variations des PIB dans l'espace, en un point donné du temps, peuvent être expliquées par le PIB des régions voisines. Rey (2001) construit des chaînes de Markov spatiales permettant d'évaluer comment les trajectoires de croissance des régions au sein de la distribution des PIB par tête sont influencées par la richesse ou la pauvreté des régions situées dans leur environnement géographique.

Dans cette section, nous appliquons ces méthodes aux données européennes. Conformément à la méthodologie suivie dans les chapitres précédents, la procédure adoptée est la suivante : détecter la présence d'autocorrélation spatiale, présenter et adopter les procédures permettant d'intégrer l'autocorrélation spatiale, et tester la significativité de cette dernière. Plus précisément, dans un premier temps, nous montrons que l'autocorrélation spatiale joue effectivement un rôle important dans les transitions des régions à travers les classes de PIB précédemment définies dans la section 51 (paragraphe 521). Dans un deuxième temps, nous « conditionnons » notre série des PIB par tête relatifs pour évaluer la similarité des PIB par tête entre une région et ses voisines (paragraphe 522). Ce conditionnement spatial, proposé par Quah (1996b), aboutit cependant à la perte de la dimension temporelle pour ne conserver que la dimension spatiale. Dans un troisième temps (paragraphe 523), nous présentons donc les méthodes permettant de conserver à la fois une dimension spatiale et une dimension temporelle. En particulier, nous appliquons la méthode suggérée par Rey (2001) à notre échantillon en estimant un processus de Markov avec conditionnement spatial. La significativité de l'autocorrélation spatiale dans ce cadre particulier n'a pas, à notre connaissance, fait l'objet de test. Aussi, nous définissons un *test de stationnarité spatiale des probabilités de transition*, destiné à tester si l'environnement géographique compte effectivement dans les possibilités de transitions des régions européennes d'une catégorie de PIB à une autre.

521 Les schémas géographiques des transitions régionales

Dans le chapitre 3, nous avons montré que les données de PIB par tête en Ecus et en SPA sont caractérisées par une très forte autocorrélation spatiale sur toute la période. En d'autres termes, il existe des concentrations spatiales persistantes de régions associées à des PIB par tête élevés ou faibles sur la période 1980-1995.

Pour illustrer encore l'importance potentielle des facteurs de localisation dans les schémas de convergence étudiés ici, nous examinons maintenant dans quelle mesure les régions qui ont subi une transition à la hausse ou à la baisse sont géographiquement concentrées. Ceci est étudié sous deux aspects. D'une part, nous évaluons l'ampleur de l'autocorrélation spatiale dans les régions qui ont subi une transition à la hausse ou à la baisse (paragraphe 521.1). D'autre part, nous montrons que les régions, dont les classements dans la distribution des PIB par tête ont fortement évolué entre 1980 et 1995, sont également géographiquement concentrées (paragraphe 521.2).

521.1 Autocorrélation spatiale et transition entre classes de PIB

Nous étudions tout d'abord le niveau d'autocorrélation spatiale dans les transitions des régions pour déterminer si la transition à la hausse ou à la baisse d'une région entre les classes de PIB par tête est reliée à celle de ses voisines géographiques. Considérons les figures 5.5a et 5.5b sur lesquelles sont représentées les transitions à la hausse et à la baisse des régions, respectivement pour les PIB en Ecus et les PIB en SPA.

Les transitions à la hausse (12 régions pour les PIB en Ecus et 20 régions pour les PIB en SPA) se situent essentiellement en Allemagne et aux Pays-Bas alors que les transitions à la baisse (28 pour les PIB en Ecus et en SPA) concernent principalement des régions françaises et anglaises. Soulignons d'ailleurs que ces régions correspondent à des concentrations spatiales ayant respectivement de forts et de faibles taux de croissance que nous avons détectées dans le troisième chapitre (cf. paragraphe 331 du chapitre 3, p. 245).

Ces résultats semblent indiquer un certain schéma spatial associé aux transitions des régions dans les classes de PIB par tête. Cette impression visuelle doit cependant être confirmée par un test formel d'autocorrélation spatiale.

Formellement, nous reprenons les matrices de poids déjà utilisées dans les deux chapitres précédents : les matrices des plus proches voisins et les matrices de distance. Nous rappelons ci-dessous les formes générales de ces matrices.

* Pour les matrices des plus proches voisins :

$$\begin{cases} w_{ij}^*(k) = 0 & \text{si } i = j \quad \forall k \\ w_{ij}^*(k) = 1 & \text{si } d_{ij} \leq d_i(k) \\ w_{ij}^*(k) = 0 & \text{si } d_{ij} > d_i(k) \end{cases} \quad k = 10, 15 \text{ et } 20 \quad [5.35]$$

où d_{ij} est la distance sphérique entre les centroïdes des régions i et j , $d_i(k)$ est la valeur-seuil définie pour chaque région, de telle sorte que chaque région a exactement k voisins.

* Pour les matrices basées sur les distances inverses :

$$\begin{cases} w_{ij}^*(k) = 0 & \text{si } i = j \quad \forall k \\ w_{ij}^*(k) = 1/d_{ij}^2 & \text{si } d_{ij} \leq D(k) \\ w_{ij}^*(k) = 0 & \text{si } d_{ij} > D(k) \end{cases} \quad k = 1, \dots, 3 \quad [5.36]$$

où $D(1) = Q_1$, $D(2) = Me$ et $D(3) = Q_3$, Q_1 , Me et Q_3 sont respectivement le premier quartile (516 km), la médiane (953 km) et le troisième quartile (1501 km) de la distribution des distances. Comme dans le chapitre 3, les résultats sont fournis pour la matrice des 10 plus proches voisins, $W(10)$, mais des indications sur la robustesse des résultats avec les autres matrices de poids sont fournies dans le texte et dans les annexes.

L'autocorrélation spatiale des transitions à la hausse des régions entre les classes de PIB est évaluée grâce au test NN , qui est le test d'autocorrélation spatiale pour variables qualitatives dichotomiques (Cliff et Ord, 1981). Ce test, présenté dans le chapitre 2 (paragraphe 213.11, p. 106), s'écrit, dans notre cas particulier de la façon suivante :

$$NN = \frac{1}{2} \sum_i \sum_j w_{ij}^* (\delta_i \delta_j) \quad [5.37]$$

où w_{ij}^* est un élément de la matrice de poids W non standardisée, $\delta_i = 1$ si la région i subit une transition à la hausse dans la distribution et $\delta_i = 0$ sinon. Ainsi, la statistique NN compte le nombre de liens entre deux régions voisines (le voisinage étant défini par la matrice de poids) subissant une transition à la hausse entre les classes de PIB.

De la même façon, pour tester l'autocorrélation spatiale entre transitions à la baisse dans la distribution, on définit $\delta_i = 1$ si la région i subit un mouvement à la baisse dans la distribution et $\delta_i = 0$ sinon.

L'inférence statistique est basée sur une permutation aléatoire des régions sur la carte. Les probabilités critiques associées à ce test pour toutes les matrices de poids et pour les deux variables (PIB par tête en Ecus et PIB par tête en SPA) sont présentées dans le tableau 5.6. Nous remarquons que l'hypothèse nulle d'absence d'autocorrélation spatiale est toujours fortement rejetée, qu'il s'agisse de transitions à la hausse ou à la baisse.

Pour affiner cette analyse, nous étudions maintenant si ces résultats sont confirmés lorsqu'on travaille, non plus sur les transitions entre classes de PIB, mais lorsque l'on prend en compte directement l'évolution des rangs occupés par les régions dans la distribution des PIB par tête.

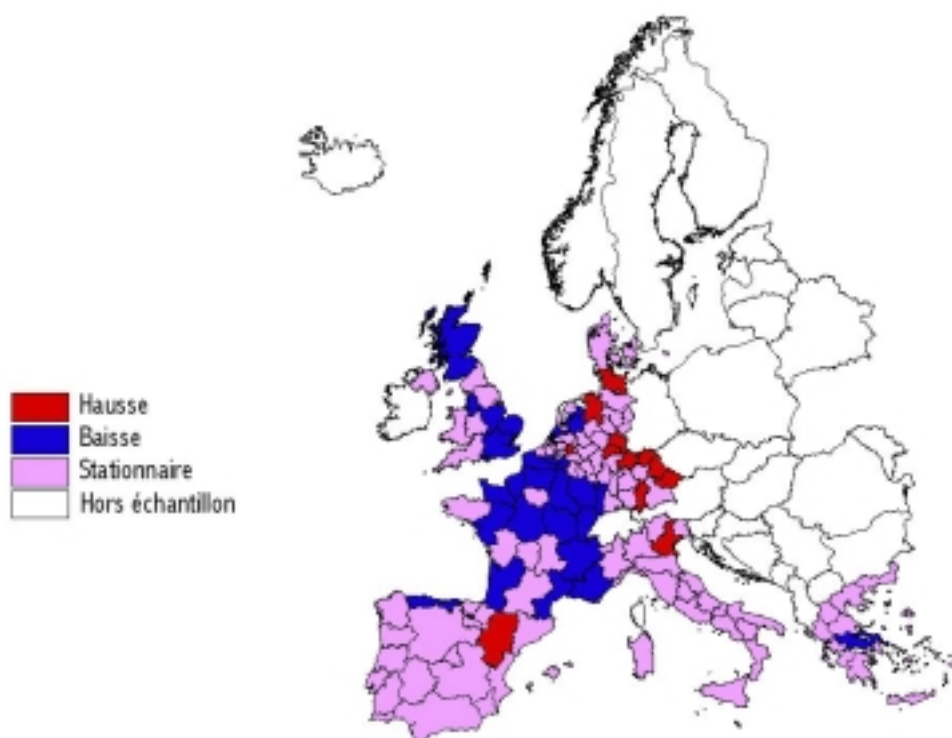


Figure 5.5a : Transitions de classe des régions entre 1980-1995 ;
PIB par tête en Ecus ; conditionnement européen

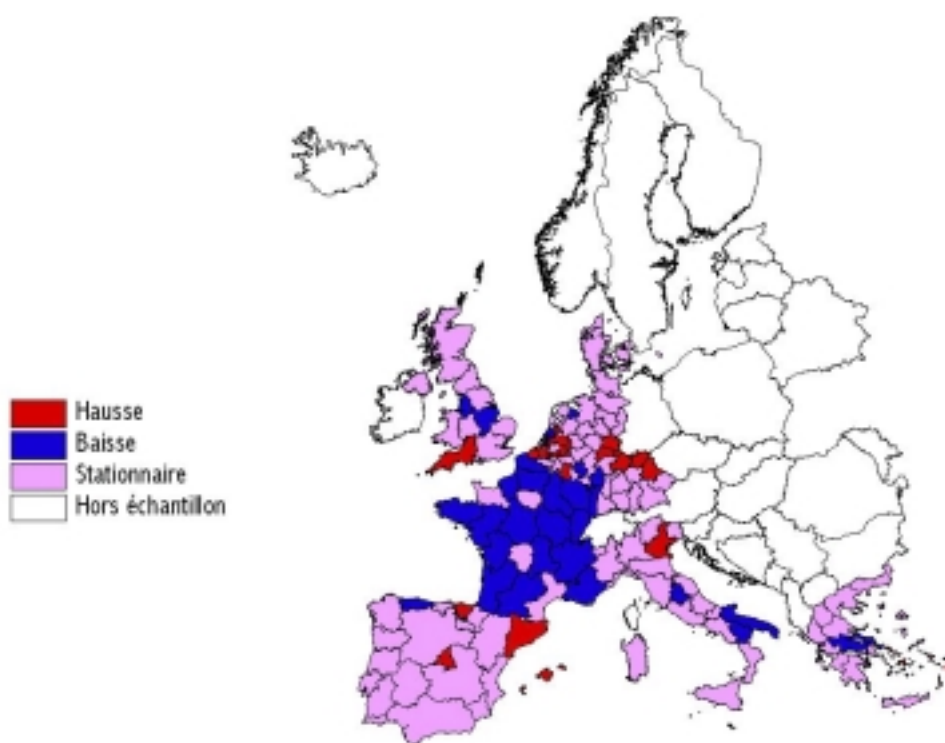


Figure 5.5b : Transitions de classe des régions entre 1980-1995 ;
PIB par tête en SPA ; conditionnement européen

		Hausse	Baisse
Ecus	W(10)	0,0008	0,0001
	W(15)	0,0005	0,0001
	W(20)	0,0007	0,0003
SPA	W(10)	0,0056	0,0003
	W(15)	0,0033	0,0002
	W(20)	0,0074	0,0003

Tableau 5.6 : Valeurs critiques pour les tests d'autocorrélation spatiale ; transitions entre classes de PIB ; PIB par tête en Ecus et en SPA

521.2 Autocorrélation spatiale et classement des régions

Dans les chaînes de Markov, la mobilité n'est visible que lorsqu'il existe des « sauts » entre les classes de PIB. Pour étudier plus largement la mobilité des régions dans la distribution des PIB, nous considérons ici les classements de toutes les régions entre 1980 et 1995.

Les figures 5.6a et 5.6b représentent les classements des régions en 1995 en fonction de leurs classements en 1980 (les codes des régions sont donnés dans l'annexe 3.1 du chapitre 3). Lorsqu'une région se trouve au-dessus de la ligne des 45°, sa position s'est dégradée dans la hiérarchie des PIB. En revanche, les régions se trouvant en dessous de cette ligne ont vu leurs classements s'améliorer. Plus une région dévie de la ligne des 45°, plus son classement dans la distribution des PIB s'est trouvé modifié entre 1980 et 1995.

Nous avons calculé l'écart-type de la série des changements de classement entre 1980 et 1995 (classement en 1995 moins le classement en 1980). Il est égal à 13,3 pour les PIB par tête en Ecus et à 16,8 pour les PIB par tête en SPA. Considérons les régions dont le classement entre 1980 et 1995 s'est modifié de plus d'un écart-type. Pour les données en Ecus, parmi les 17 régions qui ont subi une transition à la baisse, on retrouve à nouveau les 12 régions françaises. En revanche, les 17 régions dynamiques qui ont progressé dans la hiérarchie des régions sont surtout situées en Allemagne. Ce schéma est très similaire pour les données en SPA.

Le tableau 5.7 fournit les probabilités critiques lorsqu'on teste l'autocorrélation des changements de classement de ces régions. Comme précédemment, l'hypothèse nulle est fortement rejetée dans tous les cas de figure.

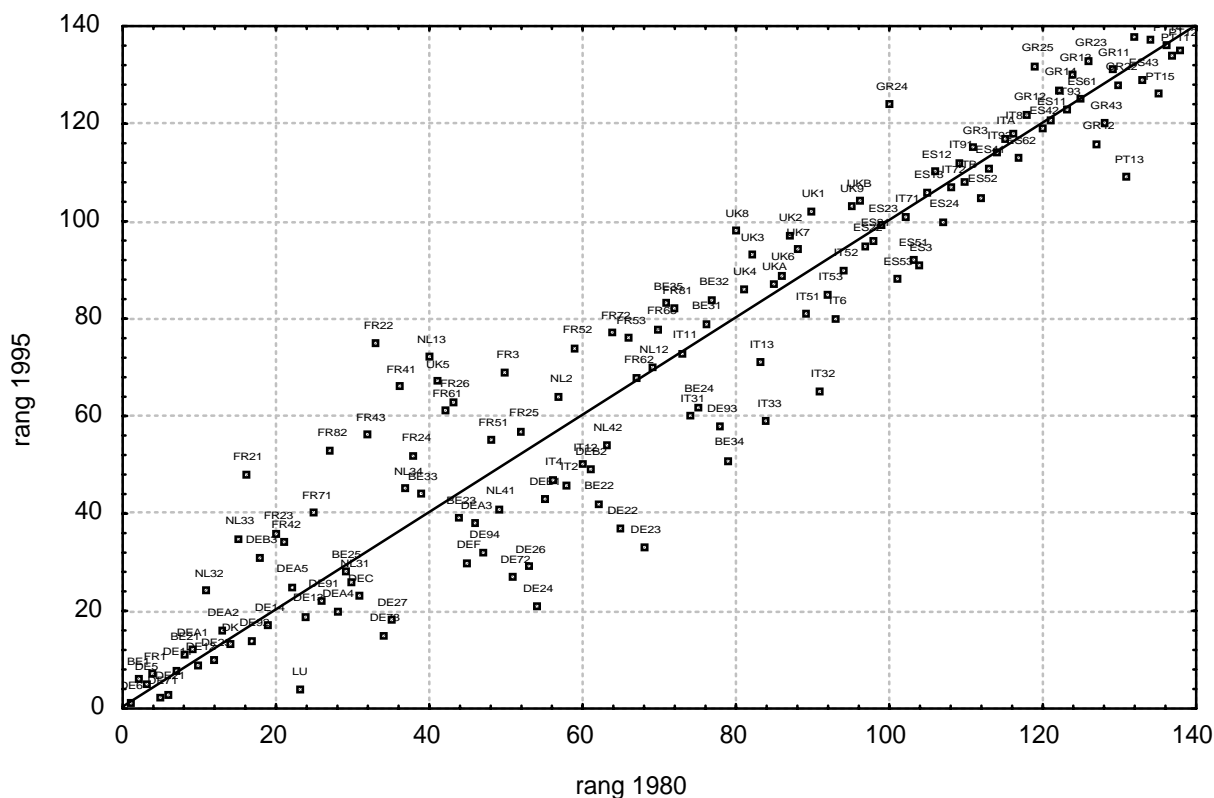


Figure 5.6a : Classement des régions en fonction du PIB par tête en 1980 et 1995 ; PIB par tête en Ecus

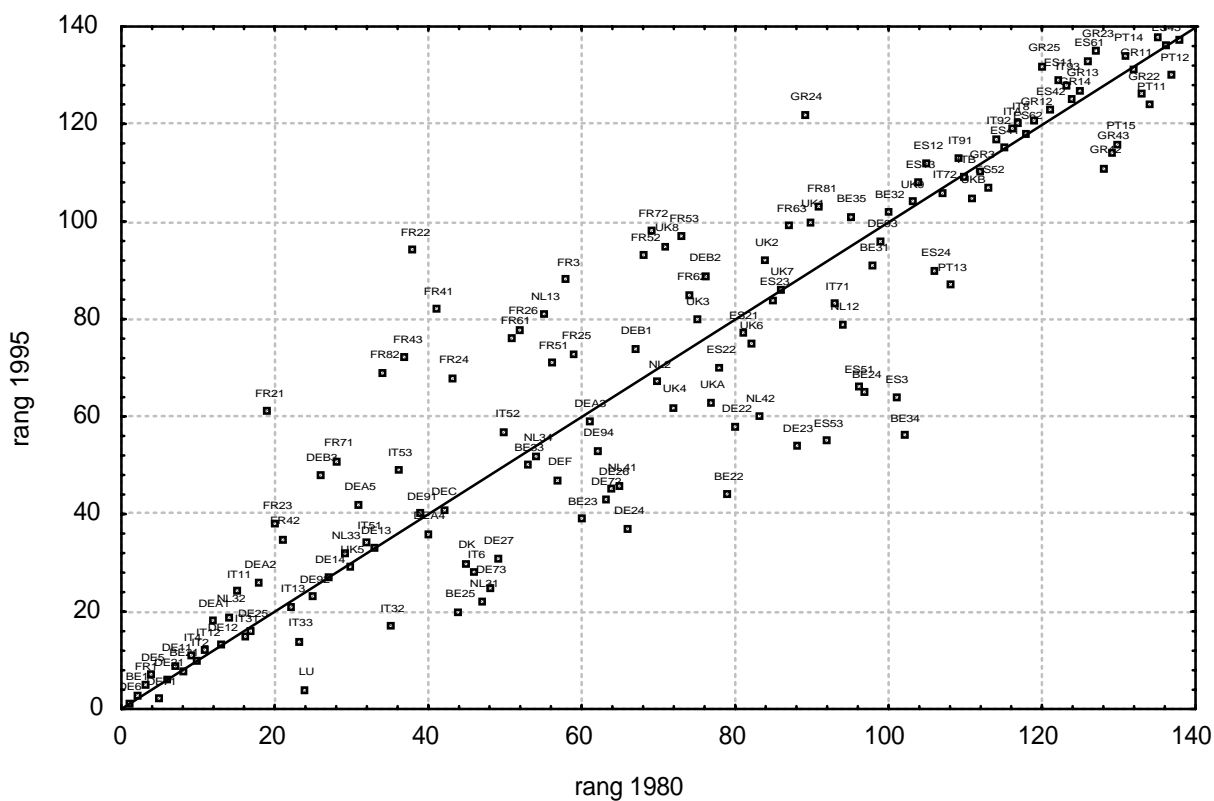


Figure 5.6b : Classement des régions en fonction du PIB par tête en 1980 et 1995 ;

PIB par tête en SPA

		Hausse	Baisse
Ecus	W(10)	0,0003	0,0016
	W(15)	0,0006	0,0013
	W(20)	0,0007	0,0006
SPA	W(10)	0,0050	0,0001
	W(15)	0,0030	0,0001
	W(20)	0,0032	0,0001

Tableau 5.7 : Valeurs critiques pour les tests d'autocorrélation spatiale ; classement des régions ; PIB par tête en Ecus et en SPA

L'ensemble des résultats présentés dans ce paragraphe 521 montre que traiter les régions et leurs positions relatives dans la distribution des PIB par tête comme si elles étaient géographiquement indépendantes n'est pas justifié. Il s'agit alors de montrer comment l'autocorrélation spatiale ainsi mise en évidence peut être explicitement prise en compte dans l'analyse de la dynamique de transition des PIB régionaux ⁷⁴.

522 Conditionnement spatial

Pour déterminer les facteurs permettant d'expliquer les caractéristiques des courbes de densité et celles de la matrice de transition exposées précédemment, Quah (1996b) suggère de « conditionner » la distribution des PIB régionaux. L'idée générale de la procédure de conditionnement est d'étudier la façon dont l'évolution du PIB d'une région a suivi celle d'un groupe de régions que l'on suppose agir d'une façon similaire. Quah (1996b) propose deux références possibles : soit les régions voisines (critère géographique), soit les régions appartenant au même pays (critère national). Les résultats de son étude indiquent que les facteurs géographiques semblent être plus importants que les facteurs nationaux pour expliquer la dynamique des inégalités de PIB parmi les régions européennes.

Nous adoptons à notre tour un critère de conditionnement géographique basé sur la notion de voisinage et nous construisons la série des *PIB par tête relatifs à la moyenne des*

⁷⁴ L'annexe 5.1 présente les résultats des différents tests d'autocorrélation spatiale pour les matrices de distance.

voisins. Cette série, référencée par la suite « conditionnement régional », est définie comme le ratio du PIB par tête d'une région à la moyenne des PIB de ses voisins. La notion de voisinage est introduite à l'aide des matrices de poids standardisées de la façon suivante :

$$w_{ij} = w_{ij}^* / \sum_j w_{ij}^* \quad [5.38]$$

Formellement, si y est le vecteur contenant les PIB par tête des régions, la moyenne pondérée des PIB des régions voisines est donnée par le vecteur de décalage spatial Wy . La série de conditionnement régional des PIB par tête est alors obtenue en divisant élément par élément y par Wy .

Quah (1996b, p. 955) suggère d'interpréter la série de PIB relatifs à la moyenne des voisins comme la partie qui n'est pas expliquée par les facteurs de localisation physique. Par conséquent, si la localisation physique explique tout, « ce qui reste » dans cette série disparaît ou est très petit. Au contraire, si la localisation physique n'explique rien, ce qui reste dans la série est très proche de la série initiale.

A l'aide de cette nouvelle série, nous analysons si l'environnement géographique des régions influence leurs positions dans la distribution des PIB par tête ainsi que l'évolution de cette position. Nous allons pour cela construire et estimer des courbes de densité conditionnelles (paragraphe 522.1) ainsi qu'une matrice de transition conditionnelle (paragraphe 522.2) ⁷⁵.

522.1 Conditionnement géographique et richesse des économies régionales

Les courbes de densité des distributions de PIB par tête relatifs à la moyenne des voisins sont représentées dans les figures 5.7a (en Ecus) et 5.7b (en SPA). Pour chaque graphique, nous avons calculé la fonction de densité pour l'année initiale 1980 et pour l'année finale 1995.

Il apparaît que, dans les deux cas, la majorité des densités de probabilité est centrée symétriquement autour de la moyenne. En comparant ces densités aux densités précédemment calculées sur les PIB relatifs à la moyenne européenne (figures 5.2a et 5.2b), on observe par ailleurs que le deuxième pic, qui se situait à environ 50% de la moyenne européenne, a disparu. En conséquence, la performance économique d'une région est bien expliquée par celle de ses voisines. Les exceptions sont les régions associées à des PIB par tête très élevés qui se trouvent dans la queue de droite des fonctions de densité. Ces régions semblent posséder une forte dynamique de croissance indépendante de celle de ses voisines. Ces résultats rejoignent ainsi, sur une base de données différente, les conclusions de Magrini (1999) et de Cheshire et Magrini (2000).

⁷⁵ Dans les paragraphes suivants, les résultats sont fournis pour la matrice des 10 plus proches voisins. L'ensemble des autres résultats pour la matrice de distance $W(1)$ est fourni dans l'annexe 5.2 pour la distribution des PIB par tête en Ecus et dans l'annexe 5.3 pour la distribution des PIB par tête en SPA.

