

PROXIMITÉ TECHNOLOGIQUE, INFRASTRUCTURES DE COMMUNICATION ET ACTIVITÉS INNOVANTES EN EUROPE

Olivier Parent *

University of Cincinnati, Department of Economics

D'après les modèles de la nouvelle économie géographique, le caractère limité de la diffusion des connaissances représente un des principaux déterminants de la localisation des activités innovantes. En s'appuyant sur la problématique des externalités dynamiques présentée par Glaeser et al. (1992), de nombreux travaux économétriques récents tentent d'évaluer les effets de la diffusion des connaissances sur la localisation des activités innovantes. Les externalités de connaissance étant favorisées par la proximité à la fois géographique et technologique, nous étudions différentes structures de dépendance spatiale intégrant l'indice technologique de Jaffe (1986) et le temps de transport entre régions européennes. Les résultats mettent en avant l'importance de l'information concernant la qualité des infrastructures de communication ainsi que de la proximité technologique dans l'explication de la polarisation des activités innovantes en Europe.

Mots clés : Économétrie spatiale. Innovation en Europe.

* Je remercie Jean-Luc Gaffard ainsi que les rapporteurs pour leurs commentaires.
olivier.parent@uc.edu

Les études empiriques sur l'innovation en Europe soulignent l'importance des phénomènes d'agglomération et de spécialisation (Brülhart et Torstensson, 1996). Paci et Usai (2000) relèvent une forte concentration des activités scientifiques et technologiques dans quelques régions européennes. Engelman et Walz (1995) justifient théoriquement ces phénomènes de concentration d'activités innovantes en évoquant le caractère local de la diffusion des externalités de connaissance. Ainsi peut-on expliquer l'apparition de régions ou pôles technologiques et industriels tels que la Silicon Valley aux États-Unis.

La diffusion de connaissances entre les régions conditionne le processus de convergence de l'activité économique. Une forte concentration des activités innovantes associée à des externalités très localisées conduit à un renforcement des inégalités régionales (Martin et Ottaviano, 1999). Ainsi, promouvoir l'équité régionale nécessite une remise en cause de la croissance globale. Il est donc intéressant de s'interroger sur les déterminants de la diffusion de connaissances pouvant stimuler la convergence entre régions. En introduisant les mécanismes de croissance endogène dans les modèles d'économie géographique, Baldwin et Forslid (2000) démontrent que ces externalités interrégionales représentent une force centrifuge encourageant la dispersion de l'activité économique et une croissance moins polarisée. Ces résultats théoriques ouvrent le champ à de nombreuses études économétriques sur l'estimation des processus de convergence entre régions et la modélisation des phénomènes de diffusion de connaissances (El Ouardighi, 2002; Baumont *et al.* 2004).

En s'interrogeant sur les formes de spécialisation des activités d'innovation, Audretsch et Feldman (1996) ont souligné l'importance de la dimension locale des spillovers technologiques. Jaffe *et al.* (1993), et Verspagen et Schoenmakers (2000) ont remarqué, en étudiant les citations de brevets, que la probabilité de citer un brevet est influencée positivement par la proximité technologique entre le nouveau brevet et le brevet cité. En s'appuyant sur ces considérations empiriques, l'enjeu de cette étude sera de tenter d'apporter des explications aux affinités repérées en évaluant la place de la proximité géographique par rapport à la proximité technologique. Cette analyse se concentre sur les interactions spatiales entre régions innovantes, à partir de différentes mesures basées sur la proximité géographique ainsi que sur la qualité des infrastructures de communication.

L'étude que nous allons proposer repose sur une approche économétrique définissant les structures de dépendance spatiale à partir de mesures tenant compte à la fois de la proximité géographique et de la

proximité technologique. Les externalités de connaissance se diffusent en effet plus facilement entre régions dont les activités de recherche et de production sont développées dans les mêmes champs scientifiques. Les matrices de pondération seront notamment construites en utilisant l'indice de proximité technologique introduit par Jaffe (1986). Nous allons utiliser un modèle bayésien hiérarchique robuste aux observations extrêmes et aux effets spatiaux. Ces effets sont de nouveau spécifiés à partir d'un modèle spatial autorégressif (SAR). L'hétéroscédasticité des dépôts de brevets est prise en compte à partir d'une méthode de mélange de lois introduite par Geweke (1993). Les résultats montrent que ces interactions entre régions sont sensibles à la prise en compte du temps de transport. Mais la proximité technologique reste le phénomène le plus important dans la formation de pôles européens d'innovation.