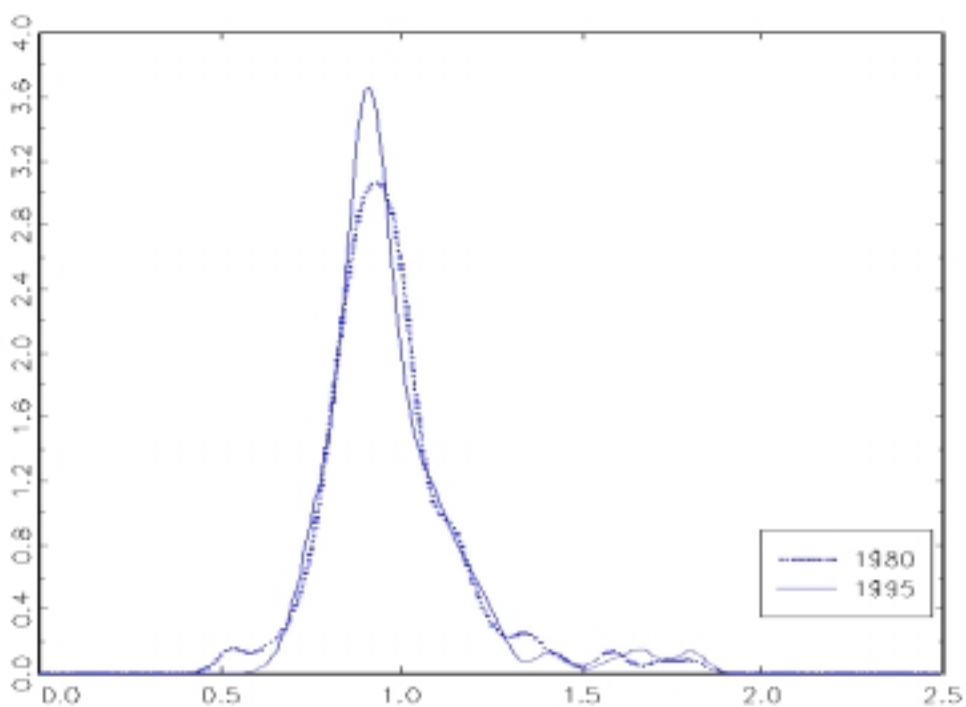


Figure 5.7a : Densités des distributions de PIB par tête mesuré en Ecus ; 1980 - 1995 ; conditionnement régional

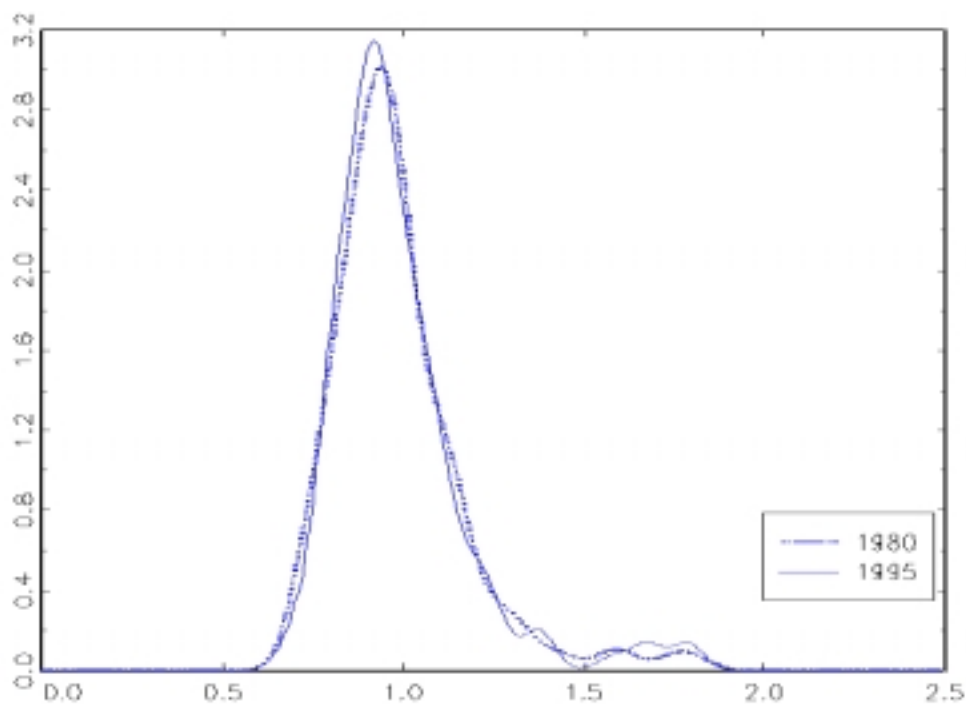


Figure 5.7b : Densités des distributions de PIB par tête mesuré en SPA ; 1980 - 1995 ; conditionnement régional

522.2 Conditionnement géographique et dynamique de transition

Les matrices de transition conditionnelles proposées par Quah (1996b) fournissent les transitions d'une région pour une année donnée entre deux distributions, d'une part, la distribution des PIB relatifs à la moyenne européenne et d'autre part, la distribution des PIB relatifs à la moyenne des voisins. Les estimations pour les deux séries de PIB sont données dans les tableaux 5.8a (en Ecus) et 5.8b (en SPA).

L'interprétation des éléments de ces nouvelles matrices de transition permet d'établir les transitions entre *deux distributions différentes au même point du temps*. Il s'agit ainsi d'une représentation différente de celle donnée par les matrices de transition précédemment estimées (tableaux 5.1a et 5.1b). En effet, dans ces dernières, nous avons établi les transitions entre *la même distribution à deux points du temps*. Dans le cas présent, Quah (1996b) qualifie les transitions estimées de transitions « interdistributionnelles » pour une année donnée. On se demande alors si les variations des PIB dans l'espace européen, en un point donné du temps, peuvent être expliquées par le conditionnement des PIB des régions voisines.

Formellement, on note p_{ik} la probabilité d'être à la fois dans la classe i de la distribution des PIB par tête relatifs à la moyenne européenne (distribution 1) au temps t et dans la classe k de la distribution des PIB par tête relatifs à la moyenne des voisins (distribution 2) au temps t . L'estimateur d'un élément p_{ik} d'une matrice de transition conditionnelle est défini de la façon suivante :

$$\hat{p}_{ik} = \frac{n_{ik}}{n_i} \quad [5.39]$$

où n_{ik} est le nombre total de régions qui sont dans la classe i de la distribution 1 à la période t et dans la classe k de la distribution 2 à la période t , n_i est le nombre total de régions dans la classe i sur les 16 années. : $n_i = \sum_k n_{ik}$.

Nous avons construit la série des PIB avec conditionnement régional en conservant les seuils des classes de la série des PIB avec conditionnement européen.

Considérons le tableau 5.8a représentant la matrice conditionnelle pour la distribution de PIB par tête en Ecus. 450 occurrences de régions apparaissent avec un PIB par tête inférieur à 65% de la moyenne européenne pour une année donnée. Seulement 9,1% de ces régions possèdent un PIB par tête inférieur à 65% de la moyenne de ses voisins pour cette même année. La majorité de ces régions (51,1%) possède un PIB par tête compris entre 65% et 95% de la moyenne de ses voisins.

Si le contexte régional ne comptait pas, chaque région pourrait être considérée comme une île indépendante de ses voisines. Dans ce cas, la matrice de transition devrait être égale à la matrice identité (Quah, 1996b, p. 956) :

« If conditioning explained nothing (...), these transition probability matrices should be the identity matrix: the distributions are invariant and, in addition, no intra-distribution movements occurs. »

En revanche, si le conditionnement spatial expliquait toutes les variations des PIB par tête régionaux, alors tous les éléments de la troisième colonne devraient être égaux à 1 (il s'agit de la colonne associée à la classe contenant la moyenne européenne).

		<i>ti</i>					Nombre d'observations
		Conditionnement régional					
<i>ti</i>		1	2	3	4	5	
		<65%	<95%	<110%	<125%	>125%	
Conditionnement européen	1	0,091	0,511	0,251	0,131	0,016	450
	2	0,000	0,641	0,277	0,049	0,033	426
	3	0,000	0,682	0,256	0,056	0,006	465
	4	0,000	0,461	0,443	0,073	0,023	397
	5	0,000	0,121	0,385	0,228	0,266	470

Tableau 5.8a : Matrice de transition conditionnelle ; PIB par tête en Ecus

		<i>ti</i>					Nombre d'observations
		Conditionnement régional					
<i>ti</i>		1	2	3	4	5	
		<75%	<95%	<105%	<120%	>120%	
Conditionnement européen	1	0,010	0,663	0,227	0,096	0,005	418
	2	0,000	0,731	0,177	0,032	0,060	464
	3	0,000	0,647	0,274	0,063	0,016	442
	4	0,000	0,353	0,577	0,068	0,002	470
	5	0,000	0,027	0,384	0,275	0,314	414

Tableau 5.8b : Matrice de transition conditionnelle ; PIB par tête en SPA

Pour notre échantillon, aucun de ces deux cas extrêmes n'est pertinent. En particulier, les éléments de la diagonale principale sont tous relativement faibles. Seule la cellule (2,2) est égale à 64,1% pour les PIB par tête en Ecus et à 73,1% pour les PIB par tête en SPA. Le conditionnement spatial explique donc une part importante de la variation des PIB régionaux. Ce résultat rejoint celui trouvé par Quah (1996b) sur les régions européennes. Les effets de

débordement géographiques apparaissent être un facteur important de la dynamique des inégalités régionales en Europe.

Le conditionnement spatial permet de capter le rôle de la dimension géographique dans la variation des PIB régionaux. Pourtant, rappelons que ces matrices de transition ne représentent pas des transitions dans le temps, mais seulement des « transitions » entre deux distributions de PIB relatifs pour une année donnée. Les matrices de Markov présentées dans le paragraphe suivant ont pour but de prendre en compte explicitement l'espace sans pour autant perdre la dynamique temporelle de l'évolution des PIB régionaux en Europe.

523 Dynamiques spatiales et temporelles des disparités régionales

Deux méthodes permettent d'étudier les effets de dépendance spatiale tout en conservant une information sur la dynamique temporelle des transitions. La première consiste à estimer une matrice de transition temporelle en utilisant exclusivement la distribution des PIB par tête relatifs à la moyenne des voisins (paragraphe 523.1). La seconde s'appuie sur la construction des *chaînes de Markov spatiales* (paragraphe 523.2).

523.1 Analyse des transitions temporelles avec conditionnement régional

Nous estimons une matrice de transition d'une année sur l'autre entre 1980 et 1995 dans laquelle le PIB par tête d'une région n'est plus relatif à la moyenne européenne mais est relatif à la moyenne des PIB par tête des régions voisines. La formule utilisée pour estimer les transitions de cette matrice est similaire à l'équation [5.28].

Les résultats de cette procédure sont présentés dans les tableaux 5.9a et 5.9b pour les PIB par tête en Ecus et en SPA. Ainsi, dans le tableau 5.9a, si l'on examine la première ligne de cette matrice, on remarque qu'il y a 41 occurrences de régions ayant un PIB par tête inférieur à 65% de la moyenne de ses voisins en début de période. Dans 73,2% des cas, le PIB de cette région était toujours inférieur à 65% de la moyenne de ses voisins à la période suivante. Au contraire, dans 26,8% des cas, cette région a vu son PIB augmenter pour aboutir dans la classe supérieure (65% et 95% de la moyenne de ses voisins) à la fin de l'année.

Si on compare ces tableaux aux tableaux 5.1a et 5.1b (calculés avec un conditionnement sur la moyenne européenne), on remarque que pour le même découpage de classes, les queues de distributions sont plus petites. Il y a ainsi 416 pauvres et 438 riches avec le conditionnement européen comparé à 41 pauvres et 148 riches avec le conditionnement géographique. En revanche, les classes moyennes sont beaucoup plus importantes que précédemment et concentrent quasiment toutes les régions. Ce dernier résultat suggère que globalement, les régions proches géographiquement évoluent de la même façon et ne se démarquent pas : on retrouve donc là un phénomène important d'autocorrélation spatiale positive entre les régions.

Les exceptions à ce phénomène général sont les régions qui restent beaucoup plus pauvres que leurs voisines tout au long de la période, c'est-à-dire les régions qui se trouvent dans la première cellule (case (1,1)) de la matrice de transition. Il s'agit de la région Ipeiros de 90 à 92, des régions portugaises Norte, Centro et Alentejo sur quasiment toute la période. Les autres exceptions sont les régions qui restent beaucoup plus riches que leurs voisines sur la période (case (5,5)). Il s'agit surtout des régions-capitales (Bruxelles, Ile de France, Madrid, Sterea Ellada) et des régions allemandes (Stuttgart, Oberbayern, Bremen, Hamburg, Darmstadt, Düsseldorf) sur toute la période. Pour les PIB par tête en SPA, les queues de classe sont plus importantes : 66 pauvres comparées à 208 riches. Pourtant, les classes médianes concentrent toujours la majorité des régions.

		<i>ti + 1</i>					Nombre d'observations
		Conditionnement régional					
<i>ti</i>		1	2	3	4	5	
		<65%	<95%	<110%	<125%	>125%	
Conditionnement régional	1	0,732	0,268	0,000	0,000	0,000	41
	2	0,007	0,930	0,063	0,000	0,000	981
	3	0,000	0,101	0,866	0,033	0,000	672
	4	0,000	0,000	0,101	0,851	0,048	228
	5	0,000	0,000	0,000	0,068	0,932	148

Tableau 5.9a : Matrice de transition ;
conditionnement régional ; PIB par tête en Ecus

		<i>ti + 1</i>					Nombre d'observations
		Conditionnement régional					
<i>ti</i>		1	2	3	4	5	
		<75%	<95%	<105%	<105%	>120%	
Conditionnement régional	1	0,758	0,242	0,000	0,000	0,000	66
	2	0,014	0,941	0,046	0,000	0,000	945
	3	0,000	0,081	0,878	0,041	0,000	558
	4	0,000	0,000	0,085	0,863	0,051	293
	5	0,000	0,000	0,000	0,067	0,933	208

Tableau 5.9b : Matrice de transition ;
conditionnement régional ; PIB par tête en SPA

A première vue, le conditionnement régional semble indiquer une convergence plus importante que le conditionnement européen. Pourtant, conditionner sur la moyenne des voisins fournit une information sur la convergence locale ou « intra-régionale », c'est-à-dire sur la façon dont les régions rattrapent leurs voisines géographiques. Par conséquent, cette méthode prend en compte explicitement le rôle de l'espace mais ne permet plus d'étudier la

position des régions dans la distribution en coupe transversale des PIB par tête. En d'autres termes, cette méthodologie ne permet pas d'analyser la convergence interrégionale.

523.2 La chaîne de Markov spatiale (Rey, 2001)

Pour étudier simultanément les dynamiques spatiale et temporelle, Rey (2001) propose de construire une *matrice de Markov spatialisée*. Il estime cette matrice sur les états américains. Ainsi, la matrice de Markov spatiale de Rey permet, d'une part, d'étudier le rôle de la localisation géographique dans la dynamique de la distribution des PIB dans le temps et d'autre part, d'analyser le processus de convergence interrégionale.

La procédure de construction de la matrice de transition spatiale est basée sur une décomposition « spatiale » de la matrice de transition traditionnelle qui permet d'obtenir des « probabilités de transitions spatiales » (paragraphe 523.21). La significativité des liens entre les probabilités d'une région transitant d'une classe de PIB à une autre et son environnement géographique, est testée en comparant les « probabilités spatiales » et les probabilités de transition traditionnelles à l'année initiale (paragraphe 523.22). Si les effets de dépendance à l'environnement géographique sont confirmés, il est alors possible d'apprécier dans quelle mesure les trajectoires des régions sont conditionnées par la richesse de leurs voisines géographiques (paragraphe 523.23).

523.21 Description de la matrice de Markov spatiale

Pour construire la matrice de Markov « spatiale », la matrice de Markov « traditionnelle » (tableaux 5.1a et 5.1b) est modifiée de telle sorte que les probabilités de transition dans le temps d'une région sont conditionnées par la classe comportant la moyenne des PIB des régions voisines au début de l'année. En d'autres termes, la matrice de transition spatiale est une matrice traditionnelle de dimension (K,K) décomposée en K sous-matrices conditionnelles de dimension (K,K) . Ici, $K = 5$ classes.

Formellement, si l'on considère la $k^{\text{ème}}$ de ces matrices conditionnelles, un élément $p_{ij}(k)$ de cette matrice représente la probabilité qu'une région dans la classe i à la période $t-1$, soit dans la classe j à la période suivante, sachant que la moyenne des PIB de ses régions voisines se trouvait dans la classe k à la période $t-1$. L'estimateur d'un élément $p_{ij}(k)$ d'une matrice de transition conditionnelle est défini de la façon suivante :

$$\hat{p}_{ij}(k) = \frac{n_{ij}(k)}{n_i(k)} \quad [5.40]$$

où $n_{ij}(k)$ est le nombre total de régions qui sont dans la classe i à la période $t-1$ et dans la classe j à la période suivante, sachant que la moyenne des PIB de leurs régions voisines se

trouvait dans la classe k à la période $t - 1$. Le terme $n_i(k)$ représente le nombre total de régions se trouvant dans la classe i , sachant que la moyenne des PIB de leurs régions voisines se trouvait dans la classe k à la période $t - 1$ sur les $T = 15$ transitions annuelles, soit :

$$n_i(k) = \sum_j n_{ij}(k).$$

La matrice de transition spatiale permet, d'une part, d'étudier les transitions temporelles des régions à travers les classes de PIB par tête et d'autre part, d'analyser la corrélation entre la direction de ces transitions et l'environnement géographique de chaque région. En d'autres termes, la matrice de transition spatiale fournit la probabilité pour une région de subir des transitions à la hausse ou à la baisse dans la distribution des PIB par tête relatifs, *conditionnellement* aux mouvements passés de ses voisines dans cette distribution.

Dans notre application avec 5 classes, la première classe contient les régions pauvres, la troisième classe contient les régions médianes et la classe finale contient les régions riches. Si $p_{35} > p_{35}(1)$, la probabilité d'une transition à la hausse d'une région médiane, sans tenir compte de ses régions voisines, est supérieure à la probabilité d'une transition à la hausse de cette région dont les voisines sont pauvres. De même, si $p_{13} < p_{13}(5)$, alors une région pauvre avec de riches voisines a une plus grande probabilité de connaître une transition à la hausse que les régions pauvres en général.

Grâce à cette matrice de transition spatiale, il est ainsi possible d'examiner l'influence, positive ou négative, des régions voisines sur les transitions d'une région. En d'autres termes, on étudie l'influence de l'autocorrélation spatiale : elle est reflétée par les différences qui existent entre les valeurs des transitions initiales (non conditionnées), calculées dans la première section (tableaux 5.1a et 5.1b), et les valeurs des différentes transitions conditionnelles.

Si le contexte régional ne comptait pas, les probabilités conditionnelles devraient être égales aux probabilités initiales :

$$p_{ij}(1) = p_{ij}(2) = \dots = p_{ij}(5) = p_{ij} \quad \forall i = 1, \dots, 5 \quad j = 1, \dots, 5 \quad [5.41]$$

Ces égalités font l'objet d'un test que nous définissons dans le paragraphe suivant.

523.22 Le test de stationnarité spatiale

Dans ce paragraphe, nous définissons un test pour l'hypothèse d'indépendance des probabilités de transition dans l'espace. En d'autres termes, notre but est de suggérer une procédure permettant de tester si les probabilités de transition spatiales sont statistiquement égales les unes aux autres et statistiquement égales à la matrice de transition spatiale.

Formellement, cela revient à tester les restrictions [5.41]. Sous l’hypothèse nulle, les probabilités de transition d’une région ne dépendent pas du décalage spatial et peuvent être estimées à l’aide de l’équation [5.28]. Sous l’hypothèse alternative, différentes probabilités de transition dépendant du décalage spatial doivent être estimées à l’aide de l’équation [5.40]. La statistique du ratio de vraisemblance peut donc être définie de la façon suivante :

$$Q_2 = -2 \log \left\{ \prod_{k=1}^K \prod_{i=1}^K \prod_{j=1}^K \left[\frac{\hat{p}_{ij}}{\hat{p}_{ij}(k)} \right]^{n_{ij}(k)} \right\} \quad [5.42]$$

où K est le nombre de cellules dans la distribution : $K = 5$, \hat{p}_{ij} est l’estimation stationnaire [5.28], $\hat{p}_{ij}(k)$ est la probabilité estimée qu’une région dans la classe i à la période $t - 1$ soit dans la classe j à la période t sachant que son décalage spatial est dans la classe k en $t - 1$, $n_{ij}(k)$ est le nombre correspondant de régions. Q_2 converge asymptotiquement vers une loi du chi-deux à $K(K-1)^2$ degrés de liberté.

523.23 Environnement géographique et convergence des régions européennes

Les matrices de transition spatiales estimées pour notre échantillon de régions européennes sont reproduites dans les tableaux 5.10a (en Ecus) et 5.10b (en SPA).

Considérons par exemple le tableau 5.10a. On observe que les régions les plus riches sont négativement affectées lorsqu’elles sont entourées de régions plus pauvres. En effet, la probabilité de descendre d’une classe augmente lorsque le PIB par tête des régions voisines diminue. En moyenne, les régions les plus riches descendent d’un cran avec une probabilité de 7,3% (cellule (5,4) dans le tableau 5.1a). En revanche, si ces régions sont entourées d’autres régions riches (classe 5), la probabilité de descendre est de 5,9% mais si les régions voisines sont moins riches (classe 4), la probabilité passe à 10,1%. De même, il apparaît que les régions les plus pauvres sont affectées négativement lorsqu’elles sont situées initialement à côté d’autres régions pauvres. Ainsi, la probabilité qu’une région pauvre monte d’une classe dans la hiérarchie européenne est de 4,1% si l’on ignore le contexte régional (cellule (1,2) dans le tableau 5.1a), mais elle n’est plus que de 0,6% pour des régions pauvres entourées de régions pauvres alors qu’elle est de 25,4% si cette région est entourée de régions moins pauvres.

Nous avons appliqué le test de stationnarité spatiale sur les deux séries (PIB par tête en Ecus et en SPA) pour nous assurer que les différences observées entre les probabilités de transition spatiales sont significatives. Rappelons que la présence d’éléments nuls dans la matrice de transition initiale et dans les différentes sous-matrices nous conduit à ajuster les degrés de liberté pour que ces éléments nuls ne soient pas pris en compte.

Décalage spatial	$t_i \backslash t_{i+1}$	Conditionnement européen					Distribution initiale
		1 <65%	2 <95%	3 <110%	4 <125%	5 >125%	
1	1	0,994	0,006	0,000	0,000	0,000	0,947
	2	0,200	0,800	0,000	0,000	0,000	0,053
	3	0,000	0,000	0,000	0,000	0,000	0,000
	4	0,000	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000
2	1	0,746	0,254	0,000	0,000	0,000	0,177
	2	0,064	0,893	0,043	0,000	0,000	0,703
	3	0,000	0,103	0,828	0,069	0,000	0,087
	4	0,000	0,000	0,182	0,818	0,000	0,033
	5	0,000	0,000	0,000	0,000	0,000	0,000
3	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,821	0,179	0,000	0,000	0,307
	3	0,000	0,182	0,802	0,017	0,000	0,442
	4	0,000	0,000	0,097	0,774	0,129	0,113
	5	0,000	0,000	0,000	0,079	0,921	0,139
4	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,862	0,138	0,000	0,000	0,096
	3	0,000	0,044	0,859	0,097	0,000	0,375
	4	0,000	0,000	0,147	0,796	0,058	0,316
	5	0,000	0,000	0,000	0,101	0,899	0,213
5	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,600	0,400	0,000	0,000	0,010
	3	0,000	0,107	0,714	0,179	0,000	0,116
	4	0,000	0,000	0,094	0,799	0,107	0,310
	5	0,000	0,000	0,000	0,059	0,941	0,563

Tableau 5.10a : Chaîne de Markov spatiale ;
conditionnement sur le décalage spatial au début de chaque année ;
PIB par tête en Ecus

Décalage spatial	$t_i \backslash t_{i+1}$	Conditionnement européen					Distribution initiale
		1 <75%	2 <95%	3 <105%	4 <120%	5 >120%	
1	1	0,989	0,011	0,000	0,000	0,000	0,884
	2	0,081	0,892	0,027	0,000	0,000	0,116
	3	0,000	0,000	0,000	0,000	0,000	0,000
	4	0,000	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000
2	1	0,890	0,110	0,000	0,000	0,000	0,392
	2	0,140	0,766	0,093	0,000	0,000	0,385
	3	0,000	0,137	0,824	0,039	0,000	0,183
	4	0,000	0,000	0,182	0,818	0,000	0,040
	5	0,000	0,000	0,000	0,000	0,000	0,000
3	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,926	0,074	0,000	0,000	0,403
	3	0,000	0,109	0,871	0,020	0,000	0,337
	4	0,000	0,000	0,048	0,762	0,190	0,070
	5	0,000	0,000	0,000	0,070	0,930	0,190
4	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,927	0,073	0,000	0,000	0,183
	3	0,000	0,069	0,857	0,074	0,000	0,263
	4	0,000	0,003	0,077	0,883	0,037	0,364
	5	0,000	0,000	0,000	0,064	0,936	0,190
5	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,933	0,067	0,000	0,000	0,043
	3	0,000	0,000	0,872	0,128	0,000	0,135
	4	0,000	0,000	0,044	0,876	0,080	0,325
	5	0,000	0,000	0,000	0,069	0,931	0,497

Tableau 5.10b : Chaîne de Markov spatiale ;
conditionnement sur le décalage spatial au début de chaque année ;
PIB par tête en SPA

Le nombre de degrés de liberté est ainsi réduit à 23 degrés de liberté pour les PIB par tête en Ecus et à 26 degrés de liberté pour les PIB par tête en SPA . On obtient $Q_2 = 133,75$ pour la distribution des PIB par tête en Ecus et $Q_2 = 93,87$ pour la distribution des PIB par tête en SPA. Dans les deux cas, la probabilité critique est inférieure à $p < 0,0001$. Ce résultat indique que l'environnement géographique d'une région influence fortement ses mouvements dans la distribution des PIB.

Pour apprécier de façon plus synthétique les informations contenues dans les chaînes de Markov spatiales, nous avons étudié la relation entre la direction de la transition d'une région dans la distribution des PIB par tête et les PIB par tête de ses voisines en ne retenant que les trois trajectoires suivantes : transition à la hausse, état stationnaire ou transition à la baisse. Ces résultats sont donnés dans le tableau 5.11a pour les PIB par tête en Ecus et dans le tableau 5.11b pour les PIB par tête en SPA pour toutes les matrices des plus proches voisins⁷⁶. On trouve dans ces tableaux la probabilité d'une transition particulière (hausse, stationnaire, baisse) d'une région conditionnée par la classe où se trouvait la moyenne des PIB de ses voisines au début de l'année (plus pauvre, identique, plus riche).