La partie 1 présente les fondements théoriques motivant l'utilisation des indices de proximité technologique et ceux de qualité des infrastructures de communication. Dans la partie 2, à partir d'une fonction de production de connaissance proposée par Griliches (1979), nous introduisons les phénomènes de dépendance spatiale. Dans la partie 3 nous discutons la méthode d'estimation et nous proposons différents modèles à comparer. Dans la partie 4, les différentes spécifications des interactions spatiales entre régions mettent en avant l'importance des informations relatives aux temps de transport ainsi qu'à la proximité technologique dans les processus d'innovation. Ces spécifications nous permettent de mesurer et d'analyser les effets des externalités de connaissance sur la localisation des activités d'innovation en Europe.

I. Mesures de proximité technologique et des coûts de transport

Les relations entre géographie et activité d'innovation représentent l'un des principaux centres d'intérêt dans le domaine de la géographie de l'innovation. De nombreuses études empiriques ont confirmé l'importance de la proximité géographique dans la transmission de connaissances (Jaffe *et al.*, 1993), et ont notamment souligné les atouts que présentent certaines régions dans la production d'innovation. Cette dimension spatiale est étudiée pour la France par Autant-Bernard (2001) qui met également en avant le rôle de la proximité technologique dans les processus d'innovation.

I.1. Modélisation des externalités

Jaffe (1989) et Feldman (1994) ont mis en évidence la corrélation entre les dépôts de brevets et les investissements en recherche et

développement en introduisant une fonction de production de l'innovation. Cette simple corrélation a été interprétée comme une mesure de l'externalité locale de connaissance. En effet, à cause de la nature « tacite » de la connaissance, les externalités sont perçues comme un phénomène local (Saxenian, 1994; Marshall, 1890). L'échange d'idées entre agents économiques est ainsi stimulé par la proximité spatiale et le contact face-à-face. Maurseth et Verspagen (2002) suggèrent l'existence de barrières empêchant les flux de connaissance de se diffuser à travers l'Europe. Ils montrent que les brevets sont plus souvent cités entre régions qui sont géographiquement proches et appartenant à un même pays. Leur étude reste essentiellement descriptive et ne tient pas compte de l'effet des externalités sur la production d'innovation.

De nombreux travaux ont mis en avant les liens existant entre localisations géographiques et disparités économiques au sein de l'espace européen (Le Gallo et Ertur, 2003; Le Gallo, 2004; Ertur et Koch, 2006). Outre l'intérêt empirique lié à l'analyse de l'espace européen et de ses disparités, ces différents travaux trouvent également une justification à travers des théories de l'économie géographique (Puga, 2002). Plus spécifiquement, ces théories aident à la compréhension de l'émergence de schémas polarisés de type centre-périphérie. Elles démontrent également que la dimension géographique n'est pas neutre sur les processus de croissance économique, plaidant ainsi pour une prise en compte plus importante de la dimension géographique dans l'analyse économique.

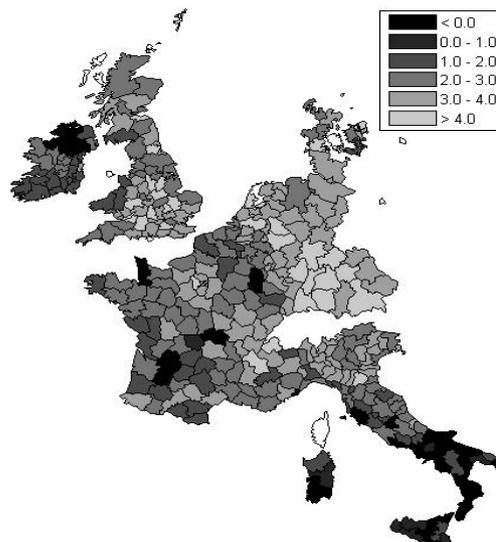
L'objectif est d'estimer à quel point l'activité innovante est sensible aux externalités interrégionales, mais aussi aux caractéristiques propres à chaque région indépendamment des localités voisines. Bottazzi et Peri (2002) proposent de modéliser ces différents paramètres à l'aide d'une fonction de production d'innovation. Ils caractérisent la solution d'équilibre (ou sentier de croissance équilibré) d'un modèle de croissance endogène par une fonction de production d'innovation. Cette fonction relie le nombre de brevets d'une région représentative i à l'investissement en R&D des régions avoisinantes, en fonction de leur distance à i . Leurs résultats indiquent une diffusion significative de la R&D jusqu'à 300 kilomètres de la région source.

1.2. Définitions et mesures de la proximité technologique

En utilisant la base de données Regio d'Eurostat, nous analysons 335 régions appartenant à 9 pays européens. Le découpage administratif des régions variant entre pays, nous avons choisi une nomenclature qui assure une homogénéité entre la taille de toutes les régions européennes. Ainsi, la nomenclature NUTS 3 (Nomenclature des Unités Territoriales Statistiques) est utilisée pour la France, la Grande-Bretagne, l'Italie, le Danemark et l'Irlande, NUTS 2 pour l'est de l'Allemagne, la

Belgique, les Pays-Bas et NUTS 1 pour le Luxembourg. La Figure 1 représente les observations des taux de croissance des dépôts de brevets pour l'année 1996.

1. Logarithme du nombre de dépôts de brevets pour 100 000 habitants pour l'année 1996



Les données sur les investissements en R&D n'existant pas au niveau départemental (NUTS 3), nous prendrons une division administrative supérieure pour les variables relatives à la R&D (tableau 1).

1. Nombre de régions administratives

Pays	Nombre de régions couvertes par toutes les variables excepté la variable de R&D	Nombre de régions couvertes par la variable de R&D
Belgique	11	3
Danemark	14	1
France	95	22
Pays-Bas	11	4
Grande-Bretagne	70	12
Irlande	8	1
Italie	95	20
Luxembourg	1	1
Allemagne ¹	30	11

1. Seule l'Allemagne de l'Ouest est prise en considération.

Source : Calculs de l'auteur.

La figure 1 représente les différents niveaux des activités d'innovation en Europe pour l'année 1996. De nombreuses méthodes ont été avancées pour distinguer les effets de la proximité géographique des effets de la proximité technologique sur les phénomènes de localisation de l'activité innovante.

Notons que le brevet est de nouveau utilisé comme indicateur d'innovation (Anselin *et al.* 1997; Hall *et al.* 2001; Jaffe 1989), même si toutes les innovations ne sont pas brevetées et que tous les brevets ne représentent pas des innovations utiles pour l'activité économique. Néanmoins, cette étude repose sur l'agrégation des dépôts de brevets en tant qu'indicateur régional de l'activité innovante. Jaffe *et al.* (1993) utilisent les citations de brevets pour analyser les phénomènes de localisation de l'innovation. Ils relèvent une forte inégalité géographique pour l'intensité d'innovation et mettent également en évidence une diffusion relativement lente de ces innovations entre régions. Maurseth et Verspagen (2002) étudient également les externalités technologiques à travers les citations de brevets. Jaffe *et al.* (1993) soulignent cependant les limites de ces approches par citations. L'obligation de citer tous les brevets précédant la nouvelle invention ne nécessite pas pour autant leur connaissance de la part de l'innovateur. L'estimation des effets de la diffusion de connaissance sur la production d'innovation à travers ces indicateurs peut donc parfois être biaisée.

Jaffe (1986) analyse la proximité technologique à partir de la distribution des brevets déposés dans différents secteurs technologiques pour un échantillon d'entreprises américaines. En distinguant 49 champs technologiques, il mesure cette proximité entre deux firmes en calculant le coefficient de corrélation entre les vecteurs représentant le nombre de brevets par domaine technologique. Cette mesure est alors utilisée pour distinguer le rôle de la proximité technologique par rapport à celui de la proximité spatiale.