Ainsi, si l'on considère le tableau 5.11a avec la matrice des 10 plus proches voisins, on s'aperçoit que la probabilité de descendre ou de monter est différente selon le contexte régional. Par exemple, la probabilité qu'une région monte d'une classe est de 13,8% si la région est entourée de régions plus riches mais elle n'est que de 2,3% si ses voisines sont plus pauvres. Inversement, la probabilité d'une transition à la baisse est plus de deux fois supérieure lorsque les régions sont entourées de voisins plus pauvres que de voisins plus riches.

Pour vérifier ces différentes indications, nous avons effectué un test du χ^2 d'indépendance entre le type de transition subie par une région et la classe de PIB de ses voisines. Formellement, soit X la variable aléatoire discrète à 3 modalités, représentant le type de transition de la région (hausse, stationnaire, baisse) et Y la variable aléatoire discrète à 3 modalités, représentant la classe contenant la moyenne pondérée des PIB par tête des régions voisines (plus riche, identique, moins riche). Nous cherchons à savoir si ces deux variables sont indépendantes :

$$H_0 = \{\forall i, j : P[(X = i) \text{ et } (Y = j)] = P(X = i) \cdot P(Y = j)\} \quad [5.43]$$

Sous H_0 , la statistique du test s'écrit de la façon suivante :

$$X^2 = \sum_{i=1}^3 \sum_{j=1}^3 \frac{(n_{ij} - NQ_{ij})^2}{NQ_{ij}} \quad [5.44]$$

⁷⁶ L'annexe 5.4 fournit les mêmes résultats pour toutes les matrices de distance.

où n_{ij} est le nombre de régions se trouvant dans la cellule (i, j) durant les 15 transitions, $N = 2070$, et Q_{ij} est défini de la façon suivante :

$$Q_{ij} = \frac{\sum_{i=1}^3 n_{ij}}{N} \cdot \frac{\sum_{j=1}^3 n_{ij}}{N} \quad [5.45]$$

Cette statistique suit une loi du chi-deux à 4 degrés de liberté. Les valeurs des statistiques ainsi que les probabilités critiques (entre parenthèses) sont données dans les tableaux 5.11a et 5.11b. Toutes les statistiques conduisent à rejeter l'hypothèse nulle d'indépendance entre le type de transition subie par une région et l'environnement géographique auquel elle est confrontée. Par conséquent, les trajectoires économiques des régions sont caractérisées par une forme de verrouillage géographique.

Notons cependant que les propriétés de tous les tests proposés dans ce chapitre doivent encore être explorées en présence d'autocorrélation spatiale. En effet, comme l'a montré Fingleton (1983a, 1983b, 1986), l'inférence statistique dans le cadre des chaînes de Markov est problématique car dans ce cas, il y a une forte inflation des statistiques du chi-deux.

Matrice	Décalage spatial	N	Mouvement			Test du chi-deux
			Baisse	Stationnaire	Hausse	
W(10)	plus pauvre	258	0,109	0,868	0,023	195,51 (0,000)
	identique	1174	0,069	0,910	0,021	
	plus riche	638	0,047	0,815	0,138	
W(15)	plus pauvre	279	0,115	0,857	0,029	153,97 (0,000)
	identique	1116	0,068	0,912	0,020	
	plus riche	675	0,046	0,822	0,132	
W(20)	plus pauvre	274	0,113	0,880	0,007	119,49 (0,000)
	identique	1069	0,069	0,905	0,026	
	plus riche	727	0,047	0,831	0,122	

Notes : pour le test du chi-deux, les probabilités critiques sont entre parenthèses.

Tableau 5.11a : Probabilités de transition conditionnées sur le décalage spatial au début de chaque année ; PIB par tête en Ecus

Matrice	Décalage spatial	N	Mouvement			Test du Chi-deux
			Baisse	Stationnaire	Hausse	
W(10)	plus pauvre	334	0,081	0,898	0,021	52,42 (0,000)
	identique	963	0,064	0,909	0,027	
	plus riche	773	0,026	0,891	0,083	
W(15)	plus pauvre	382	0,089	0,893	0,018	61,51 (0,000)
	identique	935	0,062	0,91	0,028	
	plus riche	753	0,023	0,892	0,085	
W(20)	plus pauvre	417	0,086	0,902	0,012	50,37 (0,000)
	identique	814	0,064	0,900	0,036	
	plus riche	839	0,025	0,900	0,075	

Notes : pour le test du chi-deux, les probabilités critiques sont entre parenthèses.

Tableau 5.11b : Probabilités de transition conditionnées sur le décalage spatial au début de chaque année ; PIB par tête en SPA

Conclusion de section

Dans cette section, nous avons montré que les facteurs géographiques constituent une composante importante de l'explication des trajectoires de croissance des régions européennes au sein de la distribution des PIB par tête. Par rapport aux analyses a-spatiales conduites dans la première section, les résultats obtenus soulignent l'importance du conditionnement géographique. Par exemple, le biais progressif vers la trappe à pauvreté possède principalement une explication spatiale puisque les régions pauvres sont négativement influencées en étant entourées par d'autres régions pauvres. De même, les voisines riches ont tendance à « empêcher » les régions voisines de descendre dans la hiérarchie. L'absence relative de mobilité intra-distribution peut être expliquée par des concentrations spatiales persistantes de régions avec des PIB par tête très élevés ou très bas. Plus généralement, dans la perspective d'une convergence interrégionale, les chaînes de Markov spatiales indiquent que les changements de position relative dans la distribution en coupe transversale sont fortement contraints par l'environnement géographique.

Conclusion du chapitre 5

Face aux limites du concept de β -convergence, nous nous sommes attachés à analyser l'évolution des disparités de PIB par tête et le processus de convergence entre les régions européennes sur la période 1980-1995. La méthodologie choisie est suggérée par Quah (1993a, 1993b, 1996a) : la convergence est mesurée à partir de l'évolution de la forme de la distribution et à partir des changements de position relative des régions dans cette distribution. L'évolution de la distribution en coupe transversale a été étudiée grâce au concept de σ -convergence tandis que la polarisation des régions européennes a été évaluée à l'aide d'estimations non-paramétriques de fonctions de densité et de l'indice d'Esteban et Ray (1994). La mobilité des régions au sein de la distribution des PIB par tête est mise en évidence par la modélisation du processus de croissance comme une chaîne de Markov de premier ordre stationnaire. Enfin, le rôle de l'environnement géographique est intégré explicitement, grâce à la procédure de conditionnement régional et à l'estimation de chaînes de Markov spatiales.

Concernant notre échantillon de 138 régions européennes pour la période 1980-1995 et les deux variables étudiées (PIB par tête en Ecus et PIB par tête en SPA), il apparaît que le processus de croissance économique dans l'Union Européenne durant les années 80 a globalement été caractérisé par une grande persistance des disparités régionales, une relative absence de mobilité des régions dans la distribution des PIB par tête régionaux ainsi qu'un biais progressif vers une trappe à pauvreté. Dans ces conditions, la polarisation de la distribution des PIB par tête entre les deux groupes de régions à haut et bas PIB par tête va probablement persister dans le futur et ceci même en présence du processus d'intégration européenne. En outre, la procédure de conditionnement régional et les chaînes de Markov spatiales indiquent clairement que, en dépit d'une dématérialisation progressive des économies, la localisation et la géographie physique comptent toujours pour expliquer les processus de croissance et de convergence. En effet, la convergence intra-régionale est très importante et, du point de vue de la convergence interrégionale, les changements de position relative d'une région dans la distribution sont fortement contraints par son environnement géographique. Tous ces résultats sont à rapprocher des principes de verrouillage mis en avant dans les théories de la formation des agglomérations. Plus précisément, si on s'intéresse aux théories géographiques de la croissance, les conditions géographiques, comme les conditions économiques, influencent fortement les potentialités de développement des économies régionales.

Annexe 5.1

Tests d'autocorrélation spatiale avec les matrices de poids basées sur la distance

Transitions entre classes de PIB

		Hausse	Baisse
Ecus	W(1)	0,0009	0,0052
	W(2)	0,0164	0,0011
	W(3)	0,0282	0,0112
SPA	W(1)	0,0225	0,0223
	W(2)	0,2114	0,0046
	W(3)	0,0731	0,0319

Tableau 1 : Valeurs critiques pour les tests d'autocorrélation spatiale ; transitions entre classes de PIB ; PIB par tête en Ecus et en SPA

Modification des classements des régions

		Hausse	Baisse
Ecus	W(1)	0,0005	0,0089
	W(2)	0,0012	0,0024
	W(3)	0,0980	0,1012
SPA	W(1)	0,0104	0,0025
	W(2)	0,1421	0,0012
	W(3)	0,1687	0,0774

Tableau 2 : Valeurs critiques pour les tests d'autocorrélation spatiale ; classement des régions ; PIB par tête en Ecus et en SPA

Annexe 5.2

Conditionnement régional avec la matrice des distances $W(1)$; PIB par tête en Ecus

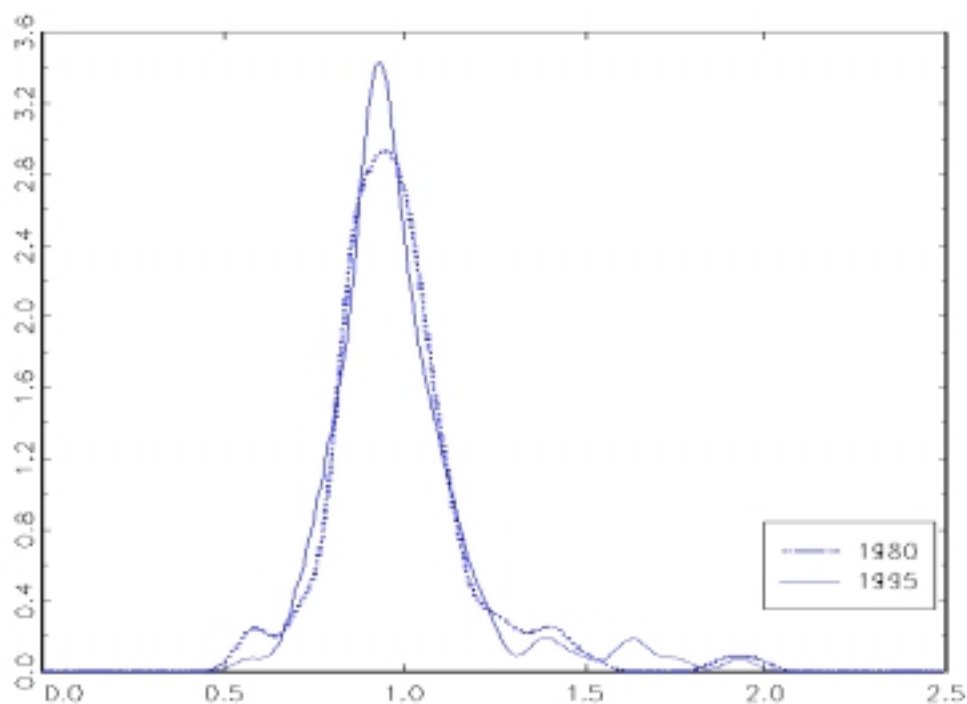


Figure 1 : Densités des distributions de PIB par tête en Ecus ;
1980 - 1995 ; conditionnement régional

		<i>t_i</i>					Nombre d'observations
		Conditionnement régional					
<i>t_i</i>		1	2	3	4	5	
		<65%	<95%	<110%	<125%	>125%	
Conditionnement européen	1	0,038	0,498	0,347	0,076	0,042	450
	2	0,031	0,533	0,336	0,061	0,040	426
	3	0,030	0,652	0,286	0,026	0,006	465
	4	0,000	0,428	0,504	0,043	0,025	397
	5	0,000	0,049	0,451	0,245	0,255	470

Tableau 1 : Conditionnement régional ;
PIB par tête en Ecus

		<i>t_{i+1}</i>					Nombre d'observations
		Conditionnement régional					
<i>t_i</i>		1	2	3	4	5	
		<65%	<95%	<110%	<125%	>125%	
Conditionnement régional	1	0,837	0,163	0,000	0,000	0,000	43
	2	0,005	0,924	0,072	0,000	0,000	878
	3	0,000	0,085	0,878	0,038	0,000	800
	4	0,000	0,000	0,153	0,805	0,042	190
	5	0,000	0,000	0,000	0,050	0,950	159

Tableau 2 : Matrice de transition ;
conditionnement régional ; PIB par tête en Ecus

Décalage spatial	t_i \ t_{i+1}	Conditionnement européen					Distribution initiale
		1	2	3	4	5	
		<65%	<95%	<110%	<125%	>125%	
1	1	0,989	0,011	0,000	0,000	0,000	0,943
	2	0,182	0,818	0,000	0,000	0,000	0,057
	3	0,000	0,000	0,000	0,000	0,000	0,000
	4	0,000	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000
2	1	0,740	0,260	0,000	0,000	0,000	0,152
	2	0,063	0,895	0,042	0,000	0,000	0,724
	3	0,000	0,227	0,682	0,091	0,000	0,067
	4	0,000	0,000	0,182	0,818	0,000	0,033
	5	0,000	0,000	0,000	0,000	1,000	0,024
3	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,776	0,224	0,000	0,000	0,269
	3	0,000	0,154	0,832	0,014	0,000	0,574
	4	0,000	0,000	0,125	0,792	0,083	0,096
	5	0,000	0,000	0,000	0,133	0,867	0,060
4	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,889	0,111	0,000	0,000	0,083
	3	0,000	0,038	0,856	0,106	0,000	0,320
	4	0,000	0,000	0,131	0,799	0,070	0,329
	5	0,000	0,000	0,000	0,097	0,903	0,269
5	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,789	0,211	0,000	0,000	0,042
	3	0,000	0,100	0,733	0,167	0,000	0,133
	4	0,000	0,000	0,105	0,789	0,105	0,294
	5	0,000	0,000	0,000	0,054	0,946	0,531

Tableau 3 : Chaîne de Markov spatiale ;
conditionnement sur le décalage spatial au début de chaque année ;
PIB par tête en Ecus

Annexe 5.3

Conditionnement régional avec la matrice des distances $W(1)$; PIB par tête en SPA

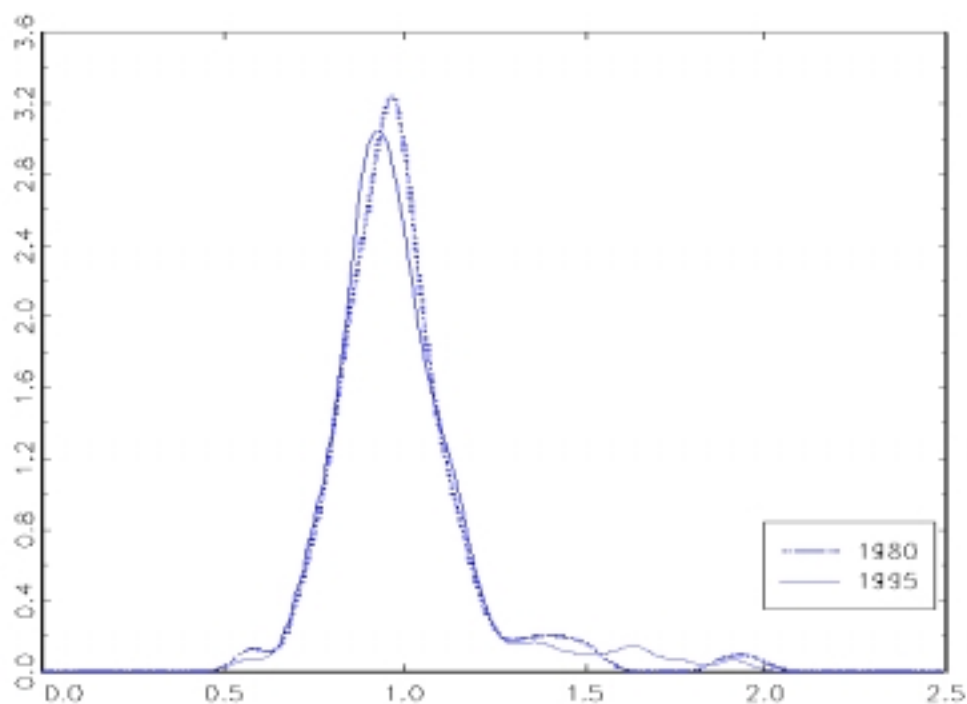


Figure 1 : Densités des distributions de PIB par tête mesuré en SPA ; 1980 - 1995 ; conditionnement régional

t_i \ t_i		Conditionnement régional					Nombre d'observations
		1 <75%	2 <95%	3 <105%	4 <120%	5 >120%	
Conditionnement européen	1	0,096	0,545	0,215	0,132	0,012	418
	2	0,125	0,530	0,220	0,052	0,073	464
	3	0,023	0,590	0,231	0,131	0,025	442
	4	0,002	0,274	0,617	0,087	0,019	470
	5	0,000	0,002	0,254	0,372	0,372	414

Tableau 1 : Conditionnement régional ;
PIB par tête en SPA

t_i \ t_i		Conditionnement régional					Nombre d'observations
		1 <75%	2 <95%	3 <105%	4 <120%	5 >120%	
Conditionnement régional	1	0,837	0,163	0,000	0,000	0,000	43
	2	0,005	0,924	0,072	0,000	0,000	878
	3	0,000	0,085	0,878	0,038	0,000	800
	4	0,000	0,000	0,153	0,805	0,042	190
	5	0,000	0,002	0,000	0,050	0,950	159

Tableau 2 : Matrice de transition ;
conditionnement régional ; PIB par tête en SPA

Décalage spatial	$t_i \backslash t_{i+1}$	Conditionnement européen					Distribution initiale
		1	2	3	4	5	
		< 75%	< 95%	< 105%	< 120%	> 120%	
1	1	0,984	0,016	0,000	0,000	0,000	0,847
	2	0,122	0,857	0,020	0,000	0,000	0,134
	3	0,000	0,000	1,000	0,000	0,000	0,019
	4	0,000	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000
2	1	0,877	0,123	0,000	0,000	0,000	0,277
	2	0,091	0,818	0,091	0,000	0,000	0,452
	3	0,000	0,148	0,820	0,033	0,000	0,209
	4	0,000	0,000	0,286	0,714	0,000	0,024
	5	0,000	0,000	0,000	0,000	1,000	0,038
3	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,897	0,103	0,000	0,000	0,375
	3	0,000	0,151	0,837	0,012	0,000	0,371
	4	0,000	0,000	0,048	0,762	0,190	0,091
	5	0,000	0,000	0,000	0,079	0,921	0,164
4	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,933	0,067	0,000	0,000	0,135
	3	0,000	0,045	0,874	0,081	0,000	0,254
	4	0,000	0,003	0,071	0,888	0,038	0,387
	5	0,000	0,000	0,000	0,086	0,914	0,224
5	1	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,000	0,955	0,045	0,000	0,000	0,146
	3	0,000	0,026	0,846	0,128	0,000	0,130
	4	0,000	0,000	0,052	0,857	0,091	0,256
	5	0,000	0,000	0,000	0,043	0,957	0,468

Tableau 3 : Chaîne de Markov spatiale ;
conditionnement sur le décalage spatial au début de chaque année ;
PIB par tête en SPA

Annexe 5.4

Probabilités de transitions conditionnées sur le décalage spatial pour les matrices de distances

	Décalage spatial	N	Mouvement			Test du Chi-deux
			Baisse	Stationnaire	Hausse	
W(1)	plus pauvre	277	0,119	0,866	0,014	182,16 (0,000)
	identique	1202	0,065	0,909	0,026	
	plus riche	591	0,047	0,810	0,142	
W(2)	plus pauvre	284	0,116	0,873	0,011	142,85 (0,000)
	identique	1135	0,068	0,907	0,025	
	plus riche	651	0,045	0,820	0,135	
W(3)	plus pauvre	299	0,114	0,876	0,010	123,05 (0,000)
	identique	1077	0,073	0,903	0,024	
	plus riche	694	0,037	0,833	0,130	

Tableau 1 : Probabilités de transition conditionnées sur le décalage spatial au début de chaque année ; PIB par tête en Ecus

	Décalage spatial	N	Mouvement			Test du chi-deux
			Baisse	Stationnaire	Hausse	
W(1)	plus pauvre	391	0,097	0,885	0,018	63,80 (0,000)
	identique	1009	0,056	0,914	0,031	
	plus riche	670	0,022	0,889	0,088	
W(2)	plus pauvre	407	0,086	0,902	0,012	61,21 (0,000)
	identique	946	0,060	0,908	0,032	
	plus riche	717	0,024	0,890	0,086	
W(3)	plus pauvre	412	0,085	0,908	0,007	69,04 (0,000)
	identique	910	0,066	0,901	0,033	
	plus riche	748	0,019	0,896	0,086	

Tableau 2 : Probabilités de transition conditionnées sur le décalage spatial au début de chaque année ; PIB par tête en SPA

Conclusion

Dans cette thèse, nous avons cherché à approfondir l'étude du processus de convergence des régions européennes en prenant en compte les schémas de répartition spatiale des activités. Ceux-ci sont définis par la façon dont les activités économiques occupent un territoire et sont le plus souvent caractérisés par un degré plus ou moins fort d'inégalité spatiale : des parties du territoire concentrent une part importante des activités, tandis que d'autres en sont relativement dépourvues. Cet objectif a été réalisé en cinq chapitres.

Nous avons d'abord justifié pourquoi il nous a semblé important de tenir compte des disparités géographiques dans l'étude de la convergence des régions européennes.

De nombreuses notions de convergence ont été proposées, chacune fournissant une information particulière sur la façon dont les disparités économiques et les modalités de leur réduction sont conceptualisées. Dans un premier temps, les deux notions de convergence-rattrapage, déduites du modèle néoclassique de croissance de Solow-Swan, ont été présentées : la convergence absolue et la convergence conditionnelle. Même si les équations visant à tester ce concept de convergence (la « β -convergence ») sont formellement déduites du modèle de croissance néo-classique et de ses extensions, nous avons montré pourquoi il est problématique de considérer ce concept de convergence comme un moyen simple permettant de discriminer les modèles de croissance exogène des modèles de croissance endogène.

Dans un second temps, les procédures d'estimation et de tests pour le concept de β -convergence et pour des concepts alternatifs de convergence ont été détaillées. En particulier, l'interprétation des tests de β -convergence est rendue difficile par l'existence de nombreux problèmes économétriques, comme l'endogénéité et la multicolinéarité des variables explicatives ou le biais lié aux variables omises. L'un de ces problèmes est l'hétérogénéité des paramètres de régression et peut être relié au concept de convergence en clubs, dans lequel les économies, partageant des conditions initiales similaires, convergent vers un même état régulier. D'autres concepts de convergence ont également été présentés, ainsi que leurs avantages et leurs inconvénients : σ -convergence, convergence stochastique ou convergence évaluée à partir de la dynamique de la répartition des PIB par tête.

Cependant, dans tous les cas, ces concepts négligent la dimension spatiale associée aux données. Ceci a pour conséquence d'ajouter une limite supplémentaire aux problèmes habituellement relevés dans les tests de convergence. Plus précisément, nous avons souligné que la concentration géographique des activités économiques dans l'espace était susceptible de conduire à une autocorrélation spatiale alors que la polarisation en deux ou plusieurs régimes pouvait refléter l'existence de clubs de convergence. D'un point de vue théorique, ces différents éléments peuvent être reliés au corpus théorique de l'économie géographique, développé à partir du modèle interrégional de Krugman (1991a, 1991b). La synthèse « géographie-croissance » permet également de préciser les liens entre les schémas géographiques et la croissance économique. Les disparités géographiques et leurs modes de formation interagissent alors avec les processus de croissance.

Il devient ainsi nécessaire de prendre en compte, dans le processus de convergence, le phénomène de l'autocorrélation spatiale et celui de l'hétérogénéité spatiale. Le premier

renvoie à une concentration géographique de valeurs similaires de PIB alors que le second est relié à une différenciation spatiale des régions selon, par exemple, un dualisme de type Nord/Sud.

La prise en compte de l'autocorrélation et de l'hétérogénéité spatiales au sein des modèles de convergence nécessite de définir les spécifications économétriques ainsi que les méthodes d'estimation et d'inférence adaptées. Dans le deuxième chapitre, nous avons donc présenté une revue des techniques de la statistique et de l'économétrie spatiales.

Tout d'abord, les méthodes récemment développées de l'analyse exploratoire des données spatiales ont été détaillées. Ces techniques permettent de décrire et de visualiser les distributions spatiales, d'identifier les localisations atypiques, de détecter les schémas d'association spatiale et enfin de suggérer les régimes spatiaux et les autres formes d'hétérogénéité spatiale. En particulier, elles fournissent une évaluation de l'autocorrélation spatiale. Nous avons ainsi présenté la notion de matrice de poids, qui est l'outil fondamental permettant de capter les interactions spatiales, ainsi que les différentes statistiques permettant de tester l'existence d'une autocorrélation spatiale globale et locale.

Les méthodes d'estimation et d'inférence adaptées aux modèles économétriques intégrant l'autocorrélation et l'hétérogénéité spatiales ont ensuite été présentées. Dans le premier cas, nous avons détaillé les trois spécifications principales permettant de capter l'autocorrélation spatiale et les interprétations associées aux paramètres, notamment en termes d'effets de débordement géographiques. Nous avons également insisté, d'une part, sur les méthodes permettant d'estimer d'une façon convergente et efficiente ces spécifications et d'autre part, sur les tests et les règles de décision permettant de choisir, parmi les différentes possibilités, la spécification la plus appropriée. Concernant l'hétérogénéité spatiale, nous avons distingué les modèles selon que l'hétérogénéité se manifeste sous la forme d'une instabilité des paramètres ou d'une hétéroscédasticité en groupes. Un paragraphe spécifique a enfin été consacré au cas où autocorrélation et hétérogénéité spatiales se trouvent conjointement dans une régression. Les méthodes d'estimation et d'inférence doivent alors être adaptées et différentes procédures ont été présentées.

Le cadre théorique et conceptuel ayant été précisé dans les deux premiers chapitres, nous avons réalisé dans les trois derniers chapitres de la thèse un ensemble d'analyses empiriques du phénomène de convergence en présence de disparités spatiales. Le terrain d'application est la convergence régionale à partir d'un échantillon de 138 régions européennes, appartenant à 11 pays, sur la période 1980-1995. Les données de PIB par tête sont exprimées en Ecus et en SPA et sont extraites de la base Eurostat-Regio. Le niveau de désagrégation spatiale est le niveau NUTS 2 pour tous les pays⁷⁷ sauf pour le Royaume-Uni. Compte tenu des limites imposées par l'utilisation de cette base de données, nous avons choisi de fournir l'ensemble des résultats empiriques pour les PIB par tête en Ecus et en SPA, qui sont les données corrigées des parités de pouvoir d'achat. L'ensemble des résultats a

⁷⁷ L'Allemagne, la Belgique, le Danemark, l'Espagne, la France, la Grèce, l'Italie, le Luxembourg, les Pays-Bas et le Portugal.

également été fourni pour tout un ensemble de matrices de poids, afin de vérifier leur robustesse au choix de la matrice. Notre objectif dans les trois derniers chapitres de la thèse n'a pas été de tester formellement les modèles d'économie géographique mais de fournir une évaluation du rôle joué par l'espace dans le processus de convergence des régions européennes. Ces analyses ont alors mis en évidence trois résultats importants : la convergence des régions européennes est relativement faible, les effets de débordement géographiques sont importants et il existe une forte polarisation Nord-Sud au sein des régions européennes. Ces résultats semblent robustes puisqu'ils ont été obtenus à l'aide de méthodologies différentes. Afin de mettre en évidence ces différents résultats, nous avons procédé en trois étapes.

La première étape a consisté à démontrer que les régions européennes sont effectivement caractérisées par un double effet d'autocorrélation et d'hétérogénéité spatiales en termes des PIB par tête et des taux de croissance. Ce résultat a été mis en évidence à l'aide d'une analyse exploratoire des données spatiales.

Dans une première section, ces méthodes ont été appliquées sur les niveaux de PIB par tête pour l'ensemble de la période 1980-1995. Une forte autocorrélation spatiale globale a été détectée, quelle que soit la matrice de poids utilisée, ce qui indique une persistance de la concentration géographique de régions similaires en niveaux. Ce résultat global a été affiné par le calcul de différents indicateurs d'association spatiale locale. Les statistiques de Getis-Ord ont ainsi permis de mettre en évidence une hétérogénéité spatiale sous la forme d'une polarisation marquée entre les régions du Nord et les régions du Sud. Le diagramme de Moran et les statistiques LISA ont confirmé ce résultat tout en détectant, en plus, des régions atypiques, se démarquant significativement de leurs voisines. Pour chaque indicateur d'association spatiale, les valeurs extrêmes ont été détectées. Certains résultats diffèrent quelque peu selon que les PIB sont exprimés en Ecus ou en SPA, notamment sur la localisation des concentrations géographiques de valeurs élevées. Dans tous les cas cependant, les résultats sont robustes au choix de la matrice de poids et indiquent la présence de deux régimes au sein de l'échantillon. De fortes disparités géographiques sont ainsi persistantes sur toute la période d'étude. Le schéma de polarisation Nord-Sud, interprété comme un schéma de type centre-périphérie, apparaît alors comme une caractéristique marquante des régions européennes.

Ces conclusions ont quelque peu été nuancées par l'analyse exploratoire sur les taux de croissance annuels moyens entre 1980 et 1995. Il apparaît d'abord que le taux de croissance annuel moyen est moins fortement spatialement autocorrélé que les niveaux de PIB par tête. Ensuite, le dynamisme de certaines régions pauvres, au sein du groupe Sud, est mis en évidence alors que d'autres concentrations de régions initialement riches ont connu des taux de croissance inférieurs à la moyenne. Cependant, cette relation inverse entre niveau de richesse initial et taux de croissance ne constitue pas une généralité puisque certaines régions pauvres sont coincées dans des trappes de sous-développement et que des régions très riches au départ ont encore connu des taux de croissance relativement élevés. Le processus de

rattrapage des régions riches par les régions pauvres apparaît ainsi, en première analyse, comme relativement faible.

Ces conclusions sont reprises formellement, dans l'analyse du processus de β -convergence à l'œuvre au sein des régions européennes.

Dans un premier temps, seule l'autocorrélation spatiale dans les régressions de β -convergence a été intégrée. Nous avons d'abord montré que les diverses spécifications incluant l'autocorrélation spatiale dans les modèles de β -convergence absolue pouvaient être interprétées comme des modèles de *β -convergence conditionnelle minimaux intégrant des variables d'environnement spatial*. Des effets de débordement géographiques sont également mis en évidence et sont associés à différentes interprétations. Ainsi, l'inclusion d'une variable endogène décalée permet de capter l'influence qu'exerce, sur le taux de croissance d'une région, le taux de croissance des régions voisines. L'influence exercée par le PIB par tête initial des régions voisines sur le taux de croissance d'une région est prise en compte à travers une variable exogène décalée. Enfin, une autocorrélation spatiale des erreurs permet de capter ces deux types d'effets de débordement géographiques. A l'aide des différents tests d'autocorrélation spatiale détaillés dans le chapitre 2, nous avons alors montré que le modèle avec autocorrélation spatiale des erreurs est le plus approprié, quelle que soit la matrice de poids utilisée. L'estimation de ce modèle par la méthode du maximum de vraisemblance et des moments généralisés a confirmé la faible vitesse de convergence entre les régions européennes. Nous avons également montré l'existence d'un effet de débordement géographique : le taux de croissance d'une région étant significativement et positivement influencé par le taux de croissance des régions voisines. Les propriétés de diffusion impliquées par le modèle avec autocorrélation des erreurs ont finalement été étudiées à l'aide de simulations. Nous avons d'abord supposé l'existence d'un choc aléatoire sur la région Ile-de-France et analysé le processus de diffusion spatiale impliqué par ce choc. Nous avons ensuite développé une procédure originale permettant d'étudier simultanément les effets d'un choc affectant toutes les régions de l'échantillon. Les simulations ont permis de détecter une certaine polarisation du processus de diffusion spatiale puisque les régions, bénéficiant le plus d'un choc de croissance et dont l'influence est la plus élevée, se trouvent au Nord alors que les régions pauvres du Sud sont à la fois les moins influentes et les moins influencées.

Dans un deuxième temps, la possibilité d'une polarisation des régions européennes en deux groupes selon un schéma de type centre-périphérie Nord/Sud. Nous avons pris en compte simultanément l'autocorrélation spatiale, sous la forme d'une autocorrélation spatiale des erreurs, et l'hétérogénéité spatiale, sous la forme d'un modèle à deux régimes. Les régimes ont été établis sur la base de critères géographiques en utilisant le diagramme de Moran, lorsque les PIB par tête sont exprimés en Ecus, et les statistiques de Getis-Ord, lorsque les PIB par tête sont exprimés en SPA. Ces régimes sont interprétés comme des clubs de convergence spatiaux. L'estimation et l'inférence statistique associées à ces modèles posent quelques problèmes, mais deux résultats forts apparaissent cependant. D'une part, il existe toujours un effet de débordement géographique significatif affectant les taux de croissance. D'autre part, il apparaît que les régions pauvres convergent entre elles à un rythme

plus rapide que celui observé dans l'échantillon total alors que les régions riches ne convergent pas entre elles. Le processus de rattrapage observé au niveau global apparaît alors être un artéfact.

Les tests de β -convergence posant de nombreux problèmes méthodologiques, nous avons choisi d'étudier, dans une troisième étape, un concept alternatif de convergence permettant de prendre en compte l'évolution des disparités spatiales dans le temps. Dans le cinquième chapitre, nous avons alors étudié et adapté le concept de convergence suggéré par Quah (1993a, 1993b, 1996a).

Dans la première section, nous avons simplement appliqué le concept de convergence de Quah sans tenir compte des disparités spatiales. L'estimation non paramétrique de la forme de la distribution des PIB par tête en début et en fin de période montre une tendance à la bimodalité, plus marquée pour les PIB par tête exprimés en Ecus que pour les PIB par tête exprimés en SPA. Ce résultat suggère la formation *d'une polarisation du processus de convergence*. Pour étudier plus précisément la mobilité des régions au sein de la distribution des PIB par tête, au cours de la période, nous avons alors discrétisé notre série de PIB par tête en cinq classes et estimé une matrice de transition associée à une chaîne de Markov stationnaire de premier ordre. Les probabilités de transition ont été estimées par le maximum de vraisemblance et un test de stabilité des probabilités de transition, entre les périodes 1980-1989 et 1989-1995, a été effectué. Il apparaît que la mobilité des régions au sein de la distribution des PIB par tête est très faible : les régions ne modifient pas d'une façon importante leur classement relatif au cours du temps et ont plutôt tendance à rester dans la même classe de PIB par tête. L'analyse des distributions de long terme fait en outre apparaître *une trappe à pauvreté* car un groupe non-négligeable de régions converge vers un niveau relativement faible de PIB par tête.

Dans la seconde section, nous prenons explicitement en compte la répartition géographique des PIB par tête dans l'analyse de la dynamique de la distribution des richesses sur la période 1980-1995. Nous justifions cette approche en remarquant que les régions, ayant le plus modifié leur classement dans la distribution des PIB par tête, sont spatialement autocorrélées. Trois méthodes, permettant d'évaluer le rôle joué par la localisation, ont été mises en œuvre. La première méthode permet de capter la façon dont les variations du PIB par tête d'une région, une année donnée, sont expliquées par les variations de PIB des régions voisines. Si cette approche permet de quantifier l'ampleur des effets de débordement géographiques, elle conduit en revanche à la perte de la dimension temporelle associée aux matrices de transition habituellement utilisées. Une deuxième méthode permet d'évaluer une notion de convergence intra-régionale, c'est-à-dire la façon dont les régions rattrapent, au cours de la période, leurs voisines. Il apparaît alors que les régions se démarquent très rarement de leur environnement géographique et qu'il existe finalement peu de régions restant beaucoup plus riches ou beaucoup plus pauvres que leurs voisines. Les résultats fournis par cette deuxième approche rejoignent ceux mis en évidence dans le troisième chapitre. En revanche, elle ne permet plus de capter la convergence au sens défini dans la première section du cinquième chapitre. Pour pouvoir étudier simultanément la mobilité des régions au sein de

la distribution des PIB par tête et le rôle joué par l'environnement géographique dans cette mobilité, nous avons alors appliqué et approfondi la méthodologie suggérée par Rey (2001). Cette méthode permet de quantifier l'influence qu'exerce, sur les probabilités de transition d'une région dans la distribution des PIB par tête, la classe de richesse dans lequel se trouvent les régions voisines. Nous avons alors mis en évidence un phénomène de verrouillage géographique puisque les changements de position relative dans la distribution en coupe transversale sont fortement contraints par l'environnement géographique d'une région : la probabilité qu'une région améliore son classement dans la distribution des PIB par tête est d'autant plus faible qu'elle est entourée de régions pauvres. Au contraire, une région entourée de voisines riches a peu de chances de connaître une dégradation importante de son classement dans la distribution des PIB par tête. Un test permettant de vérifier ces conclusions a également été suggéré.

L'ensemble de ces résultats confirme que les schémas de répartition spatiale des activités économiques ne sont pas neutres car ils conditionnent les processus d'évolution des disparités économiques. Les techniques de la statistique et de l'économétrie spatiales permettent d'étudier ces interactions. Plusieurs pistes semblent pourtant pouvoir être explorées davantage.

Sur un plan technique, différents outils utilisés dans le cadre de cette thèse doivent être améliorés ou leurs propriétés davantage précisées. Ainsi, dans le troisième chapitre, nous avons mobilisé différents indicateurs, permettant de détecter les observations extrêmes au sein de l'échantillon (mesure de levier, résidus studentisés, distance de Cook), dont on ne connaît pas les performances en présence d'autocorrélation spatiale. Dans le même ordre d'idée, différentes statistiques du chi-deux ont été utilisées dans le chapitre 5 pour tester les stationnarités temporelle et spatiale des probabilités de transition. En présence d'autocorrélation spatiale, il peut exister une forte inflation des valeurs des statistiques de ces tests (Fingleton, 1983a, 1983b, 1986). Il conviendrait alors d'étudier, par simulation, les propriétés de ces tests afin de confirmer les résultats obtenus dans ce chapitre.

Par ailleurs, les propriétés des tests d'autocorrélation spatiale ont été très largement étudiées par simulation et différentes règles de décision, permettant de choisir entre les spécifications ont été proposées (Anselin et Florax, 1995 ; Anselin *et al.*, 1996). En revanche, lorsque autocorrélation et hétérogénéité spatiales sont présentes simultanément, les propriétés des tests ne sont pas bien connues en échantillon fini. En outre, la procédure que nous avons utilisée dans la seconde section du quatrième chapitre est susceptible de conduire à des problèmes de *pretest*.

Enfin, la méthodologie permettant de tester, dans un modèle avec effets spatiaux différenciés, si les coefficients d'autocorrélation spatiale dans les deux régimes et entre les deux régimes sont significativement différents, doit également faire l'objet d'approfondissements. Ces derniers devraient porter sur les propriétés asymptotiques des estimateurs associés à ces coefficients et estimés par la méthode du maximum de vraisemblance.

Sur un plan économique, le résultat d'absence de convergence au sein des régions du Nord obtenu dans le quatrième chapitre est peut-être dû à une hétérogénéité non-observée au sein du régime Nord. Une solution consisterait alors à étudier plus en détail l'hétérogénéité au sein de ce régime afin de détecter d'autres clubs de convergence. Elle ne peut cependant être mise en œuvre que si le nombre de régions dans chaque régime est suffisamment grand. D'autres modalités de détection de clubs à partir de critères spatiaux peuvent donc être recherchées.

Ensuite, notre choix a été de négliger la dimension nationale. Nous avons alors montré que l'environnement géographique de chaque région compte et ceci qu'il soit constitué de régions appartenant au même pays ou à des pays différents. Des méthodes, évitant la méthodologie des variables muettes, tout en permettant d'évaluer si les effets de débordement géographiques sont différents entre régions d'un même pays et entre régions de différents pays, pourraient ainsi être mises en œuvre.

Enfin, une voie de recherche future consiste à réévaluer l'ensemble des résultats obtenus à la lumière de l'élargissement aux Pays d'Europe Centrale et Orientale (PECO) dans les prochaines années⁷⁸. La participation des PECO aux programmes communautaires a été préconisée par le Conseil européen de Copenhague en juin 1993. Cette possibilité a ensuite été formellement prévue dans les accords européens avec ces pays au fur et à mesure qu'ils ont été conclus, ou bien sous forme de protocoles additionnels aux accords pour ceux qui l'avaient déjà été. Les conditions et modalités concrètes de participation à certains programmes ont fait l'objet des premières décisions des Conseils d'association en 1997.

Ces pays se situent loin derrière les états membres de l'Union Européenne en termes de développement économique et leurs niveaux de PIB par tête sont bien inférieurs à la moyenne communautaire. A l'exemple de ce qui avait eu lieu lors de l'entrée de la Grèce, du Portugal et de l'Espagne, l'intégration de ces pays conduirait à un « enrichissement relatif » de certaines régions déjà dans l'Union Européenne, lorsque le PIB est calculé par référence à la moyenne communautaire. Ceci nous conduirait déjà à devoir réactualiser nos calculs effectués dans les troisième et cinquième chapitres qui étaient tous effectués par rapport à la moyenne de l'échantillon (diagramme de Moran, statistiques LISA, construction des séries de PIB par tête relatifs à la moyenne européenne...).

En outre, comme le souligne la Commission Européenne (1999, p. 161), ceci représente un défi considérable en augmentant « à coup sûr l'hétérogénéité au sein de l'Union Européenne ». Ainsi, à une hétérogénéité Nord-Sud devrait vraisemblablement s'ajouter une hétérogénéité Est-Ouest. Par ailleurs, l'environnement géographique des régions à l'Est de notre échantillon serait considérablement modifié. Les phénomènes généraux de dépendance à l'environnement et les effets de verrouillage géographique détectés dans notre étude empirique devraient alors conduire à une réévaluation des politiques de cohésion territoriale visant à réduire les déséquilibres spatiaux.

⁷⁸ Les dix PECO candidats à l'adhésion sont la Bulgarie, l'Estonie, la Hongrie, la Lettonie, la Lituanie, la Pologne, la Roumanie, la République Slovaque, la République Tchèque et la Slovénie.

Bibliographie

- Abdel-Rahman H.M. (1994) Economies of scope in intermediate goods and a system of cities, *Regional Science and Urban Economics*, 24, 497-524.
- Abdel-Rahman H.M. (1996) When do cities specialize in production ?, *Regional Science and Urban Economics*, 26, 1-22.
- Abraham F., Von Rompuy P. (1995) Regional convergence in the European Monetary Union, *Papers in Regional Science*, 74, 125-142.
- Abramovitz M. (1986) Catching up, forging ahead, and falling behind, *Journal of Economic History*, 46, 385-406.
- Ades A., Chua H.B. (1997) Thy neighbor's curse : regional instability and economic growth, *Journal of Economic Growth*, 2, 279-304.
- Aghion P., Howitt P. (1992) A model of growth through creative destruction, *Econometrica*, 60, 323-351.
- Aghion P., Howitt P. (1998) *Endogenous Growth Theory*, MIT Press, Cambridge.
- Agterberg F. (1984) Trend surface analysis, dans : Gaile G.L., Wilmot C.J. (eds.), *Spatial Statistics and Models*, Reidel, Boston.
- Akaike H. (1974) A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, AC-19, 716-723.
- Alonso W. (1964) *Location and Land-Use*, Harvard University Press, Cambridge.
- Amable B. (1999) Un survol des théories de la croissance endogène, dans : Beine M., Docquier F. (eds.), *Croissance et Convergence Economiques des Régions*, De Boeck Université, Bruxelles.
- Amable B., Guellec D. (1992) Les théories de la croissance endogène, *Revue d'Economie Politique*, 102, 313-377.
- Amemiya T. (1985) *Advanced Econometrics*, Harvard University Press, Cambridge.
- Anderson J.E. (1982) Cubic spline urban density functions, *Journal of Urban Economics*, 12, 55-67.
- Anderson J.E. (1985) The changing structure of a city : temporal changes in cubic spline urban density patterns, *Journal of Regional Science*, 25, 413-125.
- Anderson T.W., Goodman L.A. (1957) Statistical inference about Markov chains, *Annals of Mathematical Statistics*, 28, 89-110.
- Andrés J., Doménech R., Molinas C. (1996) Macroeconomic performance and convergence in OECD countries, *European Economic Review*, 40, 1683-1704.
- Anselin L. (1980) *Estimation Methods for Spatial Autoregressive Structures*, Cornell University, Regional Science Dissertation and Monograph Series #8, Ithaca, NY.
- Anselin L. (1982) A note on small sample properties of estimators in a first-order spatial autoregressive model, *Environment and Planning A*, 14, 1023-1030.
- Anselin L. (1988a) *Spatial Econometrics : Methods and Models*, Kluwer Academic Publishers, Dordrecht.
- Anselin L. (1988b) Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity, *Geographical Analysis*, 20, 1-17.

- Anselin L. (1988c) A test for spatial autocorrelation in seemingly unrelated regressions, *Economics Letters*, 28, 335-341.
- Anselin L. (1990a) Spatial dependence and spatial structural instability in applied regression analysis, *Journal of Regional Science*, 30, 185-207.
- Anselin L. (1990b) Some robust approach to testing and estimating in spatial econometrics, *Regional Science and Urban Economics*, 20, 141-163.
- Anselin L. (1992) Spatial dependence and spatial heterogeneity : model specification issues in the spatial expansion paradigm, dans : Casetti E., Jones J.P. III (eds.), *Applications of the Expansion Method*, Routledge, Londres.
- Anselin L. (1995) Local indicators of spatial association-LISA, *Geographical Analysis*, 27, 93-115.
- Anselin L. (1996) The Moran scatterplot as an ESDA tool to assess local instability in spatial association, dans : Fisher M., Scholten H.J., Unwin D. (eds.), *Spatial Analytical Perspectives on GIS*, Taylor & Francis, London.
- Anselin L. (1998a) Interactive techniques and exploratory spatial data analysis, dans : Longley P.A., Goodchild M.F., Maguire D.J., Wind D.W. (eds.), *Geographical Information Systems : Principles, Techniques, Management and Applications*, Wiley, New York.
- Anselin L. (1998b) Exploratory spatial data analysis in a geocomputational environment, dans : Longley P.A., Brooks S.M., McDonnell R., Macmillan B. (eds.), *Geocomputation, a Primer*, Wiley, New York.
- Anselin L. (1999) *SpaceStat, a software package for the analysis of spatial data*, Version 1.90, Ann Arbor, BioMedware.
- Anselin L. (2001a) Rao's score test in spatial econometrics, *Journal of Statistical Planning and Inference*, 97, 113-139.
- Anselin L. (2001b) Spatial econometrics, dans : Baltagi B. (ed.), *Companion to Econometrics*, Basil Blackwell, Oxford.
- Anselin L. (2001c) Spatial externalities, spatial multipliers and spatial econometrics, *Document de Travail*, REAL, University of Illinois of Urbana-Champaign.
- Anselin L., Bao S. (1997) Exploratory spatial data analysis linking SpaceStat and ArcView, dans : Fisher M., Getis A. (eds.), *Recent Developments in Spatial Analysis*, Springer, Berlin.
- Anselin L., Bera A.K. (1998) Spatial dependence in linear regression models with an application to spatial econometrics, dans : Ullah A., Giles D.E.A. (eds.), *Handbook of Applied Economics Statistics*, Springer, Berlin.
- Anselin L., Bera A.K., Florax R.J.G.M., Yoon M.J. (1996) Simple diagnostic tests for spatial dependence, *Regional Science and Urban Economics*, 26, 77-104.
- Anselin L., Cho W.K.T. (2000) Spatial effects and ecological inference, *Document de Travail*, REAL, University of Illinois of Urbana-Champaign.
- Anselin L., Florax R.J.G.M. (1995a) *New Directions in Spatial Econometrics*, Springer, Berlin.

- Anselin L., Florax R.J.G.M. (1995b) Small sample properties of tests for spatial dependence in regression models, dans : 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., Hudak S. (1992) Spatial econometrics in practice, a review of software options, *Regional Science and Urban Economics*, 22, 509-536.
- Anselin L., Kelejian H. (1997) Testing for spatial error autocorrelation in the presence of endogenous regressors, *International Regional Science Review*, 20, 153-182.
- Anselin L., Rey S. (1991) Properties of tests for spatial dependence in linear regression models, *Geographical Analysis*, 23, 112-131.
- Anselin L., Smirnov O. (1996) Efficient algorithms for constructing proper higher order spatial lag operators, *Journal of Regional Science*, 36, 67-89.
- Anselin L., Varga A., Acs Z. (1997) Local geographic spillovers between university research and high technology innovations, *Journal of Urban Economics*, 42, 422-448.
- Arbia G. (1989) *Spatial Data Configuration in Statistical Analysis of Regional Economic and Related Problems*, Kluwer Academic Publishers, Boston.
- Arellano M., Bond S. (1991) Some tests of specification for panel data : Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, 58, 277-297.
- Armstrong H. (1995a) Convergence among the regions of the European union, *Papers in Regional Science*, 74, 143-152.
- Armstrong H. (1995b) An appraisal of the evidence from cross-sectional analysis of the regional growth process within the European Union, dans : Armstrong H., Vickerman R. (eds.), *Convergence and Divergence among European Union*, Pion, Londres.
- Arora S., Brown M. (1977) Alternatives approaches to spatial autocorrelation : an improvement over current practice, *International Regional Science Review*, 2, 67-78.
- Arrow J.K. (1970) The organization of economic activities : issues pertinent to the choice of market versus non-market allocation, dans : Haveman R.H., Margolis J. (eds.), *Public Expenditures and Policy Analysis*, Markham, Chicago.
- Arrow J.K., Debreu G. (1954) Existence of an equilibrium for a competitive economy, *Econometrica*, 22, 265-290.
- Arthur B.W. (1989) Competing technologies, increasing returns and lock-in by historical events, *Economic Journal*, 99, 116-131.
- Aten B. (1996) Evidence of spatial autocorrelation in international prices, *Review of Income and Wealth*, 42, 149-163.
- Aten B. (1997) Does space matter ? International comparisons of the prices of tradables and nontradables, *International Regional Science Review*, 20, 35-52.
- Avery R.B., Hansen L.P., Hotz V.J. (1983) Multiperiod probit models and orthogonality condition estimation, *International Economic Review*, 24, 21-35.
- Azariadis C., Drazen A. (1990) Threshold externalities in economic development, *Quarterly Journal of Economics*, 105, 501-526.