En reprenant cette mesure, nous adoptons la classification internationale des brevets (IPC) distinguant 8 champs technologiques, ayant pour indice k . Pour chacune des 335 régions, nous classons les dépôts de brevets en fonction de leur domaine d'application. Soient F_{ki} et F_{kj} , la proportion de brevets respectivement de la région i et de la région j appartenant au domaine technologique k . L'indice de proximité technologique de Jaffe entre ces deux régions s'écrit alors :

$$S_{ij} = \frac{\sum_{k=1}^m F_{ki} F_{kj}}{\sqrt{\sum_{k=1}^m F_{ki}^2 \sum_{k=1}^m F_{kj}^2}}. \quad (1)$$

La valeur de cette mesure est d'autant plus forte que ces deux régions déposent des brevets dans les mêmes secteurs. Nous considérons que les effets d'externalité technologique sont limités aux

régions contiguës. Une des particularités de l'indice (1) consiste en la symétrie des interactions entre régions. Les externalités S_{ij} émises par i vers j sont égales aux externalités S_{ij} perçues par i et provenant de j . Nous allons maintenant proposer une nouvelle mesure de proximité en différenciant, dans le processus de diffusion des externalités, celles qui sont émises de celles qui sont perçues.

1.3 Définitions et mesures des coûts de transport

L'évolution des disparités de développement économique entre les différentes régions européennes repose également sur de nombreux facteurs autres que la proximité technologique. Ainsi de nombreuses études ont montré qu'en Europe, les coûts de transport affectent les localisations de l'activité (Parent et Riou, 2005). D'un point de vue théorique, la baisse des coûts de transport réduit l'intérêt pour une firme à se localiser en périphérie, puisque l'accès à la demande sera facilité par ces faibles coûts. Ainsi, la dynamique d'agglomération est renforcée par la réduction des coûts de transport. Nous allons ainsi spécifier une mesure de qualité des infrastructures de communication, calculée à partir du temps de transport par autoroute pour relier les principales villes administrative¹. À partir d'une étude plus approfondie, Combes et Lafourcade (2005) évaluent les coûts de transport au moyen d'une fonction de la distance et du temps de transport construite à partir du réseau routier, du prix de l'essence, des péages autoroutiers et du salaire des chauffeurs. Ils montrent qu'une baisse du coût de transport routier est synonyme d'un renforcement de la spécialisation et des disparités régionales. Dans cette étude, Nous supposons que les externalités de connaissance décroissent simplement avec la durée de transport entre les villes :

$$S_{ij} = \frac{1}{TT_{ij}}, \quad (2)$$

TT_{ij} représente le temps de transport entre les deux villes principales des régions i et j .

Afin de mieux appréhender les phénomènes d'externalité de connaissance, nous allons également introduire une spécification de référence qui ne tiendra compte que de la proximité géographique. Les interactions entre deux régions i et j seront donc mesurées par :

$$S_{ij} = \frac{1}{Dist_{ij}}, \quad (3)$$

où
$$Dist_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

1. Le logiciel Microsoft Autoroute nous permet de calculer non pas la distance la plus courte entre deux villes mais le parcours le plus rapide.

représente la distance calculée à partir de la latitude et de la longitude des principales villes administratives des régions i et j .

Par la suite, le modèle (1) comportera la matrice de poids mesurant la proximité technologique, dans le modèle (2) cette matrice ne tiendra compte que du temps de transport alors que dans le modèle (3) elle ne sera basée que sur la proximité géographique.

2. Dépendance spatiale et hétérogénéité

L'estimation de modèles à la fois robustes à la dépendance spatiale et aux observations extrêmes rend préférable l'utilisation de l'approche bayésienne hiérarchique (Parent et LeSage, 2007). L'introduction d'une loi de Student nécessite un niveau de hiérarchie supplémentaire afin de contrôler le nombre de degrés de liberté. Nous allons décrire un modèle général qui sera par la suite comparé à d'autres modèles.

2.1. Structure générale du modèle

De nombreuses études empiriques relatives à la géographie de l'innovation considèrent la formation d'innovation comme le produit d'une fonction de production de connaissance (Griliches, 1979) dont les facteurs de production sont généralement caractérisés par la part du capital et du travail allouée à l'activité de recherche et développement.

Pour chaque région $i = (1, \dots, n)$, le nombre de brevets déposés est assimilé à la quantité de connaissance produite (Jaffe, 1989; Hall et al., 2001; Anselin et al., 1997). Nous proposons une fonction de production de type Cobb-Douglas qui relie la production de connaissance aux facteurs de production mesurés par deux ensembles de variables. Le premier regroupe les investissements en $R\&D$ et le nombre de personnes travaillant dans ce domaine, alors que le second représente un indicateur de décomposition sectorielle, capturant la diversité ou la spécialisation des secteurs innovants dans chaque région. L'ensemble des facteurs de production est donc représenté par les trois variables suivantes : le pourcentage des dépenses en $R\&D$ dans le secteur privé (EXP), le nombre d'employés dans le domaine de la $R\&D$ (équivalent temps plein) dans le secteur privé ($NPriv$) et la variable de diversité sectorielle. Cette dernière variable nous permet de tester dans quelle mesure la structure de l'industrie influence la production d'innovations. Elle est construite à partir du rapport entre le nombre de brevets des quatre secteurs ayant le plus grand nombre de brevets déposés, et le nombre total de brevets déposés pour chaque région.