- Azzoni C. (2001) Economic growth and regional income inequality in Brazil, *Annals of Regional Science*, 35, 133-152.
- Bailey T., Gatrell A.C. (1995) *Interactive Spatial Data Analysis*, Longman, Harlow.
- Bairoch P. (1985) *De Jericho à Mexico – Villes et Economie dans l'Histoire*, Gallimard, Paris.
- Bajo-Rubio O. (2000) A further generalization of the Solow growth model : the role of the public sector, *Economics Letters*, 68, 79-84.
- Baldwin R.E., Forslid R. (2000) The core-periphery model and endogenous growth : stabilizing and destabilizing integration, *Economica*, 67, 307-324.
- Baldwin R.E., Martin P., Ottaviano G.I.P. (2001) Global income divergence, trade, and industrialization : the geography of growth take-offs, *Journal of Economic Growth*, 6, 5-37.
- Baller R.D., Anselin L., Messner S.F., Deane G., Hawkins D.F. (2001) Structural covariates of U.S. county homicide rates : incorporating spatial effects, *Criminology*, 39, 561-590.
- Baltagi B.H. (1995) *Econometric Analysis of Panel Data*, Wiley, Chichester.
- Baltagi B.H., Li D. (1998) Double length artificial regressions for testing spatial dependence, *Document de Travail*, Department of Economics, Texas A&M University.
- Baltagi B.H., Li D. (2001) LM tests for functional form and spatial error correlation, *International Regional Science Review*, 24, 194-225.
- Baltagi B.H., Li D. (2002) Prediction in the panel data model with spatial correlation, dans : Anselin L., Florax R.J.G.M., *Advances in Spatial Econometrics*, Springer, Berlin.
- Baltagi B.H., Song S.H., Koh W. (2001) Testing panel data regression models with spatial error correlation, communication, 56th European Meeting of the Econometric Society, Lausanne (Suisse), 25-29 Août.
- Banerjee A. (1993) The economics of rumours, *Review of Economic Studies*, 60, 309-327.
- Barro R.J. (1990) Government spending in simple model of endogenous growth, *Journal of Political Economy*, 98, S103-S125.
- Barro R.J. (1991) Economic growth in a cross-section of countries, *Quarterly Journal of Economics*, 106, 407-444.
- Barro R.J. (1997) *Determinants of Economic Growth. A Cross-Country Empirical Study*, MIT Press, Cambridge.
- Barro R.J., Sala-I-Martin X. (1991) Convergence across states and regions, *Brooking Papers on Economic Activity*, 1, 107-182.
- Barro R.J., Sala-I-Martin X. (1992) Convergence, *Journal of Political Economy*, 100, 223-251.
- Barro R.J., Sala-I-Martin X. (1995) *Economic Growth*, McGraw-Hill, New York.
- Barro R.J., Sala-I-Martin X. (1997) Technological diffusion, convergence and growth, *Journal of Economic Growth*, 2, 1-26.
- Barry B., Pace R.K. (1999) A Monte-Carlo estimator of the log-determinant of large sparse matrices, *Linear Algebra and its applications*, 289, 41-54.
- Basawa I.Y., Prakasa Rao B.L.S. (1980) *Statistical Inference for Stochastic Process*, Londres.

- Basu S., Thibodeau T. (1998) Analysis of spatial autocorrelation in house prices, *Journal of Real Estate Finance and Economics*, 17, 61-85.
- Basu S., Weil D.N. (1998) Appropriate technology and growth, *Quarterly Journal of Economics*, 1025-1054.
- Baumol W.J. (1986) Productivity growth, convergence and welfare : what the long-run data show, *American Economic Review*, 76, 1072-1085.
- Baumol W.J., Oates W.E. (1975) *The Theory of Environmental Policy*, Prentice-Hall, Englewood Cliffs.
- Baumol W.J., Wolff E. (1988) Productivity growth, convergence and welfare : reply, *American Economic Review*, 78, 1155-1159.
- Baumont C. (1997) Croissance endogène et espace, dans : Célimène F., Lacour C. (eds.), *L'Intégration Régionale des Espaces*, Economica, Paris.
- Baumont C. (1998a) Economie, géographie et croissance : quelles leçons pour l'intégration régionale européenne ?, *Revue Française de Géographie*, 5, 35-56.
- Baumont C. (1998b) Economie géographique et intégration régionale. Quels enseignements pour les Pays d'Europe Centrale et Orientale ?, *Document de travail du LATEC*, Dijon, n°9811.
- Baumont C., Ertur C., Le Gallo J. (2002a) 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.
- Baumont C., Ertur C., Le Gallo J. (2002b) Spatial convergence clubs and the European growth process, 1980-1995, dans : Fingleton B. (ed.), *European Regional Growth*, Springer, Berlin, à paraître.
- Baumont C., Guillain R., Huriot J.-M. (1998) Proximités et formation des villes : le rôle des externalités d'information, dans : Huriot J.-M. (ed.), *La Ville ou la Proximité Organisée*, Anthropos, Paris.
- Baumont C., Huriot J.-M. (1995) Agglomération, industrie et ville, *Document de travail du LATEC*, Dijon, n°9508.
- Baumont C., Huriot J.-M. (1999) L'interaction agglomération-croissance en économie géographique, dans : Bailly A., Huriot J.-M. (eds.), *Villes et Croissance : Théories, Modèles, Perspectives*, Anthropos, Paris.
- Baumont C., Huriot J.-M. (2000) Urban economy in retrospect : continuity or change ?, dans : Huriot J.-M., Thisse J.-F. (eds.), *Economics of Cities. Theoretical Perspectives*, Cambridge University Press, Cambridge.
- Bavaud F. (1998) Models for spatial weights : a systematic look, *Geographical Analysis*, 30, 153-171.
- Becker G.S., Murphy K.M., Tamura R. (1990) Human capital, fertility, and economic growth, *Journal of Political Economy*, 98, S12-37.
- Beine M., Docquier F., Hecq A. (1999) Convergence des groupes en Europe : une analyse sur données régionales, *Revue d'Economie Régionale et Urbaine*, 1, 45-62.
- Beine M., Jean-Pierre P. (2000) L'apport des tests de racine unitaire en panel à l'identification des clubs de convergence, dans : Beine M., Docquier F. (eds.), *Croissance et Convergence Economiques des Régions*, De Boeck Université, Bruxelles.

- Bell K.B., Bockstael N.E. (2000) Applying the generalized-moments estimation approach to spatial problems involving microlevel data, *The Review of Economics and Statistics*, 82, 72-82.
- Belsley D., Kuh E., Welsch R. (1980) *Regression Diagnostics : Identifying Influential Data and Sources of Collinearity*, Wiley, New York.
- Benhabib J., Spiegel M. (1994) The role of human capital in economic development. Evidence from cross-country data, *Journal of Monetary Economics*, 34, 143-173.
- Benirschka A., Binkley J. (1996) Land price volatility in a geographically dispersed market, *American Journal of Agricultural Economics*, 76, 185-195.
- Bentivogli C., Pagano P. (1999) Regional disparities and labour mobility : the Euro-11 versus USA, *Labour*, 13, 737-760.
- Bera A.K., Yoon M. (1993) Specification testing with locally misspecified alternatives, *Econometric Theory*, 9, 649-658.
- Berchtold A. (1998) *Chaînes de Markov et Modèles de Transition*, Editions Hermès, Paris.
- Bernard A.B. (1991) Empirical implications of the convergence hypothesis, *CPER Working Paper*, n°23.
- Bernard A.B., Durlauf S.N. (1995) Convergence in international output, *Journal of Applied Econometrics*, 10, 97-108.
- Bernard A.B., Durlauf S.N. (1996) Interpreting tests of the convergence hypothesis, *Journal of Econometrics*, 71, 161-173.
- Bernard A.B., Jones C. (1996) Technology and growth, *Economic Journal*, 106, 1037-1044.
- Bernat G.A. (1996) Does manufacturing matter ? A spatial econometric view of Kaldor's laws, *Journal of Regional Science*, 36, 463-477.
- Beron K.J., Vijverberg W.P.M. (2002) Probit in a spatial context : A Monte Carlo analysis, dans : Anselin L., Florax R.J.G.M. (eds.), *Advances in Spatial Econometric*, Springer, Berlin.
- Bertazzon S., Zanninotto E. (1996) A trend surface analysis of retail location : an Italian case study, *Ricerche Economiche*, 50, 193-208.
- Berthélemy J.-C., Varoudakis A. (1995) Clubs de convergence et croissance : le rôle du développement financier et du capital humain, *Revue Economique*, 46, 217-235.
- Berthélemy J.-C., Varoudakis A. (1996) Economic growth, convergence clubs, and the role of financial development, *Oxford Economic Papers*, 48, 300-328.
- Bianchi M. (1997) Testing for convergence : evidence from nonparametric multimodality tests, *Journal of Applied Econometrics*, 12, 393-409.
- Bishop Y.M., Fienberg S.E., Holland P.W. (1975) *Discrete Multivariate Analysis : Theory and Practice*, MIT Press, Cambridge.
- Bivand R.S. (1984) Regression modeling with spatial dependence : an application of some class selection and estimation methods, *Geographical Analysis*, 16, 25-37.
- Bivand R.S. (1992) Systat compatible software for modeling spatial dependence among observations, *Computers and Geoscience*, 18, 951-963.
- Bliss C. (1999) Galton's fallacy and economic convergence, *Oxford Economic Papers*, 51, 4-14.

- Blommestein H.J. (1983) Specification and estimation of spatial econometric models, *Regional Science and Urban Economics*, 13, 251-270.
- Blommestein H.J. (1985) Elimination of circular routes in spatial dynamic regression equations, *Regional Science and Urban Economics*, 15, 121-130.
- Blommestein H.J., Koper N.A.M. (1992) Recursive algorithms for the elimination of redundant paths in spatial lag operators, *Journal of Regional Science*, 32, 91-111.
- Blommestein H.J., Koper N.A.M. (1997) The influence of sample size on the degree of redundancy in spatial lag operators, *Journal of Econometrics*, 82, 317-333.
- Blommestein H.J., Nijkamp P. (1986) Testing the spatial scale and the dynamic structure in regional models (a contribution to spatial econometric specification analysis), *Journal of Regional Science*, 26, 1-17.
- Blomström M., Lipsey R.E., Zejean M. (1996) Is fixed investment the key to economic growth ?, *Quarterly Journal of Economics*, 111, 269-276.
- Boarnet M.G. (1998) Spillovers and the locational effects of public infrastructure, *Journal of Regional Science*, 38, 381-400.
- Bolduc D., Laferrière R., Santarossa G. (1995) Spatial autoregressive error components in travel flow models : an application to aggregate mode choice, dans : *New Directions in Spatial Econometrics*, Anselin L., Florax R.J.G.M. (eds.), Springer, Berlin.
- Boots B., Tiefelsdorf M. (2000) Global and local spatial autocorrelation in bounded regular tessellations, *Journal of Geographical Systems*, 2, 319-348.
- Bowman A.W. (1984) An alternative method for cross-validation for the smoothing of density estimates, *Biometrika*, 71, 353-360.
- Brackman S., Garretsen H., Gigenback R., Van Marrewijk C., Wagenvoort R. (1996) Negative feedbacks in the economy and industrial location, *Journal of Regional Science*, 36, 631-651.
- Brandsma A.S., Kelletaper R.H. (1979) A biparametric approach to spatial autocorrelation, *Environment and Planning A*, 11, 51-58.
- Braüninger M., Pannenberg M. (2002) Unemployment and productivity growth : an empirical analysis within an augmented Solow model, *Economic Modelling*, 19, 105-120.
- Breinan L., Friedman J., Olshen R., Stone C. (1984) *Classification and Regression Trees*, Wadsworth, Belmont.
- Breusch T. (1980) Useful invariance results for generalized regression models, *Journal of Econometrics*, 13, 327-340.
- Breusch T., Pagan A. (1979) A simple test for heteroskedasticity and random coefficient variation, *Econometrica*, 47, 1287-1294.
- Brock W.A., Durlauf S.N. (2001) Growth empirics and reality, *World Bank Economic Review*, 15, 229-272.
- Bueckner J.K. (1981) A vintage model of urban growth, *Journal of Urban Economics*, 8, 389-402.
- Bueckner J.K. (1986) A switching regression analysis of urban population densities, *Journal of Urban Economics*, 19, 174-189.

- Brueckner J.K. (1998) Testing for strategic interaction among local governments : the case of growth controls, *Journal of Urban Economics*, 44, 438-467.
- Brühlhart M., Torstensson, J. (1996) Regional integration, scale economies and industry location, *CEPR Discussion Paper*, n°1435.
- Brundson C., Fotheringham A.S., Charlton M. (1996) Geographically weighted regression : a method for exploring spatial nonstationarity, *Geographical Analysis*, 28, 281-298.
- Brundson C., Fotheringham A.S., Charlton M. (1998a) Spatial nonstationarity and autoregressive models, *Environment and Planning A*, 30, 957-973.
- Brundson C., Fotheringham A.S., Charlton M. (1998b) Geographically weighted regression-modelling spatial non-stationarity, *The Statistician*, 47, 431-443.
- Brundson C., Fotheringham A.S., Charlton M. (1999) Some notes on parametric significance tests for geographically weighted regression, *Journal of Regional Science*, 39, 497-524.
- Bulli S. (2001) Distribution dynamics and cross-country convergence : a new approach, *Scottish Journal of Political Economy*, 48, 226-243.
- Burridge P. (1980) On the Cliff-Ord test for spatial autocorrelation among regression residuals, *Geographical Analysis*, 4, 267-284.
- Burridge P. (1981) Testing for a common factor in a spatial autoregressive model, *Environment and Planning A*, 13, 795-800.
- Button K., Pentecost E. (1995) Testing for convergence of the EU regional economies, *Economic Inquiry*, 33, 664-671.
- Button K., Pentecost E. (1999) *Regional Economic Performance within the European Union*, Edward Elgar, Cheltenham.
- Calmette M.F., Le Pottier J. (1995) Localisation des activités : un modèle bi-sectoriel avec coûts de transport, *Revue Economique*, 46, 901-909.
- Can A. (1990) The measurement of neighborhood dynamics in urban house prices, *Economic Geography*, 66, 254-272.
- Can A. (1992) Specification and estimation of hedonic housing price models, *Regional Science and Urban Economics*, 22, 453-474.
- Can A., Megboluge I. (1997) Spatial dependence and house price index construction, *Journal of Real Estate Finance and Economics*, 14, 203-222.
- Cappelen A., Fagerberg J., Verspagen B. (1999) Lack of regional convergence, dans : Fagerberg J., Guerrieri P., Verspagen B. (eds.), *The Economic Challenge for Europe : Adapting to Innovation-Based Growth*, Edward Elgar, Aldershot.
- Cappellin R. (1988) Transaction costs and urban agglomeration, *Regional Science and Urban Economics*, 2, 261-278.
- Capron H. (2000) Disparités de croissance entre régions européennes : résorption ou amplification ?, dans : Beine M., Docquier F. (eds.), *Croissance et Convergence Economiques des Régions*, De Boeck Université, Bruxelles.
- Cárdenas M., Pontón A. (1995) Growth and convergence in Colombia : 1950-1990, *Journal of Development Economics*, 47, 5-37.
- 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.
- Casellas A., Galley C. (1999) Regional definitions in the European Union : a question of disparities, *Regional Studies*, 33, 551-558.
- Caselli F., Esquivel G., Lefort F. (1996) Reopening the convergence debate : a new look at cross-country growth empirics, *Journal of Economic Growth*, 1, 363-390.
- Casetti E. (1972) Generating models by the expansion method : applications to geographical research, *Geographical Analysis*, 4, 81-91.
- Casetti E. (1982) Drift analysis of regression parameters : an application to the investigation of fertility development relations, *Modeling and Simulation*, 3, 961-966.
- Casetti E. (1997) The expansion method, mathematical modeling, and spatial econometrics, *International Regional Science Review*, 20, 9-33.
- Casetti E., Can A. (1999) The econometric estimation and testing of DARP models, *Journal of Geographical Systems*, 1, 91-106.
- Casetti E., Jones J.P. III (1988) Spatial parameter variation by orthogonal trend surface expansions : an application to the analysis of welfare program participation rates, *Social Science Research*, 16, 285-300.
- Casetti E., Jones J.P. III (1992) *Applications of the Expansion Method*, Routledge, Londres.
- Chamberlain G. (1982) Multivariate regression models for panel data, *Journal of Econometrics*, 18, 5-46.
- Chamberlain G. (1983) Panel data, dans : Griliches Z., Intriligator M. (eds.), *Handbook of Econometrics*, North Holland, Amsterdam.
- Charlot S. (1999) *Economie Géographique et Croissance Régionale : le Rôle des Infrastructures Publiques*, Thèse en Science Economique, Université de Bourgogne.
- Chatterji M. (1992) Convergence clubs and endogenous growth, *Oxford Review of Economic Policy*, 8, 57-69.
- Chatterji M., Dewhurst J. (1996) Convergence clubs and relative economic performance in Great Britain : 1977-1991, *Regional Studies*, 30, 31-40.
- Chen J., Fleisher B. (1996) Regional income inequality and economic growth in China, *Journal of Comparative Economics*, 22, 141-164.
- Cheshire P.C., Carbonaro G. (1995) Convergence-divergence in regional growth rates : an empty black box ? dans : Armstrong H., Vickerman R. (eds.), *Convergence and Divergence among European Regions*, Londres, Pion.
- Cheshire P.C., Hay D.G. (1989) *Urban Problems in Western Europe : An Economic Analysis*, Routledge, London.
- Cheshire P.C., Magrini S. (2000), Endogenous processes in European regional growth, convergence and policy, *Growth and Change*, 31, 455-479.
- Cho D. (1996) An alternative interpretation of conditional convergence results, *Journal of Money, Credit, and Banking*, 28, 669-681.

- Chou Y.H. (1991) Map resolution and spatial autocorrelation, *Geographical Analysis*, 23, 228-246.
- Chua H.B. (1993) *On Spillovers and Convergence*, Ph.D. Dissertation, Harvard University.
- Chung K.L. (1960) *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag, Berlin.
- Cleveland W.S. (1979) Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, 74, 829-836.
- Cleveland W.S., Devlin S.J. (1988) Locally weighted regression : an approach to regression analysis by local fitting, *Journal of the American Statistical Association*, 83, 596-610.
- Cleveland W.S., Devlin S.J., Grosse E. (1988) Regression by local fitting, methods, properties, and computational algorithms, *Journal of Econometrics*, 37, 87-114.
- Cliff A.D., Ord J.K. (1969) The problem of spatial autocorrelation, dans : Scott A.J. (ed.), *Studies in Regional Science*, Pion, Londres.
- Cliff A.D., Ord J.K. (1971) Evaluating the percentage points of a spatial autocorrelation coefficient, *Geographical Analysis*, 3, 51-62.
- 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.
- Combes P.-P., Duranton G. (2001) Labor pooling, labor poaching, and spatial clustering, *CEP Discussion Paper*, n°510.
- Combes P.-P., Lafourcade M. (2001), Transportation costs decline and regional inequalities : evidence from France, *CEPR Discussion Paper*, n°2894.
- Commission Européenne (1991) *Les Régions dans les Années 90, Quatrième Rapport Périodique sur la Situation et l'Évolution Socio-Economique des Régions dans la Communauté*, Office des publications officielles des communautés européennes, Luxembourg.
- Commission Européenne (1994) *Compétitivité et Cohésion : Tendances dans les Régions, Cinquième Rapport Périodique sur la Situation et l'Évolution Socio-Economique des Régions dans la Communauté*, Office des publications officielles des communautés européennes, Luxembourg.
- Commission Européenne (1996) *Premier Rapport sur la Cohésion Economique et Sociale*, Office des publications officielles des communautés européennes, Luxembourg.
- Commission Européenne (1999) *Sixième Rapport Périodique sur la Situation et le Développement Economique et Social des Régions de l'Union Européenne*, Office des publications officielles des communautés européennes, Luxembourg.
- Commission Européenne (2001) *Unité de l'Europe, Solidarité des Peuples, Diversité des Territoires, Deuxième Rapport sur la Cohésion Economique et Sociale*, Office des publications officielles des Communautés européennes, Luxembourg.
- Cook R. (1977) Detection of influential observations in linear regression, *Technometrics*, 19, 15-18.
- Cressie N. (1993) *Statistics for Spatial Data*, John Wiley, New York.

- Dalgaard C.-J., Vastrup J. (2001) On the measurement of σ -convergence, *Economics Letters*, 70, 283-287.
- Dardanoni V. (1993) Measuring social mobility, *Journal of Economic Theory*, 61, 372-394.
- Davidson R., Mackinnon J.G. (1984) Model specification tests based on artificial linear regressions, *International Economic Review*, 25, 485-502.
- Davidson R., Mackinnon J.G. (1985) Heteroskedasticity-robust tests in regression directions, *Annales de l'INSEE*, 59-60, 183-217.
- Davidson R., Mackinnon J.G. (1988) Double-length artificial regressions, *Oxford Bulletin of Economics and Statistics*, 50, 203-217.
- Davidson R., Mackinnon J.G. (1993) *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.
- De Jong P., Sprenger C., Van Veen F. (1984) On extreme values of Moran's *I* and Geary's *c*, *Geographical Analysis*, 16, 17-24.
- De la Fuente A. (1995) Catch-up, growth and convergence in the OECD, *CEPR Discussion Paper*, n°1274.
- De la Fuente A. (1996) On the sources of convergence : a close look at the Spanish regions, *CEPR Discussion Paper*, n°1543.
- De la Fuente A. (1997) The empirics of growth and convergence : a selective review, *Journal of Economic Dynamics and Control*, 21, 23-74.
- De la Fuente A. (1998) What kind of regional convergence ?, *CEPR Discussion Paper*, n°1924.
- De la Fuente A. (2000) Convergence across countries and regions : theory and empirics, *CEPR Discussion Paper*, n°2465.
- De Long B. (1988) Productivity growth, convergence and welfare : comment, *American Economic Review*, 78, 1138-1154.
- De Long B., Summers L. (1991) Equipment investment and economic growth, *Quarterly Journal of Economics*, 106, 445-502.
- Desdoigts A. (1997) Vers une convergence globale ou locale ?, *Economie Internationale*, 71, 43-74.
- Desdoigts A. (1999) Patterns of economic development and the formation of clubs, *Journal of Economic Growth*, 4, 305-330.
- Des Rosiers F., Thériault M. (1995) Household preferences and the changing structure of residential markets : combining hedonic modeling, GIS and spatial statistics, *Document de travail*, n°1995-32, Faculté des Sciences de l'Administration, Université Laval, Québec.
- Des Rosiers F., Thériault M. (1996) House prices, neighborhood factors and spatial dynamics : improving hedonic modeling using trend surface analysis, *Document de travail*, n°1996-04, Faculté des Sciences de l'Administration, Université Laval, Québec.
- Des Rosiers F., Thériault M. (1999) House prices and spatial dependence : towards an integrated procedure to model neighborhood dynamics, *Document de travail*, n°1999-002, Faculté des Sciences de l'Administration, Université Laval, Québec.

- Dewhurst J.H.L., Mutis-Gaitan H. (1995) Varying speeds of regional GDP per capita convergence in the European Union, 1981-1991, dans : Armstrong H., Vickerman R. (eds.), *Convergence and Divergence among European Regions*, Londres, Pion.
- Dickey D.A., Fuller W.A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 49, 1057-1072.
- Dignan T. (1995) Regional disparities and regional policy in the European Union, *Oxford Review of Economic Policy*, 11, 64-95.
- Dixit A.K., Stiglitz J.E. (1977) Monopolistic competition and optimum product diversity, *American Economic Review*, 67, 297-308.
- Donald S.G., Maddala G.S. (1993) Identifying outliers and influential observations in econometric models, dans : Maddala G.S., Rao C.R., Vinod H.D. (eds.) *Handbook of Statistics II*, Elsevier Science Publishers B.V., North-Holland.
- Dowrick S., Nguyen D.T. (1989) OECD comparative economic growth 1950-85 : catch-up and convergence, *American Economic Review*, 79, 1010-1030.
- Dubin R.A. (1992) Spatial autocorrelation and neighborhood quality, *Regional Science and Urban Economics*, 22, 433-452.
- Dubin R.A. (1995) Estimating logit models with spatial dependence, dans : Anselin L., Florax R.J.G.M. (eds.), *New Directions in Spatial Econometrics*, Springer, Berlin.
- Dubin R.A. (1997) A note on the estimation of spatial logit models, *Geographical Systems*, 4, 181-193.
- Dubin R.A. (1998) Predicting house prices using multiple listings data, *Journal of Real Estate Finance and Economics*, 17, 35-59.
- Dubin R.A., Pace R.K., Thibodeau T.G. (1999) Spatial autoregression techniques for real estate data, *Journal of Real Estate Literature*, 7, 79-95.
- Dunford M., Smith A. (2000) Catching up or falling behind ? Economic performance and regional trajectories in the “New Europe”, *Economic Geography*, 76, 169-196.
- Duranton G. (1997) La nouvelle économie géographique : agglomération et dispersion, *Economie et Prévision*, 131, 1-24.
- Durlauf S.N., Johnson P.A. (1995) Multiple regimes and cross-country growth behaviour, *Journal of Applied Econometrics*, 10, 365-384.
- Durlauf S.N., Quah D. (1999) The new empirics of economic growth, dans : Taylor J., Woodford M. (eds.), *Handbook of Macroeconomics*, North-Holland Elsevier Science.
- Easterly W., Levine R. (1997) Africa’s growth tragedy : policies and ethnic divisions, *Quarterly Journal of Economics*, 112, 1203-1250.
- Easterly W., Levine R. (2001) It’s not factor accumulations : stylized facts and growth models, *World Bank Economic Review*, 15, 177-219.
- Elhorst J.P. (2001) Dynamic models in space and time, *Geographical Analysis*, 33, 119-140.
- Englmann F.C., Walz U. (1995) Industrial centers and regional growth in the presence of local inputs, *Journal of Regional Science*, 35, 3-27.
- Ertur C., Le Gallo J. (2002) An exploratory spatial data analysis of European regional disparities, 1980-1995, dans : Fingleton B. (ed.), *European Regional Growth*, Springer, Berlin, à paraître.

- Esteban J.M., Ray D. (1994) The measurement of polarization, *Econometrica*, 62, 819-852.
- Eurostat (1999) *Regio database, User's Guide, Methods and Nomenclatures*, Office des publications officielles des communautés européennes, Luxembourg.
- Eurostat (2000) *Parités de Pouvoirs d'Achat et Indicateurs Economiques Dérivés*, Office des publications officielles des communautés européennes, Luxembourg.
- Evans P. (1996) Using cross-country variances to evaluate growth theories, *Journal of Economic Dynamics and Control*, 20, 1027-1049.
- Evans P. (1998) Using panel data to evaluate growth theories, *International Economic Review*, 39, 295-306.
- Evans P., Karras G. (1993) Do standards of living converge ? Some cross-country evidence, *Economics Letters*, 43, 149-155.
- Evans P., Karras G. (1996) Convergence revisited, *Journal of Monetary Economics*, 37, 249-265.
- Fagerberg J., Verspagen B. (1996) Heading for divergence ? Regional growth in Europe reconsidered, *Journal of Common Market Studies*, 34, 431-448.
- Faini R. (1983) Cumulative process of deindustrialization in an open region : the case of Southern Italy, *Journal of Development Economics*, 12, 277-301.
- Fingleton B. (1983a) Independence, stationarity, categorical spatial data and the chi-squared test, *Environment and Planning A*, 15, 483-499.
- Fingleton B. (1983b) Log-linear models with dependent spatial data, *Environment and Planning A*, 15, 801-814.
- Fingleton B. (1986) Analyzing cross-classified data with inherent spatial dependence, *Geographical Analysis*, 18, 48-61.
- Fingleton B. (1997) Specification and testing of Markov chain models : an application to convergence in the European Union, *Oxford Bulletin of Economics and Statistics*, 59, 385-403.
- Fingleton B. (1999) Estimates of time to economic convergence : an analysis of regions of the European Union, *International Regional Science Review*, 22, 5-34.
- Fingleton B. (2000) Spatial econometrics, economic geography, dynamics and equilibrium : a third way ?, *Environment and Planning A*, 32, 1481-1498.
- Fingleton B. (2001) Equilibrium and economic growth : spatial econometric models and simulation, *Journal of Regional Science*, 41, 117-147.
- Fingleton B. (2002) Regional economic growth and convergence : insights from a spatial econometric perspective, dans : Anselin L., Florax R.J.G.M., *Advances in Spatial Econometrics*, Springer, Berlin.
- Fingleton B., Lewney R., Pinelli D. (1997) *Regional Growth and Convergence*, dans : Commission Européenne, *The single Market Review*, subseries 6, Aggregate and Regional Impact, Vol. 1, Kogan Page, London.
- Fingleton B., McCombie J.S.L. (1998) Increasing returns and economic growth : some evidence for manufacturing from the European Union regions, *Oxford Economic Papers*, 50, 89-105.

- Florax R.J.G.M., Folmer H. (1992) Specification and estimation of spatial linear regression models, *Regional Science and Urban Economics*, 22, 405-432.
- Florax R.J.G.M., Rey S. (1995) The impacts of misspecified spatial interaction in linear regression models, dans : Anselin L., Florax R.J.G.M. (eds.), *New Directions in Spatial Econometrics*, Springer, Berlin.
- Fotheringham A.S., Brundson C. (1999) Local forms of spatial analysis, *Geographical Analysis*, 31, 340-358.
- Fotheringham A.S., Brundson C., Charlton M. (2000) *Quantitative Geography, Perspectives on Spatial Data Analysis*, Sage Publications, Londres.
- Fotheringham A.S., Charlton M., Brundson C. (1996a) The geography of parameter space : an investigation of spatial non-stationarity, *International Journal of Geographical Information Systems*, 10, 605-627.
- Fotheringham A.S., Charlton M., Brundson C. (1996b) Two techniques for exploring non-stationarity in geographical data, *Geographical Systems*, 28, 281-298.
- Fotheringham A.S., Charlton M., Brundson C. (1997) Measuring spatial variations in relationships with geographically weighted regression, dans : Fisher M.M., Getis A. (eds.), *Recent Developments in Spatial Analysis*, Springer, Berlin.
- Fotheringham A.S., Charlton M., Brundson C. (1998) Geographically weighted regression : a natural evolution of the expansion method for spatial data analysis, *Environment and Planning A*, 30, 1905-1927.
- Fotheringham A.S., Wong D.W.S. (1991) The modifiable areal unit problem in multivariate statistical analysis, *Environment and Planning A*, 23, 1025-1044.
- Frankel J., Romer D. (1996) Trade and growth : an empirical investigation, *NBER Working Paper*, n°5476.
- Friedman M. (1992) Do old fallacies ever die ?, *Journal of Economic Literature*, 30, 2129-2132.
- Fujita M. (1989) *Urban Economic Theory – Land Use and City Size*, Cambridge University Press, Cambridge.
- Fujita M., Thisse J.-F. (1997) Economie géographique, problèmes anciens et nouvelles perspectives, *Annales d'Economie et de Statistique*, 45, 37-87.
- Fujita M., Thisse J.-F. (2000a) Les théories économiques de l'agglomération, dans : Beine M., Docquier F. (eds.), *Croissance et Convergence Economiques des Régions*, De Boeck Université, Bruxelles.
- Fujita M., Thisse J.-F. (2000b) The formation of economic agglomerations, old problems and new perspectives, dans : Huriot J.-M., Thisse J.-F. (eds.), *Economics of Cities. Theoretical Perspectives*, Cambridge University Press, Cambridge.
- Fujita M., Thisse J.-F. (2002a) *Economics of Agglomeration : Cities, Industrial Location, and Regional Growth*, Cambridge University Press, Cambridge.
- Fujita M., Thisse J.-F. (2002b) Does geographical agglomeration foster economic growth ? And who gains and loses from it ?, *CEPR Discussion Paper*, n°3135.
- Fuss C. (1999) Mesures et tests de la convergence : une revue de la littérature, *Revue de l'OFCE*, 69, 221-249.

- Galor O. (1996) Convergence ? Inference from theoretical models, *Economic Journal*, 106, 1056-1069.
- Geary R. (1954) The contiguity ratio and statistical mapping, *The Incorporated Statistician*, 5, 115-145.
- Getis A. (1990) Screening for spatial dependence in regression analysis, *Papers of the Regional Science Association*, 69, 69-81.
- Getis A. (1995) Spatial filtering in a regression framework : examples using data on urban crime, regional inequality, and government expenditures, dans : Anselin L., Florax R.J.G.M. (eds.), *New Directions in Spatial Econometrics*, Springer, Berlin.
- Getis A., Ord J.K. (1992) The analysis of spatial association by use of distance statistics, *Geographical Analysis*, 24, 189-206.
- Geweke J., Marshall R.C., Zarkin G.A. (1986) Mobility indices in continuous time Markov chains, *Econometrica*, 54, 1407-1423.
- Ghio S. (1999a) Intégration économique et concentration spatiale : une revue de la littérature récente, *Région et Développement*, 10, 127-149.
- Ghio S. (1999b) Economies d'agglomération, concentration spatiale et croissance, dans : Baumont C., Combes P.-P., Derycke P.-H., Jayet H. (eds.), *Economie Géographique. Les Théories à l'Épreuve des Faits*, Economica, Paris.
- Gillen K., Thibodeau T., Wachter S. (2001) Anisotropic autocorrelation in house prices, *Journal of Real Estate Finance and Economics*, 23, 5-30.
- Glaeser E.L. (1994) Economic growth and urban density : a review essay, *Document de travail*, n°E-94-7, Hoover Institution, Stanford University.
- Goddard J., Wilson J. (2001) Cross sectional and panel estimation of convergence, *Economics Letters*, 70, 327-333.
- Greene W.H. (2000) *Econometric Analysis*, quatrième édition, Prentice-Hall, Londres.
- Grier K.B., Tullock G. (1989) An empirical analysis of cross-national economic growth, 1951-1980, *Journal of Monetary Economics*, 24, 259-276.
- Griffith D.A. (1978) A spatially adjusted ANOVA model, *Geographical Analysis*, 10, 296-301.
- Griffith D.A. (1980) Towards a theory of spatial statistics, *Geographical Analysis*, 12, 325-339.
- Griffith D.A. (1981) Modelling urban population density in a multi-centered city, *Journal of Urban Economics*, 9, 298-310.
- Griffith D.A. (1983) The boundary value problem in spatial statistical analysis, *Journal of Regional Science*, 23, 377-387.
- Griffith D.A. (1985) An evaluation of correction techniques for boundary effects in spatial statistical analysis : contemporary, *Geographical Analysis*, 17, 81-88.
- Griffith D.A. (1987) Toward a theory of spatial statistics : another step forward, *Geographical Analysis*, 19, 69-82.
- Griffith D.A. (1988a) *Advanced Spatial Statistics, Special Topics in the Exploration of Quantitative Spatial Data Series*, Kluwer Academic Publishers, Dordrecht.

- Griffith D.A. (1988b) Estimating spatial autoregressive model parameters with commercial statistical packages, *Geographical Analysis*, 20, 176-186.
- Griffith D.A. (1992a) A spatially adjusted *N*-way ANOVA model, *Regional Science and Urban Economics*, 22, 347-369.
- Griffith D.A. (1992b) What is spatial autocorrelation ? Reflections on the past 25 years of spatial statistics, *L'Espace Géographique*, 21, 265-280.
- Griffith D.A. (1992c) Simplifying the normalizing factor in spatial autoregressions for irregular lattices, *Papers in Regional Science*, 71, 71-86.
- Griffith D.A. (1993) *Spatial Regression on the PC : Spatial Statistics using SAS*, Association of the American Geographers, Washington, DC.
- Griffith D.A. (1995) Some guidelines for specifying the geographic weights matrix contained in spatial statistics, dans : Arlinghaus S.L. (ed.), *Practical Handbook of Spatial Statistics*, CRC Press, Boca Raton.
- Griffith D.A. (1999) Statistical and mathematical sources of regional science theory : map pattern analysis as an example, *Papers in Regional Science*, 78, 21-45.
- Griffith D.A., Amrhein C.G. (1983) An evaluation of correction techniques for boundary effects in spatial statistical analysis : traditional methods, *Geographical Analysis*, 15, 352-360.
- Griffith D.A., Can A. (1996) Spatial Statistical/Econometric Versions of Simple Urban Population Density Models, dans : Griffith D.A., Arlinghaus S.L. (eds.), *Practical Handbook of Spatial Statistics*, Boca Raton, CRC Press.
- Griffith D.A., Layne L.J. (1999) *A Casebook for Spatial Statistical Data Analysis*, University Press Oxford, Oxford.
- Griffith D.A., Sone A. (1995) Trade-offs associated with normalizing constant computational simplifications for estimating spatial statistical models, *Journal of Statistical Computation and Simulation*, 51, 165-183.
- Grossman G.M., Helpman E. (1991) *Innovation and Growth in the Global Economy*, MIT Press, Cambridge.
- Guillain R. (2000) *Villes et Echanges d'Informations*, Thèse en Science Economique, Université de Bourgogne.
- Guillain R., Huriot J.-M. (2001) The local dimension of information spillovers : a critical review of empirical evidence in the case of innovation, *Canadian Journal of Regional Science*, 24, 294-319.
- Haining R. (1977) Model specification in stationary random fields, *Geographical Analysis*, 9, 107-129.
- 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.
- Hall P., Hay D. (1980) *Growth Centres in the European Urban System*, Heinemann Educational Book, Londres.

- Hall R., Jones C. (1999) Why do some countries produce so much output per worker than others, *Quarterly Journal of Economics*, 114, 83-116.
- Hanson G.H. (1998a) North american economic integration and industry location, *Oxford Review of Economic Policy*, 14, 30-44.
- Hanson G.H. (1998b) Market potential, increasing returns, and geographic concentration, *NBER Working Paper*, n°6429.
- Hart P.E. (1995) Galtonian regressions across countries and the convergence of productivity, *Oxford Bulletin of Economics and Statistics*, 57, 287-293.
- Harvey A. (1976) Estimating regression models with multiplicative heteroskedasticity, *Econometrica*, 47, 387-406.
- Heijmans R.D.H, Magnus J.R. (1986a) Asymptotic normality of maximum likelihood estimators obtained from normally distributed but dependent observations, *Econometric Theory*, 12, 374-412.
- Heijmans R.D.H, Magnus J.R. (1986b) Consistent maximum-likelihood estimation with dependent observations : the general (non-normal) case and the normal case, *Journal of Econometrics*, 32, 253-285.
- Helpman H. (1996) The size of regions, *Document de travail*, Foerder Institute for Economic Research, 14-95, Université de Tel-Aviv.
- Henderson J.V. (1987) Systems of cities and inter-city trade, dans : Hansen *et al.* (eds.), *Systems of Cities and Facility Location*, Harwood Academic Publishers, Chur.
- Henderson J.V., Shalizi Z., Venables A.J. (2001) Geography and development, *Journal of Economic Geography*, 1, 81-105.
- Hénin P.-Y., Le Pen Y. (1995) Les épisodes de la convergence européenne, *Revue Economique*, 46, 667-677.
- Henry M.S., Schmitt B., Piguet V. (2001) Spatial econometric models for simultaneous systems: application to rural community growth in France, *International Regional Science Review*, 24, 171-193.
- Hepple L. (1995a) Bayesian techniques in spatial and network econometrics : 1. Model comparison and posterior odds, *Environment and Planning A*, 27, 447-469.
- Hepple L. (1995b) Bayesian techniques in spatial and network econometrics : 2. Computational methods and algorithms, *Environment and Planning A*, 27, 615-644.
- Hepple L. (1998) Exact testing for spatial autocorrelation among regression residuals, *Environment and Planning A*, 30, 85-108.
- Hildreth C., Houck J. (1968) Some estimators for a linear model with random coefficients, *Journal of the American Statistical Association*, 63, 584-595.
- Hirschman A. (1958) *The Strategy of Economic Development*, CT Yale University Press, New Haven.
- Hoaglin D., Welsch R. (1978) The hat matrix in regression and ANOVA, *The American Statistician*, 32, 17-22.
- Hohenberg P.M., Lees L.H. (1985) *The Making of Urban Europe, 1000-1950*, Harvard University Press, Cambridge.

- Hoover E.M. (1936) *Location Theory and the Shoe and Leather Industry*, Harvard University Press, Cambridge.
- Hordijk L. (1979) Problems in estimating econometric relations in space, *Papers of the Regional Science Association*, 42, 99-115.
- Hordijk L., Nijkamp P. (1977) Dynamic models of spatial autocorrelation, *Environment and Planning A*, 9, 505-519.
- Howitt P. (2000) Endogenous growth and cross-country income differences, *American Economic Review*, 90, 829-846.
- Hsiao C. (1986) *Analysis of Panel Data*, Cambridge University Press, Cambridge.
- Huang J.S. (1984) The autoregressive moving average model for spatial analysis, *Australian Journal of Statistics*, 26, 169-178.
- Hubert L.J., Arabie P. (1991) The assessment of spatial autocorrelation through constrained multiple regression, *Geographical Analysis*, 23, 95-111.
- Hubert L.J., Golledge R.J., Costanzo C.M. (1981) Generalized procedures for evaluating spatial autocorrelation, *Geographical Analysis*, 13, 224-233.
- Hubert L.J., Golledge R.J., Costanzo C.M., Gale N. (1985) Measuring association between spatially defined variables : an alternative procedure, *Geographical Analysis*, 17, 36-46.
- IAURIF (2001) 40 ans en Ile-de-France. Rétrospective 1960-2000, *Etudes et Documents*, Paris IAURIF.
- Inada K. (1963) On a two-sector model of economic growth : comments and a generalization, *Review of Economic Studies*, 30, 119-127.
- Islam N. (1995) Growth empirics : a panel data approach, *Quarterly Journal of Economics*, 110, 1127-1170.
- Islam N. (1998a) Convergence : variation in concept and empirical results, *mimeo*, Department of Economics, Emory University.
- Islam N. (1998b) Growth empirics : a panel data approach – a reply, *Quarterly Journal of Economics*, 113, 325-329.
- Izenman A.J. (1991) Recent developments in nonparametric density estimation, *Journal of the Japanese and International Economics*, 13, 61-72.
- Jarque C.M., Bera A.K. (1987) A test for normality of observations and regression residuals, *International Statistical Review*, 55, 163-172.
- Jayet H. (1993) *Analyse Spatiale Quantitative*, Economica, Paris.
- Jayet H. (1997) Quels modèles asymptotiques en statistique et en économétrie spatiales ?, *L'Espace Géographique*, 2, 165-172.
- Jean-Pierre P. (1997) Sélection et tests de seuils de convergence, *Revue Economique*, 48, 429-440.
- Jean-Pierre P. (1999) La convergence régionale européenne : une approche empirique par les clubs et les panels, *Revue d'Economie Régionale et Urbaine*, 1, 21-44.
- Johansen S., Juselius K. (1990) Maximum likelihood estimation and inference on cointegration with applications to the demand of money, *Oxford Bulletin of Economic and Statistics*, 52, 169-210.

- Johnson M.S., Ragas W.R. (1987) CBD land values and multiple externalities, *Land Economics*, 63, 335-347.
- Johnson P.A. (2000) A nonparametric analysis of income convergence across the US states, *Economics Letters*, 69, 219-223.
- Jones C.I. (1995a) R&D-based models of economic growth, *Journal of Political Economy*, 103, 759-784.
- Jones C.I. (1995b) Time-series tests of endogenous growth models, *Quarterly Journal of Economics*, 110, 495-525.
- Jones C.I. (1997) On the evolution of the world income distribution, *Journal of Economic Perspectives*, 11, 19-36.
- Jones C.I. (1999) Growth : with or without scale effects, *AEA Papers and Proceedings*, 89, 139-144.
- Jones J.P. III (1983) Parameter variation via the expansion method with tests for autocorrelation, *Modeling and Simulation*, 14, 853-857.
- Jones L., Manuelli R. (1990) A convex model of equilibrium growth : theory and policy implications, *Journal of Political Economy*, 98, 1008-1038.
- Kangasharju A. (1999) Relative economic performance in Finland : regional convergence, 1934-1993, *Regional Studies*, 33, 207-217.
- Kawagoe M. (1999) Regional dynamics in Japan : a reexamination of Barro regressions, *Journal of the American Statistical Association*, 86, 205-224.
- Kelejian H.H., Prucha I.R. (1997) Estimation of spatial regression models with autoregressive errors by two-stage least squares procedures : a serious problem, *International Regional Science Review*, 20, 103-111.
- Kelejian H.H., Prucha I.R. (1998) A generalized spatial least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances, *Journal of Real Estate Finance and Economics*, 17, 99-121.
- Kelejian H.H., Prucha I.R. (1999) A generalized moments estimator for the autoregressive parameter in a spatial model, *International Economic Review*, 40, 509-534.
- Kelejian H.H., Prucha I.R. (2001) On the asymptotic distribution of the Moran *I* test statistic with applications, *Journal of Econometrics*, 104, 219-257.
- Kelejian H.H., Robinson D.P. (1992) Spatial autocorrelation : a new computationally simple test with an application to per capita country police expenditures, *Regional Science and Urban Economics*, 22, 317-331.
- Kelejian H.H., Robinson D.P. (1993) A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure country police expenditure, *Papers in Regional Science*, 72, 297-312.
- Kelejian H.H., Robinson D.P. (1995) Spatial correlation : a suggested alternative to the autoregressive model, dans : Anselin L., Florax R.J.G.M. (eds.), *New Directions in Spatial Econometrics*, Springer, Berlin.
- Kelejian H.H., Robinson D.P. (1997) Infrastructure productivity estimation and its underlying econometric specifications : a sensitivity analysis, *Papers in Regional Science*, 76, 115-131.

- Kelejian H.H., Robinson D.P. (1998) A suggested test for spatial autocorrelation and/or heteroskedasticity and corresponding Monte Carlo results, *Regional Science and Urban Economics*, 28, 389-417.
- Kemeny J., Snell L. (1976) *Finite Markov Chains*, Springer-Verlag, New York.
- 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., Rodríguez-Clare A. (1997a) Economic growth : a review essay, *Journal of Monetary Economics*, 40, 597-618.
- Klenow P.J., Rodríguez-Clare A. (1997b) The neoclassical revival in growth economics : has it gone too far ?, *NBER Macroeconomics Annual 1997*, 73-102.
- Knight M., Loyaza N., Villanueva D. (1993) Testing for neoclassical theory of economic growth, *IMF Staff Papers*, 40, 512-541.
- Knowles S., Owen D.P. (1995) Health capital and cross-country variation in income per capita in the Mankiw-Romer-Weil model, *Economics Letters*, 48, 99-106.
- Kocherlakota N., Yi K. (1995) Can convergence regressions distinguish between exogenous and endogenous growth models ?, *Economics Letters*, 49, 211-215.
- Kocherlakota N., Yi K. (1996) A simple time series test of endogenous vs. exogenous growth models : an application to the United States, *Review of Economics and Statistics*, 78, 126-134.
- Kocherlakota N., Yi K. (1997) Is there endogenous long run growth ? Evidence from the U.S. and the U.K., *Journal of Money, Credit and Banking*, 29, 235-262.
- Koenker R. (1981) A note on studentizing a test for heteroskedasticity, *Journal of Econometrics*, 17, 107-112.
- Koenker R., Basset G. (1982) Robust tests for heteroskedasticity based on regression quantiles, *Econometrica*, 50, 43-61.
- Kortum S. (1997) Research, patenting, and technological change, *Econometrica*, 65, 1389-1419.
- Krugman P. (1981) Trade, accumulation, and uneven development, *Journal of Development Economics*, 8, 149-161.
- Krugman P. (1991a) Increasing returns and economic geography, *Journal of Political Economy*, 99, 483-499.
- Krugman P. (1991b) *Geography and Trade*, MIT Press, Cambridge.
- Krugman P. (1991c) History versus expectations, *Quarterly Journal of Economics*, 106, 651-667.
- Krugman P. (1993a) First nature, second nature and metropolitan location, *Journal of Regional Science*, 33, 129-144.
- Krugman P. (1993b) On the relationship between trade theory and location theory, *Review of International Economics*, 1, 110-122.
- Krugman P. (1999) The role of geography in development, *International Regional Science Review*, 22, 142-161.
- Krugman P., Venables A.J. (1995) Globalization and the inequality of nations, *Quarterly Journal of Economics*, 110, 857-880.

- Krugman P., Venables A.J. (1996) Integration, specialization and adjustment, *European Economic Review*, 40, 959-967.
- Kubo Y. (1995) Scale economies, regional externalities, and the possibility of uneven development, *Journal of Regional Science*, 35, 29-42.
- Kullback S., Kupperman M., Ku H.H. (1962) Tests for contingency tables and Markov chains, *Technometrics*, 4, 573-608.
- Lahiri S.N. (1996) On inconsistency of estimators based on spatial data under infill asymptotics, *Sankhya*, 58A, 403-417.
- Land K., Deane G. (1992) On the large sample estimation of regression models with spatial - or network - effects terms : a two stage least squares approach, dans : Marsden P. (ed.), *Sociological Methodology*, Josey Bass, San Francisco.
- Leamer E.E. (1978) *Specification Searches : Ad Hoc Inferences From Non-Experimental Data*, Wiley, New York.
- Leamer E.E. (1983) Let's take the con out of econometrics, *American Economic Review*, 73, 31-43.
- Leamer E.E. (1985) Sensitivity analysis would help, *American Economic Review*, 75, 308-313.
- Leamer E.E., Leonard H. (1983) Reporting the fragility of regression estimates, *Review of Economics and Statistics*, 65, 306-317.
- Lebart L. (1969) *Les variables socio-économiques départementales et régionales*, Publication de l'ISUP.
- Lee K., Pesaran M., Smith R. (1997) Growth and convergence in a multi-country empirical stochastic Solow model, *Journal of Applied Econometrics*, 12, 357-392.
- Lee K., Pesaran M., Smith R. (1998) Growth empirics : a panel data approach – a comment, *Quarterly Journal of Economics*, 113, 319-323.
- Le Gallo J. (2001a) Économétrie spatiale, 2. Hétérogénéité spatiale, *Document de travail du LATEC*, Dijon, n°2001-01 (Janvier 2001).
- Le Gallo J. (2001b) Space-time analysis of GDP disparities among European regions : a Markov chains approach, *Document de travail du LATEC*, Dijon, n°2001-06 (Mars 2001).
- Le Gallo J. (2002) Économétrie spatiale : l'autocorrélation spatiale dans les modèles de régression linéaire, *Economie et Prévision*, à paraître.
- Le Gallo J., Ertur C. (2002) Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995, *Papers in Regional Science*, à paraître.
- Le Gallo J., Ertur C., Baumont C. (2002) A spatial econometric analysis of convergence across European regions, 1980-1995, dans : Fingleton B. (ed.), *European Regional Growth*, Springer, Berlin, à paraître.
- Legendre P., Oden N.L., Sokal R.R., Vaudor A., Kim J. (1990) Approximate analysis of variance of spatially autocorrelated regional data, *Journal of Classification*, 7, 53-75.
- Le Pen Y. (1997) Convergence internationale des revenus par tête : un tour d'horizon, *Revue d'Economie Politique*, 107, 715-756.

- LeSage J.P. (1997) Bayesian estimation of spatial autoregressive models, *International Regional Science Review*, 20, 113-129.
- LeSage J.P. (1999a) *Spatial Econometrics*, WebBook of Regional Science.
- LeSage J.P. (1999b) Bayesian estimation of limited dependent variable spatial autoregressive models, *Geographical Analysis*, 32, 19-35.
- Leung Y., Mei C., Zhang W. (2000) Statistical tests for spatial non-stationarity based on the geographically weighted regression model, *Environment and Planning A*, 32, 9-32.
- Levine R., Revelt D. (1992) A sensitivity analysis of cross-country growth regressions, *American Economic Review*, 82, 942-963.
- Levine R., Zervos S. (1993) What we have learned about policy and growth from cross-country regressions, *American Economic Association Papers and Proceedings*, 83, 426-430.
- Li B. (1996) Implementing spatial statistics on parallel computers, dans : Arlinghaus S. (ed.), *Practical Handbook of Spatial Statistics*, CRC Press, Boca Raton.
- López-Bazo E., Vayá E., Mora A.J., Suriñach J. (1999) Regional economic dynamics and convergence in the European Union, *Annals of Regional Science*, 33, 343-370.
- Lordon F. (1991) Théories de la croissance : quelques développements récents, 2^{ème} partie, *Observations et Diagnostics Economiques*, 37, 193-239.
- Lucas R. (1988) On the mechanisms of economic development, *Journal of Monetary Economics*, 22, 3-42.
- Maasoumi E. (1998) On mobility, dans : Ullah A., Giles D. (eds.), *Handbook of Applied Economic Statistics*, Marcel Dekker, New York..
- Macedo P.B.R. (1998) Hedonic price models with spatial effects : an application to the housing market of Belo Horizonte, Brazil, *Revista Brasileira de Economia*, 52, 63-81.
- Maddison A. (1982) *Phases of Capitalist Development : A Long-Run Comparative View*, Oxford University Press, Oxford.
- Magnus, J. (1978) Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix, *Journal of Econometrics*, 7, 281-312 (Corrigenda, *Journal of econometrics*, 10, 261).
- Magrini S. (1999) The evolution of income disparities among the regions of the European Union, *Regional Science and Urban Economics*, 29, 257-281.
- Mankiw N.G. (1995) The growth of nations, *Brooking Papers on Economic Activity*, 1, 275-310.
- Mankiw N.G., Romer D., Weil D.N. (1992) A contribution to the empirics of economic growth, *Quarterly Journal of Economics*, 107, 407-437.
- Martin P.J., Ottaviano G.I.P. (1999) Growing locations : industry location in a model of endogenous growth, *European Economic Review*, 43, 281-302.
- Martin P.J., Ottaviano G.I.P. (2001) Growth and agglomeration, *International Economic Review*, 42, 947-968.
- Martin P.J., Rogers C.A. (1995) Industrial location and public infrastructure, *Journal of International Economics*, 39, 335-351.