Rappelons que le manque de données concernant les mesures de $R\&D$ au niveau régional nous oblige à construire les variables EXP et

$NPriv$ à partir de divisions administratives plus larges (voir tableau 1). Les variables explicatives prises sous forme logarithmique constituent la matrice \mathbf{X} du modèle à estimer. La variable dépendante y correspond au logarithme du nombre de brevets déposés pour 100 000 habitants. Il semble approprié d'introduire des effets non observables dans l'explication des phénomènes de localisation de l'activité innovante. Le fait de définir ces coefficients en fonction de leur voisinage améliore significativement la qualité du modèle.

Nous pouvons réécrire la fonction de production de connaissance sous la forme :

$$\mathbf{y} = \alpha + \mathbf{X}\beta + \phi + \nu \quad (4)$$

où α est la constante du modèle et β , le vecteur des effets des variables explicatives à estimer. Rappelons que β représente le vecteur des effets fixes. Le vecteur ν correspond au vecteur d'hétérogénéité non observable. Nous prendrons en compte les observations extrêmes en spécifiant ce vecteur de paramètres à partir d'une loi de Student. Le vecteur ϕ représente les effets spatiaux non observables. Nous supposons ainsi que pour chaque région, la fonction de production de connaissance dépend non seulement de ses propres efforts en termes de recherche et développement mais aussi de connaissances issues d'autres régions. Les externalités relatives au capital humain, à l'économie d'agglomération ou encore aux réseaux peuvent être internes à chaque région. Avec la spécification des effets spatiaux que nous proposons, tous ces facteurs non observables peuvent également provenir des régions avoisinantes. Cette approche se distingue des études traditionnelles mesurant la dépendance spatiale au niveau de la variable dépendante ou au niveau des perturbations (Anselin *et al.*, 1997). Cette spécification des effets spatiaux est donc basée sur la différence $y - \alpha - \mathbf{X}\beta$ entre la production de connaissance et les inputs observés de chaque région. L'intensité de ses effets mesure ainsi les spillovers spatiaux de la fonction de production de connaissance.

2.2. Spécification des effets spatiaux

Afin de modéliser les interactions spatiales, il est nécessaire de spécifier les liens spatiaux qui existent entre chaque région. L'ensemble de ces liens est alors résumé dans une matrice que l'on appelle matrice de pondération spatiale. Cette matrice \mathcal{W} est une matrice carrée dont la dimension correspond au nombre de régions étudiées et où chaque élément w_{ij} représente la façon dont la région i et la région j sont connectées spatialement. Différentes structures seront ainsi proposées à partir des indices de proximité technologique et des mesures des temps de transport entre régions. Nous considérons que la région i appartient au voisinage δ_j de j si et seulement si ces deux régions sont contiguës.

Les modèles que nous allons estimer se distinguent alors uniquement par leur matrice de pondération spatiale. On a alors respectivement :

$$(1) \begin{cases} w_{ij} = \frac{\sum_{k=1}^m F_{ki} F_{kj}}{\sqrt{\sum_{k=1}^m F_{ki}^2 \sum_{k=1}^m F_{kj}^2}}, \forall i \in \delta_j \\ w_{ij} = 0 \text{ sinon} \end{cases}$$

$$(2) \begin{cases} w_{ij} = \frac{1}{T_{ij}}, \forall i \in \delta_j \\ w_{ij} = 0 \text{ sinon} \end{cases}$$

$$(3) \begin{cases} w_{ij} = \frac{1}{Dist_{ij}}, \forall i \in \delta_j \\ w_{ij} = 0 \text{ sinon} \end{cases}$$

À partir de la structure de connectivité que nous venons d'établir, la loi jointe des effets spatiaux est définie pour chacun des modèles comme étant :

$$\phi \sim N(0, \sigma_\phi^2 B_\rho^{-1} B_\rho^{-1}), \quad (5)$$

où $B_\rho = (I_n - \rho W)$ et le paramètre ρ mesure l'intensité de la dépendance spatiale.

Cette spécification originale des effets spatiaux nous permet donc d'introduire les informations relatives à la proximité technologique et à la proximité spatiale en utilisant une matrice de poids dont les coefficients sont nuls si les régions ne sont pas contiguës.

2.3. Loi de Student et observations extrêmes

Les lois de Student sont utilisées depuis de nombreuses années afin d'estimer des modèles robustes aux observations extrêmes (e.g., Lange *et al.*, 1989). Ces observations aberrantes ont une forte influence sur l'estimation des paramètres d'intérêt, surtout lorsque les observations sont spatialement liées. Des effets artificiels de dépendance peuvent apparaître en certaines localisations et « contaminer » l'estimation des effets spatiaux sur plusieurs régions. Cette approche a été introduite en économétrie spatiale par LeSage (1997) qui spécifie la distribution des termes d'erreur comme étant :

$$v \sim N(0, \sigma_v^2 \mathbf{V}), \quad \text{où } \mathbf{V} = \text{diag}(v_1, v_2, \dots, v_k) \quad (6)$$

représente une matrice de n variances v_i . Intuitivement, un problème d'identification apparaît pour estimer cette matrice \mathbf{V} comprenant n paramètres non-contraints. Les n paramètres de variance v_i sont alors supposés indépendamment et identiquement distribués selon une loi Gamma inverse :

$$p(\mathbf{V}) = \prod_{i=1}^n Ga^{-1}(v_i | \frac{\lambda}{2}, \frac{\lambda}{2}). \quad (7)$$

Il ne reste plus que le seul paramètre λ à estimer, résolvant ainsi le problème d'identification. Geweke (1993) montre que cette approche revient à spécifier une loi de Student pour les termes d'erreur, où λ représente le nombre de degrés de liberté.

Cette étude repose sur des méthodes de simulation de Monte Carlo par chaînes de Markov, appelées communément MCMC, et à leur application à des problèmes de dépendance spatiale. Une idée fondamentale de cette approche est d'utiliser l'information *a priori* disponible, en modélisant cette connaissance par des distributions de probabilités *ad hoc*. La méthode utilisée est développée en annexe.

3. Résultats empiriques

Nous commençons cette section en utilisant la méthode de Chib et Jeliazkov pour comparer les trois modèles décrits précédemment. Nous générons une chaîne de 20 000 itérations dont les 5 000 premières sont supprimées pour estimer les paramètres². La convergence des estimations est contrôlée à l'aide de différents indicateurs.

3.1. Estimation des critères de comparaison

Les estimations de la vraisemblance marginale apparaissent dans la dernière ligne du tableau 2. Ces résultats révèlent tout d'abord que les informations relatives à la proximité technologique améliorent la qualité du modèle. En effet, la valeur du critère pour le modèle 1 (-512.314) est supérieure à celle du modèle 2 (-517.316) et du modèle 3 (-521.566). Or la seule différence entre ces deux modèles repose sur la spécification des effets spatiaux : pour le modèle 1, la matrice de poids est construite à partir de l'indice de proximité technologique, alors que pour le modèle 2, cette matrice ne tient compte que de la proximité spatiale à travers le temps de transport.

2. Tous les calculs sont effectués à partir d'un PC Pentium-IV, 2.3 GHz, en utilisant le logiciel Matlab.

2. Lois *a posteriori* des trois modèles

	Modèle 1 (Techno)	Modèle 2 (TT)	Modèle 3 (Géo)
<i>Constante</i>	-3.566 (0.247)	-3.547 (0.248)	-3.594 (0.245)
<i>EXP</i>	0.529 (0.082)	0.542 (0.083)	0.543 (0.083)
<i>NPriv</i>	-0.028 (0.088)	-0.033 (0.088)	-0.031 (0.087)
<i>Div</i>	-4.246 (0.817)	-4.366 (0.773)	-4.287 (0.789)
σ_v	0.198 (0.078)	0.202 (0.041)	0.204 (0.056)
σ_ϕ	1.002 (0.412)	1.032 (0.476)	1.015 (0.456)
ρ	0.096 (0.040)	0.060 (0.027)	0.178 (0.35)
λ	18.848 * (10.366)	30.4 * (16.184)	25.8 * (14.451)
$m(y M)$	-512.314	-517.316	-521.566

* Erreur numérique *a posteriori* (« Numerical Standard Error ») supérieure à 1 %.
Source : Estimations de l'auteur.