- Martin R. (1999) The new "geographical turn" in economics : some critical reflections, *Cambridge Journal of Economics*, 23, 65-91.
- Martin R. (2001) EMU versus the regions ? Regional convergence and divergence in Euroland, *Journal of Economic Geography*, 1, 51-80.
- Martin R., Sunley P. (1998) Slow convergence ? The new endogenous growth theory and regional development, *Economic Geography*, 74, 201-227.
- Martin R.J. (1993) Approximations to the determinant term in gaussian maximum likelihood estimation of some spatial models, *Communications in Statistics : Theory and Methods*, 22, 189-205.
- Mas M., Pérez F., Uriel E. (1995) Growth and convergence in the spanish provinces, dans : Armstrong H., Vickerman R. (eds.), *Convergence and Divergence among European Union*, Pion, Londres.
- Mathsoft (1996) *S+SpatialStats. User's Manual for Windows and Unix*, Mathsoft, Inc, Seattle, WA.
- Matsuyama K. (1995) Complementarities and cumulative processes in models of monopolistic competition, *Journal of Economic Literature*, 33, 701-729.
- Maurseth P.B. (2001) Convergence, geography and technology, *Structural Change and Economic Dynamics*, 12, 247-276.
- McCallum B.T. (1996) Neoclassical vs. endogenous growth analysis : an overview, *NBER Working Paper*, n°5844.
- McDonald S., Roberts J. (2002) Growth and multiple forms of human capital in an augmented Solow model : a panel data investigation, *Economics Letters*, 271-276.
- McKinnon J., White H. (1985) Some heteroskedasticity-consistent covariance matrix estimators with improved finite samples properties, *Journal of Econometrics*, 29, 305-325.
- McMillen D.P. (1992) Probit with spatial autocorrelation, *Journal of Regional Science*, 32, 335-348.
- McMillen D.P. (1994) Vintage growth and population density : an empirical investigation, *Journal of Urban Economics*, 36, 333-352.
- McMillen D.P. (1995a) Spatial effects in probit models, dans : Anselin L., Florax R.J.G.M. (eds.), *New Directions in Spatial Econometrics*, Springer, Berlin.
- McMillen D.P. (1995b) Selection bias in spatial econometric models, *Journal of Regional Science*, 35, 417-436.
- McMillen D.P. (1996) One hundred fifty years of land values in Chicago : a nonparametric approach, *Journal of Urban Economics*, 40, 100-124.
- McMillen D.P., McDonald J.F. (1997) A nonparametric analysis of employment density in a polycentric city, *Journal of Regional Science*, 37, 591-612.
- McMillen D.P., McDonald J.F. (1998) Population density in Chicago : a bid rent approach, *Urban Studies*, 7, 1119-1130.
- Molle W. (1980) *Regional Disparity and Economic Development in the European Community*, Westmead, Saxon House.

- Montfort P., Nicolini R. (2000) Regional convergence and international integration, *Journal of Urban Economics*, 48, 286-306.
- Monfort P., Ottaviano G.I.P. (2000) Local labor markets, skill accumulation and regional disparities, mimeo, IRES, Université Catholique de Louvain, Louvain-La-Neuve.
- Moore M. (1988) Spatial linear models, *Communications in Statistics, Stochastic models*, 4, 45-75.
- Moran P. (1948) The interpretation of statistical maps, *Journal of the Royal Statistical Society*, 10B, 243-251.
- Moran P. (1950a) Notes on continuous stochastic phenomena, *Biometrika*, 37, 17-23.
- Moran P. (1950b) A test for serial independence of residuals, *Biometrika*, 37, 178-181.
- Moreno R., Artís M., López-Bazo J., Suriñach J. (1998) Evidence on the complex link between infrastructure and regional growth, *International Journal of Development Planning Literature*, 12, 81-108.
- Moreno R., Trehan B. (1997) Location and the growth of nations, *Journal of Economic Growth*, 2, 399-418.
- Mulligan C.B., Sala-I-Martin X. (1993) Transitional dynamics in two-sector models of endogenous growth, *Quarterly Journal of Economics*, 108, 739-773.
- Mur J. (1999) Testing for spatial autocorrelation : moving average versus autoregressive process, *Environment and Planning A*, 31, 1371-1382.
- Muth R.F. (1969) *Cities and Housing*, The University of Chicago Press, Chicago.
- Myrdal G. (1957) *Economic Theory and Underdeveloped Regions*, Duckworth, Londres.
- Neary J.P. (2001) Of hype and hyperbolas : introducing the new economic geography, *Journal of Economic Literature*, 39, 536-561.
- Neven D., Gouyette C. (1994) European integration and regional growth, *Revue Economique*, 3, 703-713.
- Neven D., Gouyette C. (1995) Regional convergence in the European Community, *Journal of Common Market Studies*, 33, 47-65.
- Nonneman W., Vanhoudt P. (1996) A further augmentation of the Solow model and the empirics of economic growth for OECD countries, *Quarterly Journal of Economics*, 111, 943-953.
- Oden N.L. (1984) Assessing the significance of a spatial correlogram, *Geographical Analysis*, 16, 1-16.
- Odland J. (1988) *Spatial Autocorrelation*, Scientific Geography Series, Volume 9, SAGE Publications, Newbury Park.
- Ogawa H., Fujita M. (1980) Equilibrium land use patterns in a nonmonocentric city, *Journal of Regional Science*, 4, 455-475.
- Ogawa H., Fujita M. (1989) Nonmonocentric urban configuration in a two-dimensional space, *Environment and Planning A*, 21, 363-374.
- Olmo J.C. (1995) Spatial estimation of housing prices and locational rents, *Urban Studies*, 32, 1331-1344.

- Openshaw S., Taylor P.J. (1979) A million or so correlation coefficients : three experiments on the modifiable areal unit problem, dans : Wrigley N. (eds.), *Statistical Applications in the Spatial Sciences*, Pion, Londres.
- Ord J.K. (1975) Estimation methods for models of spatial interaction, *Journal of the American Statistical Association*, 70, 120-126.
- Ord J.K., Getis A. (1995) Local spatial autocorrelation statistics : distributional issues and an application, *Geographical Analysis*, 27, 286-305.
- Ottaviano G.I.P., Puga D. (1997) L'agglomération dans l'économie mondiale, *Economie Internationale*, 71, 75-100.
- Ottaviano G.I.P., Puga D. (1998) Agglomeration in the global economy : a survey of the "new economic geography", *World Economy*, 21, 707-731.
- Ottaviano G.I.P., Tabuchi T., Thisse J.-F. (2002) Agglomeration and trade revisited, *International Economic Review*, 43, 101-127.
- Paap R., Van Dijk H. (1998) Distribution and mobility of the wealth of nations, *European Economic Review*, 42, 1269-1293.
- Pace R.K. (1997) Performing large spatial regressions and autoregressions, *Economics Letters*, 54, 283-291.
- Pace R.K., Barry R. (1997a) Sparse spatial autoregressions, *Statistics and Probability Letters*, 33, 291-297.
- Pace R.K., Barry R. (1997b) Quick computation of spatial autoregressive estimators, *Geographical Analysis*, 29, 232-246.
- Pace R.K., Barry R. (1998) Simulating mixed regressive spatially autoregressive estimators, *Computational Statistics*, 13, 397-418.
- Pace R.K., Barry R., Sirmans C.F. (1998) Spatial Statistics and Real Estate, *Journal of Real Estate Finance and Economics*, 17, 5-13.
- Pace R.K., Gilley O.W. (1997) Using the spatial configuration of the data to improve estimation, *Journal of Real Estate Finance and Economics*, 14, 333-340.
- Pace R.K., Zou D. (2000) Closed-form likelihood estimates of nearest neighbor spatial dependence, *Geographical Analysis*, 32.
- Paci R., Pigliaru F. (1999) Technological catch-up and regional convergence in Europe, *Document de travail*, CRENOS, Université de Cagliari.
- Paci R., Pigliaru F. (2001) Technological diffusion, spatial spillovers and regional convergence in Europe, *Document de travail*, CRENOS, Université de Cagliari.
- Pack H. (1997) Endogenous growth theory : intellectual appeal and empirical shortcomings, *Journal of Economic Perspectives*, 8, 55-72.
- Paelinck J.H.P., Klaassen L.H. (1979) *Spatial Econometrics*, Saxon House, Farnborough.
- Papageorgiou Y.Y., Smith T.R. (1983) Agglomeration as local instability of spatially uniform steady-states, *Econometrica*, 51, 1109-1119.
- Pavlov A.D. (2000) Space-varying regression coefficients : a semi-parametric approach applied to real estate markets, *Real Estate Economics*, 2, 249-283.

- Pereira J., Carreiras J., Perestrello de Vasconcelos M. (1998) Exploratory data analysis of the spatial distribution of wildfires in Portugal, 1980-1989, *Geographical Systems*, 5, 355-390.
- Perron P. (1989) The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, 57, 1361-1401.
- Persson J. (1997) Convergence across the Swedish counties, 1911-1993, *European Economic Review*, 41, 1835-1852.
- Petrakos G., Saratsis Y. (2000) Regional inequalities in Greece, *Papers in Regional Science*, 79, 57-74.
- Phillips P.C.B., Perron P. (1988) Testing for a unit root in time series regression, *Biometrika*, 75, 347-353.
- Pinkse J. (1999) Asymptotic properties of the Moran and related tests and a test for spatial correlation in probit models, *Document de Travail*, Department of Economics, University of British Columbia, Vancouver, BC.
- Pinkse J. (2002) Moran-flavoured tests with nuisance parameters, dans : Anselin L., Florax R.J.G.M. (Eds.), *Advances in Spatial Econometrics*, Springer, Heidelberg.
- Pinkse J., Slade M.E. (1998) Contracting in space : an application of spatial statistics to discrete-choice models, *Journal of Econometrics*, 85, 125-154.
- Poirier D.J., Ruud P.A. (1988) Probit with dependent observations, *Review of Economic Studies*, 55, 593-614.
- Puga D. (1999) The rise and fall of regional inequalities, *European Economic Review*, 43, 303-334.
- Puga D. (2001) European regional policies in light of recent location theories, *CEPR Discussion Paper*, n°2767.
- Puga D., Venables A.J. (1997) Preferential trading arrangements and industrial location, *Journal of International Economics*, 43, 347-368.
- Puga D., Venables A.J. (1998) Trading arrangements and industrial development, *CEP Discussion Paper*, n°319.
- Puga D., Venables A.J. (1999) Agglomeration and economic development : import substitution versus trade liberalisation, *Economic Journal*, 109, 293-311.
- Quah D. (1993a) Empirical cross-section dynamics in economic growth, *European Economic Review*, 37, 426-434.
- Quah D. (1993b) Galton's fallacy and tests of the convergence hypothesis, *Scandinavian Journal of Economics*, 94, 427-443.
- Quah D. (1996a) Empirics for economic growth and convergence, *European Economic Review*, 40, 1353-1375.
- Quah D. (1996b) Regional convergence clusters across Europe, *European Economic Review*, 40, 951-958.
- Quah D. (1996c) Convergence empirics across economics with (some) capital mobility, *Journal of Economic Growth*, 1, 95-124.
- Quah D. (1996d) Aggregate and regional disaggregate fluctuations, *Empirical Economics*, 21, 137-159.

- Quah D. (1996e) Twin peaks : growth and convergence in models of distribution dynamics, *Economic Journal*, 106, 1045-1055.
- Quah D. (1996f) Convergence as distribution dynamics (with or without growth), *CEP Discussion Paper*, n°317 (Novembre 1996).
- Quah D. (1997) Empirics for growth and distribution : stratification, polarization, and convergence clubs, *Journal of Economic Growth*, 2, 27-59.
- Quandt R. (1958) The estimation of the parameters of a linear regression system obeying two separate regimes, *Journal of the American Statistical Association*, 53, 873-880.
- Rao C.R. (1947) Large sample tests to statistical hypotheses concerning several parameters with applications to problems of estimation, *Proceedings of the Cambridge Philosophical Society*, 44, 50-57.
- Rappaport J. (2000) Is the speed of convergence constant ? *Document de Travail*, Research Division, Federal Reserve Bank of Kansas City.
- Rauch J.E. (1993) Productivity gains from geographic concentration of human capital : evidence from the cities, *Journal of Urban Economics*, 34, 380-400.
- Rebelo S. (1991) Long-run policy analysis and long-run growth, *Journal of Political Economy*, 99, 500-521.
- Rey S.J. (2001) Spatial empirics for economic growth and convergence, *Geographical Analysis*, 33, 195-214.
- Rey S.J., Boarnet M.G. (2002) A taxonomy of spatial econometric models for simultaneous equations systems, dans : Anselin L., Florax R.J.G.M., *Advances in Spatial Econometrics*, Springer, Berlin.
- Rey S.J., Montouri B.D. (1999) U.S. regional income convergence : a spatial econometric perspective, *Regional Studies*, 33, 145-156.
- Rietveld P., Wintershoven P. (1998) Border effects and spatial autocorrelation in the supply of network infrastructure, *Papers in Regional Science*, 77, 265-276.
- Ripley B. (1981) *Spatial Statistics*, New York, Wiley.
- Rodríguez-Posé A. (1998) *The Dynamics of Regional Growth in Europe*, Clarendon Press, Oxford.
- Romer P. (1986) Increasing returns and long run growth, *Journal of Political Economy*, 94, 1002-1037.
- Romer P. (1990) Endogenous technological change, *Journal of Political Economy*, 98, S71-S102.
- Romer P. (1994) The origins of endogenous growth, *Journal of Economic Perspectives*, 8, 3-22.
- Rousseeuw P.J., Leroy A.M. (1987) *Robust Regression and Outlier Detection*, Wiley, New York.
- Saavedra L. (2000), Test for spatial lag dependence based on method of moments estimation, communication, Econometric Society World Conference, Seattle (USA), 11-16 Août.
- Sala-I-Martin X. (1994) Cross-sectional regressions and the empirics of economic growth, *European Economic Review*, 38, 739-747.

- Sala-I-Martin X. (1996a) Regional cohesion : evidence and theories of regional growth and convergence, *European Economic Review*, 40, 1325-1352.
- Sala-I-Martin X. (1996b) The classical approach to convergence analysis, *Economic Journal*, 106, 1019-1036.
- Sala-I-Martin X. (1997) I just ran two million regression, *American Economic Association Papers and Proceedings*, 87, 178-183.
- Samuelson P.A. (1954) The transfer problem and transport costs, II : analysis of trade impediments, *Economic Journal*, 53, 202-211.
- Savin N.E. (1984) Multiple hypotheses testing, dans : Griliches Z., Intriligator M.D. (eds.), *Handbook of Econometrics*, volume II, Elsevier Science Publishers.
- Schmutzler A. (1999) The new economic geography, *Journal of Economic Surveys*, 13, 355-379.
- Schroeder L.D., Sjoquist D.L. (1976) Investigation of population density gradients using trend surface analysis, *Land Economics*, 52, 382-392.
- Schwarz G. (1978) Estimating the dimension of a model, *The Annals of Statistics*, 6, 461-464.
- Scitovsky T. (1954) Two concepts of external economies, *Journal of Political Economy*, 62, 143-151.
- Scotchmer S., Thisse J.-F. (1992) Space and competition : a puzzle, *Annals of Regional Science*, 26, 269-286.
- Segerstrom P. (1998) Endogenous growth without scale effects, *American Economic Review*, 88, 1290-1310.
- Sen A. (1976) Large sample-size distribution of statistics used in testing for spatial correlation, *Geographical Analysis*, 8, 175-184.
- Shioji E. (1997a) Convergence in panel data : evidence from the skipping estimation, *mimeo*, Universitat Pompeu Fabra, Barcelona.
- Shioji E. (1997b) It's still 2%. Evidence of convergence from 116 years of the US states panel data, *mimeo*, Universitat Pompeu Fabra, Barcelona.
- Shorrocks A. (1978) The measurement of mobility, *Econometrica*, 46, 1013-1024.
- Silverman B. (1986) *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, Londres.
- Siriopoulos C., Asteriou D. (1998) Testing for convergence across the Greek regions, *Regional Studies*, 537-546.
- Smirnov O., Anselin L. (2001) Fast maximum likelihood estimation of very large spatial autoregressive models : a characteristic polynomial approach, *Computational Statistics and Data Analysis*, 35, 301-319.
- Sokal R.R., Oden N.L. (1978) Spatial autocorrelation in biology. 1. Methodology, *Biological of the Linnean Society*, 10, 229-249.
- Sokal R.R., Oden N.L., Thomson B.A., Kim J. (1993) Testing for regional differences in means : distinguishing inherent from spurious spatial autocorrelation by restricted randomization, *Geographical Analysis*, 25, 199-209.
- Solow R. (1956) A contribution to the theory of economic growth, *Quarterly Journal of Economics*, 70, 65-94.

- Starrett D. (1978) Market allocation of location choice in a model of free mobility, *Journal of Economic Theory*, 17, 21-37.
- Student (1914) The elimination of spurious correlation due to position in time or space, *Biometrika*, 5, 351-360.
- Suits D.B., Mason A., Chan L. (1978) Spline functions fitted by standard regression methods, *Review of Economics and Statistics*, 60, 132-139.
- Summers R., Heston A. (1988) A new set of international comparisons of real product and price levels estimates for 130 countries, 1950-1985, *Review of Income and Wealth*, 34, 1-25.
- Summers R., Heston A. (1991) The penn world table (mark 5) : an expanded set of international comparisons, 1950-1988, *Quarterly Journal of Economics*, 106, 327-368.
- Swan T. (1956) Economic growth and capital accumulation, *The Economic Record*, 32, 334-361.
- Tabuchi T., Thisse J.-F. (2002) Regional specialization and transport costs, *mimeo*, CORE, Université Catholique de Louvain, Louvain-La-Neuve.
- Tamura R. (1991) Income convergence in an endogenous growth model, *Journal of Political Economy*, 99, 522-540.
- Teixeira A.C. (2001) Transport policies in light of new economic geography : the portuguese experience, *mimeo*, CORE, Université Catholique de Louvain, Louvain-La-Neuve.
- Temple J. (1998) Robustness tests of the augmented Solow model, *Journal of Applied Econometrics*, 13, 361-375.
- Temple J. (1999a) The new growth evidence, *Journal of Economic Literature*, 37, 112-156.
- Temple J. (1999b) A positive effect of human capital on growth, *Economics Letters*, 65, 131-134.
- Terui N., Kikuchi M. (1994) The size-adjusted critical region of Moran's *I* test statistics for spatial autocorrelation and its application to geographical areas, *Geographical Analysis*, 26, 213-227.
- Thériault M., Des Rosiers F., Vandersmissen M.-H. (1999) GIS-based simulation of accessibility to enhance hedonic modeling and property value appraisal : an application to the Quebec city metropolitan area, dans : *Integrating GIS & CAMA 1999 Conference Proceedings*, Urban and Regional Information Systems Association (URISA) and International Association of Assessing Officers (IAAO), Chicago.
- Thisse J.-F. (1992) Espace et concurrence, une cohabitation difficile ?, dans : Derycke P.-H. (ed.), *Espace et Dynamiques Territoriales*, Economica, Paris.
- Thisse J.-F. (1997) L'oubli de l'espace dans la pensée économique, *Région et Développement*, 6, 13-39.
- Thisse J.-F. (2000) Agglomeration and regional imbalance : why? And is it bad?, *EIB-Papers*, 5, 47-67.
- Thisse J.-F., Van Ypersele (1999) Métropoles et concurrence territoriale, *Economie et Statistique*, 6/7, 19-30.
- Thisse J.-F., Zénou (1995) Appariement et concurrence spatiale sur la marché du travail, *Revue Economique*, 46, 615-624.

- Tiefelsdorf M. (1998) Some practical approximations of Moran I 's exact distribution, *Papers in Regional Science*, 77, 101-129.
- Tiefelsdorf M. (2000) *Modelling Spatial Processes*, Lecture Notes in Earth Sciences, n°87, Springer, Berlin.
- Tiefelsdorf M., Boots B. (1995) The exact distribution of Moran's I , *Environment and Planning A*, 27, 985-999.
- Tiefelsdorf M., Boots B. (1996) Letter to the editor : the exact distribution of Moran's I , *Environment and Planning A*, 28, 1900.
- Tobler W. (1979) Cellular geography, dans : Gale S., Olsson G. (eds.), *Philosophy in Geography*, Reidel, Dordrecht.
- Upton G.J.G., Fingleton B. (1985) *Spatial Data Analysis by Example*, vol.1, John Wiley, New York.
- Uzawa (1965) Optimal technical change in an aggregative model of economic growth, *International Economic Review*, 6, 18-31.
- Vanhoudt P., Matha T., Smith B. (2000), How productive are capital investments in Europe, *EIB papers*, 5, 81-106.
- Varga A. (1998) *University Research and Regional Innovation : A Spatial Econometric Analysis of Academic Technology Transfers*, Kluwer Academic Publishers, Boston.
- Varga A. (2000) Local academic knowledge transfers and the concentration of economic activity, *Journal of Regional Science*, 40, 289-309.
- Venables D. (1996) Equilibrium location of vertically linked industries, *International Economic Review*, 37, 341-359.
- Walz U. (1996) Transport costs, intermediate goods, and localized growth, *Regional Science and Urban Economics*, 26, 671-695.
- Walz U. (1998) Does an enlargement of a common market stimulate growth and convergence, *Journal of International Economics*, 45, 297-321.
- Weisberg S. (1985) *Applied Linear Regression*, New York, Wiley.
- White H. (1980) A heteroskedastic-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, 48, 817-838.
- Wolfson M.C. (1994) When inequalities diverge, *American Economic Review*, 84, 353-358.
- Young A. (1998) Growth without scale effects, *Journal of Political Economy*, 106, 41-63.
- Zellner A. (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, *Journal of the American Statistical Association*, 57, 348-368.

Liste des tableaux et des figures

Liste des tableaux

Tableau 2.1 : Espérances et variances de <i>NN</i> et <i>NB</i>	107
Tableau 2.2 : Espérances et variances des statistiques de Moran et de Geary.....	110
Tableau 2.3 : Résumé des différents tests.....	147
Tableau 3.1 : Correspondance entre les niveaux NUTS et les divisions administratives nationales pour les 15 pays de l'Union Européenne.....	206
Tableau 3.2 : Auteurs ayant utilisé Eurostat-Regio pour étudier les disparités de PIB par tête entre les régions européennes.....	212
Tableau 3.3 : Disparités de PIB par tête entre 138 régions européennes.....	213
Tableau 3.4a : Evolution de la statistique de Moran pour les PIB par tête en logarithmes et en Ecus sur la période 1980-1995.....	220
Tableau 3.4b : Evolution de la statistique de Moran pour les PIB par tête en logarithmes et en SPA sur la période 1980-1995.....	220
Tableau 3.5a : Valeurs extrêmes de la statistique Getis-Ord ; PIB par tête en logarithmes et en Ecus pour 1980 et 1995.....	225
Tableau 3.5b : Valeurs extrêmes de la statistique Getis-Ord ; PIB par tête en logarithmes et en SPA pour 1980 et 1995.....	226
Tableau 3.6a : Cohésion et flux dans les diagrammes de Moran ; PIB par tête en logarithmes et en Ecus sur la période 1980-1995.....	232
Tableau 3.6b : Cohésion et flux dans les diagrammes de Moran ; PIB par tête en logarithmes et en SPA sur la période 1980-1995.....	232
Tableau 3.7a : Observations extrêmes du diagramme de Moran ; PIB par tête en logarithmes et en Ecus pour 1980 et 1995.....	235
Tableau 3.7b : Observations extrêmes du diagramme de Moran ; PIB par tête en logarithmes et en SPA pour 1980 et 1995.....	235
Tableau 3.8 : Valeurs extrêmes des statistiques LISA ; PIB par tête en logarithmes, en Ecus et SPA pour 1980 et 1995.....	241
Tableau 3.9a : Analyse de robustesse pour les statistiques Getis-Ord ; PIB par tête en logarithmes et en Ecus sur la période 1980-1995.....	242
Tableau 3.9b : Analyse de robustesse pour les statistiques Getis-Ord ; PIB par tête en logarithmes et en SPA sur la période 1980-1995.....	242
Tableau 3.10a : Analyse de robustesse pour les statistiques LISA ; PIB par tête en logarithmes et en Ecus sur la période 1980-1995.....	243
Tableau 3.10b : Analyse de robustesse pour les statistiques LISA ; PIB par tête en logarithmes et en SPA sur la période 1980-1995.....	244
Tableau 3.11 : Valeurs extrêmes de la statistique Getis-Ord ; taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus et en SPA.....	247
Tableau 3.12 : Valeurs extrêmes du diagramme de Moran ; taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus et en SPA.....	250

Tableau 3.13 : Valeurs extrêmes des statistiques LISA ; taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus et en SPA.....	253
Tableau 3.14a : Statistiques Getis-Ord et LISA pour la matrice $W(10)$; PIB par tête en logarithmes et en Ecus pour 1980, 1995 et pour le taux de croissance annuel moyen	255
Tableau 3.14b : Statistiques Getis-Ord et LISA pour la matrice $W(10)$; PIB par tête en logarithmes et en SPA pour 1980, 1995 et pour le taux de croissance annuel moyen	256
Tableau 4.1 : Convergence dans les régions européennes	304
Tableau 4.2a : Résultats d'estimation avec $W(1)$ et PIB par tête en Ecus	307
Tableau 4.2b : Résultats d'estimation avec $W(1)$ et PIB par tête en SPA.....	308
Tableau 4.3a : Résultats d'estimation avec $W(10)$ et PIB par tête en Ecus	309
Tableau 4.3b : Résultats d'estimation avec $W(10)$ et PIB par tête en SPA	310
Tableau 4.4 : Détection des clubs spatiaux basée sur le diagramme de Moran	325
Tableau 4.5 : Détection des clubs spatiaux basée sur les statistiques Getis-Ord	327
Tableau 4.6 : Résultats d'estimation pour le modèle de β -convergence absolue.....	330
Tableau 4.7 : Résultats d'estimation pour le modèle de β -convergence avec autocorrélation des erreurs	331
Tableau 4.8 : Résultats d'estimation pour le modèle de β -convergence avec hétéroscédasticité en groupes.....	333
Tableau 4.9 : Résultats d'estimation pour le modèle à deux régimes	335
Tableau 4.10 : Résultats d'estimation pour le modèle de β -convergence avec autocorrélation des erreurs et 2 régimes.....	338
Tableau 4.11 : Résultats d'estimation avec régimes et effets spatiaux différenciés	340
Tableau 4.12 : Résultats d'estimation avec effets spatiaux différenciés.....	341
Tableau 5.1a : Matrice de transition ; PIB par tête en Ecus ; conditionnement européen	378
Tableau 5.1b : Matrice de transition ; PIB par tête en SPA ; conditionnement européen	378
Tableau 5.2a : Matrices de transition par sous-période ; PIB par tête en Ecus ; conditionnement européen.....	380
Tableau 5.2b : Matrices de transition par sous-période ; PIB par tête en SPA ; conditionnement européen.....	380
Tableau 5.3a : Matrice des temps moyens de premier passage ; PIB par tête en Ecus.....	383
Tableau 5.3b : Matrice des temps moyens de premier passage ; PIB par tête en SPA.....	383
Tableau 5.4 : Distribution ergodique et distribution initiale ; PIB par tête en Ecus et en SPA ; conditionnement européen.....	384
Tableau 5.5 : Indices de mobilité ; PIB par tête en Ecus et en SPA	386
Tableau 5.6 : Valeurs critiques pour les tests d'autocorrélation spatiale ; transitions entre classes de PIB ; PIB par tête en Ecus et en SPA.....	392
Tableau 5.7 : Valeurs critiques pour les tests d'autocorrélation spatiale ; classement des régions ; PIB par tête en Ecus et en SPA.....	394
Tableau 5.8a : Matrice de transition conditionnelle ; PIB par tête en Ecus.....	398
Tableau 5.8b : Matrice de transition conditionnelle ; PIB par tête en SPA	398
Tableau 5.9a : Matrice de transition ; conditionnement régional ; PIB par tête en Ecus.....	400

Tableau 5.9b : Matrice de transition ; conditionnement régional ; PIB par tête en SPA.....	400
Tableau 5.10a : Chaîne de Markov spatiale ; conditionnement sur le décalage spatial au début de chaque année ; PIB par tête en Ecus	404
Tableau 5.10b : Chaîne de Markov spatiale ; conditionnement sur le décalage spatial au début de chaque année ; PIB par tête en SPA	405
Tableau 5.11a : Probabilités de transition conditionnées sur le décalage spatial	407
au début de chaque année ; PIB par tête en Ecus	407
Tableau 5.11b : Probabilités de transition conditionnées sur le décalage spatial	408
au début de chaque année ; PIB par tête en SPA.....	408

Liste des figures

Figure 1.1 : Dynamique dans le modèle de Solow-Swan – économies identiques	21
Figure 1.2 : Dynamique dans le modèle de Solow-Swan – économies différentes.....	22
Figure 1.3 : Le modèle AK.....	30
Figure 1.4 : Relation entre écart initial et écart final	53
Figure 1.5 : β - et σ -convergence	57
Figure 1.6 : Représentation des conditions de β - et de σ -convergence.....	59
Figure 1.7a : σ -convergence, limites (1).....	60
Figure 1.7b : σ -convergence, limites (2).....	60
Figure 1.8 : Convergence et évolution de la distribution	62
Figure 1.9 : Causalité circulaire dans la formation d'une agglomération d'entreprises et de travailleurs.....	78
Figure 2.1 : Les définitions de la contiguïté pour une configuration spatiale régulière	102
Figure 2.2 : Diagramme de Moran	114
Figure 2.3 : Echiquier	117
Figure 2.4 : Autocorrélation spatiale et effet d'échelle	118
Figure 2.5 : Autocorrélation spatiale et effet de forme.....	119
Figure 2.6 : Un exemple d'utilisation de la méthode TSA	163
Figure 3.1a : PIB par tête en Ecus relatifs à la moyenne européenne en 1980	215
Figure 3.1b : PIB par tête en SPA relatifs à la moyenne européenne en 1980.....	215
Figure 3.2a : PIB par tête en Ecus relatifs à la moyenne européenne en 1995	216
Figure 3.2b : PIB par tête en SPA relatifs à la moyenne européenne en 1995.....	216
Figure 3.3a : Statistiques de Getis-Ord significatives ; PIB par tête en logarithmes et en Ecus pour 1980	223
Figure 3.3b : Statistiques de Getis-Ord significatives ; PIB par tête en logarithmes et en SPA pour 1980	223
Figure 3.4a : Statistiques de Getis-Ord significatives ; PIB par tête en logarithmes et en Ecus pour 1995	224
Figure 3.4b : Statistiques de Getis-Ord significatives ; PIB par tête en logarithmes et en SPA pour 1995	224
Figure 3.5a : Diagramme de Moran ; PIB par tête en logarithmes et en Ecus pour 1980.....	229
Figure 3.5b : Diagramme de Moran ; PIB par tête en logarithmes et en SPA pour 1980.....	229
Figure 3.6a : Diagramme de Moran ; PIB par tête en logarithmes et en Ecus pour 1995.....	230
Figure 3.6b : Diagramme de Moran ; PIB par tête en logarithmes et en SPA pour 1995.....	230
Figure 3.7a : Statistiques LISA significatives ; PIB par tête en logarithmes et en Ecus pour 1980.....	237
Figure 3.7b : Statistiques LISA significatives ; PIB par tête en logarithmes et en SPA pour 1980.....	237

Figure 3.8a : Statistiques LISA significatives ; PIB par tête en logarithmes et en Ecus pour 1995.....	238
Figure 3.8b : Statistiques LISA significatives ; PIB par tête en logarithmes et en SPA pour 1995.....	238
Figure 3.9a : Statistiques de Getis-Ord significatives ; taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus.....	248
Figure 3.9b : Statistiques de Getis-Ord significatives ; taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en SPA.....	248
Figure 3.10a : Diagramme de Moran pour le taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus.....	251
Figure 3.10b : Diagramme de Moran pour le taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en SPA.....	251
Figure 3.11a : Statistiques LISA significatives ; taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en Ecus.....	254
Figure 3.11b : Statistiques LISA significatives ; taux de croissance annuel moyen entre 1980 et 1995 ; PIB par tête en SPA.....	254
Figure 4.1a : Diffusion dans le modèle avec autocorrélation des erreurs avec $W(1)$ et PIB par tête en Ecus.....	314
Figure 4.1b : Diffusion dans le modèle avec autocorrélation des erreurs avec $W(1)$ et PIB par tête en SPA.....	314
Figure 4.2a : Diffusion dans le modèle avec autocorrélation des erreurs avec $W(10)$ et PIB par tête en Ecus.....	315
Figure 4.2b : Diffusion dans le modèle avec autocorrélation des erreurs avec $W(10)$ et PIB par tête en SPA.....	315
Figure 4.3a : Quartiles de la distribution des impacts médians d'émission pour $W(1)$ et PIB par tête en Ecus.....	319
Figure 4.3b : Quartiles de la distribution des impacts médians d'émission pour $W(1)$ et PIB par tête en SPA.....	319
Figure 4.4a : Quartiles de la distribution des impacts médians de réception pour $W(1)$ et PIB par tête en Ecus.....	320
Figure 4.4b : Quartiles de la distribution des impacts médians de réception pour $W(1)$ et PIB par tête en SPA.....	320
Figure 5.1a : σ -convergence pour la période 1980 - 1995 ; PIB par tête en Ecus.....	360
Figure 5.1b : σ -convergence pour la période 1980 - 1995 ; PIB par tête en SPA.....	360
Figure 5.2a : Densités des distributions de PIB par tête en Ecus ; conditionnement européen.....	364
Figure 5.2b : Densités des distributions de PIB par tête en SPA ; conditionnement européen.....	364
Figure 5.3 : Indice de polarisation ($\alpha = 1.5$) ; 1980-1995.....	365
Figure 5.4 : Stratification, polarisation et clubs de convergence.....	366

Figure 5.5a : Transitions de classe des régions entre 1980-1995 ; PIB par tête en Ecus ; conditionnement européen	391
Figure 5.5b : Transitions de classe des régions entre 1980-1995 ; PIB par tête en SPA ; conditionnement européen	391
Figure 5.6a : Classement des régions en fonction du PIB par tête en 1980 et 1995 ; PIB par tête en Ecus	393
Figure 5.6b : Classement des régions en fonction du PIB par tête en 1980 et 1995 ; PIB par tête en SPA.....	394
Figure 5.7a : Densités des distributions de PIB par tête mesuré en Ecus ; 1980 - 1995 ; conditionnement régional.....	396
Figure 5.7b : Densités des distributions de PIB par tête mesuré en SPA ; 1980 - 1995 ; conditionnement régional.....	396

Table des matières

INTRODUCTION	1
CHAPITRE 1 : DISPARITÉS GÉOGRAPHIQUES ET CONVERGENCE - UN NOUVEAU DÉBAT	12
SECTION 11 : CONVERGENCE ET MODELES DE CROISSANCE	16
111 La convergence dans le modèle néoclassique.....	16
111.1 Les mécanismes de la croissance dans le modèle de Solow-Swan	17
111.2 Le processus de convergence vers l'état régulier : présentation et tests.....	20
111.21 Convergence pour des économies identiques.....	20
111.22 Convergence pour des économies différentes	22
111.23 Le test de l'hypothèse de convergence	23
111.3 Le processus de convergence en présence de capital humain	25
112 La convergence dans les théories de la croissance endogène	27
112.1 Les origines de la croissance endogène.....	27
112.2 Le modèle AK.....	29
112.3 Les premiers modèles de croissance endogène	31
113 Convergence et discrimination entre les modèles de croissance.....	32
113.1 Convergence absolue <i>versus</i> convergence conditionnelle	32
113.2 Croissance endogène et convergence.....	33
Conclusion de section.....	35
SECTION 12 : ANALYSE EMPIRIQUE DE LA CONVERGENCE	36
121 La convergence-rattrapage ou β -convergence	36
121.1 La β -convergence absolue et le « rattrapage » des économies riches par les économies pauvres.....	37
121.2 La β -convergence conditionnelle ou l'individualisation de la notion de rattrapage.....	39
121.21 Le test de l'hypothèse de β -convergence conditionnelle	39
121.22 Critiques de l'approche en termes de β -convergence conditionnelle.....	42
121.23 Estimation par les données de panel.....	47
121.3 Les clubs de convergence.....	50
121.31 Equilibres multiples dans les modèles de croissance	51
121.32 Détection empirique des clubs de convergence.....	51
122 Les autres concepts de convergence	55
122.1 La σ -convergence	55
122.11 La relation entre σ - et β -convergence	56
122.12 Les limites du concept de σ -convergence	58
122.2 L'étude de l'évolution des distributions en coupe transversale.....	61
122.3 La convergence comme persistance des écarts de PIB par tête.....	62
Conclusion de section.....	63
SECTION 13 : ECONOMIE GEOGRAPHIQUE DES DISPARITES REGIONALES ..	64
131 Concentrations spatiales et tests de convergence.....	64
132 La formation des agglomérations : fondements et principes.....	68
132.1 L'agglomération en présence d'hétérogénéités	68

132.2 Concurrence monopolistique et externalités	70
132.3 L'agglomération : un équilibre des forces centripètes et centrifuges	71
132.31 Les forces d'agglomération	72
132.32 Les forces de dispersion	73
132.33 Les équilibres géographiques	74
133 Schéma centre-périphérie et intégration économique	75
133.1 L'équilibre centre-périphérie	76
133.2 D'un modèle d'agglomération à l'autre	79
134 La synthèse géographie-croissance	84
134.1 Les fondements de la synthèse	85
134.2 Effets de débordement et équilibres spatiaux	86
134.3 Croissance et agglomération	89
Conclusion de section	91
Conclusion du chapitre 1	93
CHAPITRE 2 : ECONOMÉTRIE SPATIALE	95
SECTION 21 : LES EFFETS SPATIAUX - DEFINITION ET DETECTION	98
211 Autocorrélation et hétérogénéité spatiales	99
211.1 L'autocorrélation spatiale	99
211.2 L'hétérogénéité spatiale et ses deux facettes	100
212 Les matrices de poids et les variables spatiales décalées	101
212.1 Les matrices de contiguïté	102
212.2 Autres matrices de poids	103
212.3 Standardisation des matrices de poids	104
212.4 Variable spatiale décalée	105
213 Analyse exploratoire des données spatiales	105
213.1 Autocorrélation spatiale globale	105
213.11 Variables qualitatives dichotomiques	106
213.12 La statistique de Moran	108
213.13 La statistique de Geary	109
213.14 La statistique générale Γ	111
213.2 Autocorrélation spatiale locale	112
213.21 Les statistiques de Getis-Ord	112
213.22 Le diagramme de Moran	114
213.23 Les indicateurs locaux d'association spatiale (LISA)	115
213.3 Quelques précautions méthodologiques	117
213.31 La dépendance à la définition de la matrice de poids	117
213.32 Le problème « MAUP »	118
Conclusion de section	119
SECTION 22 : AUTOCORRELATION SPATIALE	120
221 Autocorrélation spatiale et modèles économétriques	120
221.1 Le modèle autorégressif spatial : variable endogène décalée	121
221.2 Modèle régressif croisé : variable exogène décalée	124
221.3 Modèle avec autocorrélation spatiale des erreurs	124
221.31 Les erreurs suivent un processus autorégressif	124
221.32 Deux alternatives	126
221.4 Généralisation	127

Table des matières

222 Estimation des modèles spatiaux	129
222.1 La non-convergence des MCO et ses conséquences	129
222.2 Estimation par le maximum de vraisemblance	130
222.21 La fonction de vraisemblance dans le cas général	131
222.22 Le jacobien	132
222.3 Autres méthodes d'estimation	134
222.31 La méthode des variables instrumentales	134
222.32 La méthode des moments généralisés	136
223 Les tests en économétrie spatiale	136
223.1 Principes de tests en économétrie	137
223.2 Le test de l'autocorrélation spatiale de Moran	139
223.3 Les tests d'une hypothèse simple	140
223.31 Modèle avec autocorrélation spatiale des erreurs	140
223.32 Spécification de Kelejian-Robinson	141
223.33 Test d'une variable endogène décalée	141
223.4 Les tests en présence d'une autocorrélation des erreurs et d'une variable décalée	141
223.41 Test joint	142
223.42 Tests conditionnels	142
223.43 Tests robustes	143
223.5 Puissance et robustesse des tests	144
223.6 Autres tests d'autocorrélation spatiale	145
224 A la recherche de la spécification du modèle	146
224.1 Le test du facteur commun	146
224.2 Les règles de décision	147
225 Quelques extensions récentes	149
225.1 Données de panel et autocorrélation spatiale	149
225.11 Modèles dynamiques dans le temps et dans l'espace	149
225.12 Modèle SUR spatial	151
225.13 Modèles à erreurs composées	152
225.2 Les modèles à variables qualitatives	153
225.3 Les modèles à équations simultanées	155
Conclusion de section	155
SECTION 23 : HETEROGENEITE SPATIALE	157
231 Modélisation et estimation de l'instabilité spatiale des paramètres	157
231.1 Variation spatiale discrète	158
231.11 ANOVA spatiale	158
231.12 Régimes spatiaux	159
231.13 Fonctions <i>spline</i>	160
231.2 Variation spatiale continue	161
231.21 La méthode <i>Trend Surface Analysis</i> (TSA)	162
231.22 La méthode <i>Variable Expansion</i> (VE)	163
231.23 La méthode non-paramétrique <i>Geographically Weighted Regression</i> (GWR)	165
232 Modélisation et estimation de l'hétéroscédasticité	167
232.1 La méthode de White	168
232.2 Spécification de l'hétéroscédasticité	169
232.21 La variance est fonction d'un ensemble de variables	169
232.22 Hétéroscédasticité en groupes	170
232.3 Hétéroscédasticité et instabilité des paramètres	172
232.31 Régimes spatiaux et hétéroscédasticité	172

232.32	Variation spatiale continue et hétéroscédasticité	172
233	L'interaction entre autocorrélation et hétérogénéité spatiales	174
233.1	Les liens entre hétérogénéité et autocorrélation spatiales	175
233.2	Le test de la dépendance spatiale en présence d'hétéroscédasticité	176
233.21	Le test de l'autocorrélation spatiale en présence d'hétéroscédasticité de forme connue	177
233.22	Le test de l'autocorrélation spatiale en présence d'hétéroscédasticité de forme inconnue	177
233.3	Le test de l'hétéroscédasticité en présence de dépendance spatiale	180
233.31	Test joint d'hétéroscédasticité et d'autocorrélation spatiale des erreurs	180
233.32	Le test de l'hétéroscédasticité en présence d'autocorrélation spatiale des erreurs	181
233.4	Le test de l'instabilité des paramètres en présence de dépendance spatiale	182
233.41	Les différents cas possibles	182
233.42	Le test	185
	Conclusion de section	186
	Conclusion du chapitre 2	187
	Annexe 2.1 : Algorithmes de construction d'une matrice générale des distances à partir d'une matrice de contiguïté d'ordre 1	189
	Annexe 2.2 : Estimation par le maximum de vraisemblance : étude de deux cas particuliers	197
	CHAPITRE 3 : ANALYSE EXPLORATOIRE DES DISPARITÉS RÉGIONALES DE PIB PAR TÊTE EN EUROPE	201
	SECTION 31 : LES DONNEES ET LA MATRICE DE POIDS	204
311	La base de données Eurostat-Regio	204
311.1	La Nomenclature des Unités Territoriales Statistiques	204
311.2	Ecus et parités de pouvoir d'achat	207
312	L'échantillon	209
312.1	Le choix du découpage spatial	209
312.2	Disparités économiques et mesure du PIB par tête	211
313	La matrice de poids	214
313.1	Matrices basées sur la distance	214
313.2	Distance ou contiguïté entre les régions européennes : quelques éléments de comparaison	217
	Conclusion de section	218
	SECTION 32 : ANALYSE EXPLORATOIRE SPATIALE DES PIB PAR TETE REGIONAUX	219
321	Une forte autocorrélation spatiale globale	219
322	Concentrations locales de régions pauvres et de régions riches : statistiques Getis-Ord $G_i(k)$	221
323	Concentrations de régions similaires et détection de régions atypiques : le diagramme de Moran	226
323.1	Les diagrammes de Moran en 1980 et 1995	226
323.2	Evolution temporelle des diagrammes de Moran	228
323.3	Analyse des observations extrêmes	232
324	La significativité des concentrations locales : les statistiques LISA	236
325	Des résultats robustes au choix de la matrice de poids	241
	Conclusion de section	244
	SECTION 33 : ANALYSE EXPLORATOIRE SPATIALE DES TAUX DE CROISSANCE ANNUELS MOYENS	245
331	Les concentrations spatiales de taux de croissance	245

331.1 Une forte autocorrélation spatiale globale positive	246
331.2 Concentrations locales et statistiques Getis-Ord	246
331.3 Concentrations locales, régions atypiques et diagramme de Moran	249
331.4 Concentrations locales, régions atypiques et statistiques LISA	252
332 Analyse exploratoire de la relation entre taux de croissance et PIB par tête initial	257
Conclusion de section	258
Conclusion du chapitre 3	259
Annexe 3.1 : Les régions de l'échantillon	261
Annexe 3.2 : Résultats pour les sections 32 et 33 avec les matrices des distances inverses ; PIB par tête en Ecus	267
Annexe 3.3 : Résultats pour les sections 32 et 33 avec les matrices des distances inverses ; PIB par tête en SPA	274
Annexe 3.4 : Statistiques $G_i(k)$ et LISA pour $W(10)$; PIB par tête en logarithmes et en Ecus ; 1980-1995 ...	281
Annexe 3.5 : Statistiques $G_i(k)$ et LISA pour $W(10)$; PIB par tête en logarithmes et en SPA ; 1980-1995 ...	285
CHAPITRE 4 : ANALYSE ÉCONOMÉTRIQUE SPATIALE DE LA β-CONVERGENCE ENTRE LES RÉGIONS EUROPÉENNES	289
SECTION 41 : β-CONVERGENCE ET AUTOCORRELATION SPATIALE	292
411 Raisons et implications du choix du modèle de β -convergence absolue	292
412 L'autocorrélation spatiale dans le modèle de β -convergence absolue	295
412.1 Le modèle avec variable endogène décalée	295
412.2 Le modèle régressif croisé	296
412.3 Le modèle avec autocorrélation des erreurs	297
413 Convergence et effets de débordement géographiques	298
413.1 Une forte autocorrélation spatiale dans le modèle de β -convergence absolue	299
413.2 Le modèle avec autocorrélation spatiale des erreurs	301
414 Analyse du processus de diffusion spatiale d'un choc aléatoire	311
414.1 Analyse du choc affectant une seule région	312
414.2 Emission et réception au sein du système régional européen	313
414.21 Présentation analytique de la diffusion globale des chocs aléatoires	313
414.22 La géographie des chocs aléatoires dans le système régional européen	317
Conclusion de section	318
SECTION 42 : β-CONVERGENCE ET HETEROGENEITE SPATIALE	321
421 Les clubs de convergence en Europe	321
421.1 Les clubs de convergence en Europe, un tour d'horizon	322
421.2 Détection des clubs spatiaux	323
421.21 Instabilité spatiale des PIB par tête en Ecu en 1980	323
421.22 Instabilité spatiale des PIB par tête en SPA en 1980	324
422 Résultats des estimations	328
422.1 Estimation par les MCO du modèle de β -convergence absolue et tests	328
422.2 Traitement de l'autocorrélation spatiale	329
422.3 Hétérogénéité spatiale : hétéroscédasticité en groupes et/ou instabilité structurelle	332
422.4 Autocorrélation et hétérogénéité spatiales	336
422.5 Effets spatiaux différenciés	339
Conclusion de section	341
Conclusion du chapitre 4	342

Annexe 4.1 : Estimation du modèle avec autocorrélation spatiale des erreurs avec $W(2)$, $W(3)$, $W(15)$ et $W(20)$; PIB par tête en Ecus	343
Annexe 4.2 : Estimation du modèle avec autocorrélation spatiale des erreurs avec $W(2)$, $W(3)$, $W(15)$ et $W(20)$; PIB par tête en SPA.....	344
Annexe 4.3 : Boîtes à moustaches pour les impacts d'émission et de réception.....	345
Annexe 4.4 : Estimation du modèle avec autocorrélation spatiale des erreurs et régimes spatiaux avec $W(2)$, $W(3)$, $W(15)$ et $W(20)$; PIB par tête en Ecus.....	349
Annexe 4.5 : Estimation du modèle avec autocorrélation spatiale des erreurs et régimes spatiaux avec $W(2)$, $W(3)$, $W(15)$ et $W(20)$; PIB par tête en SPA.....	350
Annexe 4.6 : Estimation du modèle avec effets spatiaux différenciés par le maximum de vraisemblance	351

CHAPITRE 5 : CONVERGENCE DES RÉGIONS EUROPÉENNES ET CHAÎNES DE MARKOV

355

SECTION 51 : LA DYNAMIQUE DE LA DISTRIBUTION DES PIB REGIONAUX EN EUROPE

358

511 Evolution de la forme de la distribution des PIB par tête	358
511.1 Le test de la σ -convergence	359
511.2 Estimation non paramétrique des fonctions de densité : principes et résultats.....	361
511.3 Etude de la polarisation.....	363
511.4 Limites de ces approches	366
512 Chaînes de Markov : principes	367
512.1 Définitions.....	367
512.2 Typologie des chaînes de Markov.....	370
512.21 Classification des états	370
512.22 Classification des chaînes de Markov	372
512.3 Comportement limite des chaînes de Markov	373
513 Convergence et processus de Markov : résultats empiriques.....	374
513.1 Principes de modélisation	375
513.2 Estimation des matrices de transition.....	376
513.3 Test de stabilité	379
513.4 Caractéristiques de la dynamique de transition.....	381
513.41 Une forte persistance des situations régionales	382
513.42 Une fluidité des transitions assez faible	382
513.43 L'équilibre régional de long terme : un désavantage pour les régions pauvres.....	384
513.44 Indice de mobilité en convergence et indice de mobilité à l'équilibre	385
Conclusion de section.....	386

SECTION 52 : L'INTEGRATION DE LA DIMENSION SPATIALE DANS LES PROCESSUS DE MARKOV.....

388

521 Les schémas géographiques des transitions régionales.....	389
521.1 Autocorrélation spatiale et transition entre classes de PIB	389
521.2 Autocorrélation spatiale et classement des régions	392
522 Conditionnement spatial	394
522.1 Conditionnement géographique et richesse des économies régionales	395
522.2 Conditionnement géographique et dynamique de transition	397
523 Dynamiques spatiales et temporelles des disparités régionales	399

Table des matières

523.1 Analyse des transitions temporelles avec conditionnement régional	399
523.2 La chaîne de Markov spatiale (Rey, 2001)	401
523.21 Description de la matrice de Markov spatiale	401
523.22 Le test de stationnarité spatiale.....	402
523.23 Environnement géographique et convergence des régions européennes.....	403
Conclusion de section.....	408
Conclusion du chapitre 5	409
Annexe 5.1 : Tests d'autocorrélation spatiale avec les matrices de poids basées sur la distance.....	410
Annexe 5.2 : Conditionnement régional avec la matrice des distances $W(1)$; PIB par tête en Ecus.....	411
Annexe 5.3 : Conditionnement régional avec la matrice des distances $W(1)$; PIB par tête en SPA	414
Annexe 5.4 : Probabilités de transitions conditionnées sur le décalage spatial pour les matrices de distances ..	417
CONCLUSION	418
BIBLIOGRAPHIE	426
LISTE DES TABLEAUX ET DES FIGURES	457
TABLE DES MATIÈRES	464