3.2. Résultats des estimations

Le tableau 2 présente les estimations du modèle 1 et du modèle 2. Les estimations des effets des variables explicatives révèlent que la variable de dépense en R&D dans le secteur privé, *EXP*, influence positivement l'activité d'innovation (+ 0.529 pour le modèle 1, + 0.542 pour le modèle 2 et + 0.543 pour le modèle 3), alors que les effets de la variable *NPriv* mesurant le nombre d'employés dans le domaine de la R&D ne sont pas significativement différents de zéro. Les effets de la variable de diversité *Div* mettent en évidence l'influence positive de la diversité des secteurs d'activité sur la production d'innovation. Ces externalités dites « de Jacobs » (1969) confirment les conclusions d'Autretsch et Feldman (2004) : les externalités de connaissance se diffusent plus intensément dans les régions à technologies diversifiées que dans les régions spécialisées.

3.3. Effets spatiaux

Les résultats des estimations des paramètres mesurant l'intensité de la dépendance spatiale ρ sont également représentés dans le tableau 2.

Ces estimations révèlent que la relation de dépendance entre les régions est fortement positive. En effet, pour le modèle 1, l'inverse de la plus grande valeur propre de la matrice W est égale à $\psi_{max}^{-1} = 0.14$, alors que l'estimation du paramètre ρ est de 0.096. Ce résultat est sensiblement identique pour les modèles 2 et 3 où $\psi_{max}^{-1} = 0.11$ (resp. $\psi_{max}^{-1} = 0.11$) et les estimations du paramètre ρ sont de 0.060 (resp. 0.178). Remarquons également que l'intensité de la dépendance spatiale est relativement plus forte dans le cas de la proximité technologique.

La figure 2 représente les estimations des effets spatiaux en Europe. Les résultats font apparaître des effets de dépendance positifs et négatifs entre régions contiguës. Avant d'analyser plus en profondeur ces résultats, il est important de rappeler que toutes les variables relatives à la R&D sont basées sur un niveau géographique supérieur (voir tableau 1). Ce manque de données implique une homogénéité des dépenses de R&D pour tous les départements au sein d'une même région NUTS2. Cette homogénéité peut donc se refléter au niveau des estimations des effets spatiaux localisés dans une même région NUTS2.

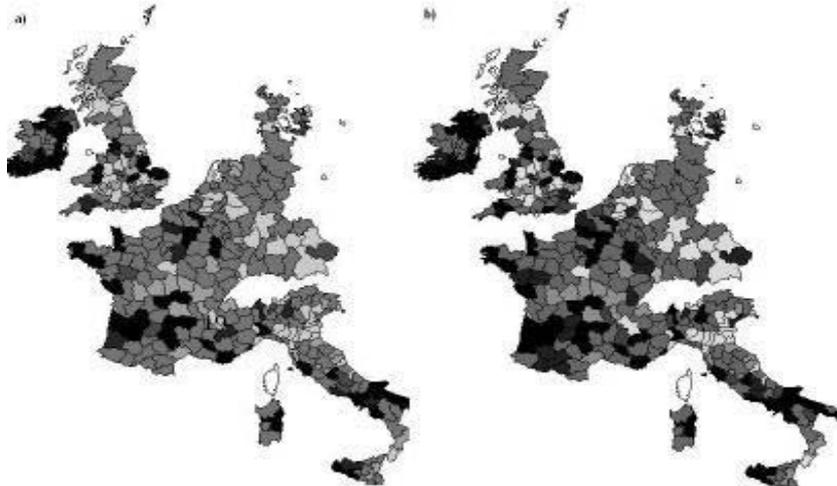
Les effets positifs se situent principalement dans le sud-ouest de l'Allemagne (province du Saarland, Unterfranken, Mittelfranken et dans la région de Karlsruhe), au nord de l'Italie (région de l'Émilie-Romagne), au sud de la région parisienne en France, autour de Londres en Angleterre (région du West-Midlands), les régions au sud de l'Écosse et les régions au nord des Pays-Bas (Gelderland). Les effets négatifs se localisent principalement dans le sud-ouest de la France, le sud de l'Italie et en Irlande. Notons que les estimations des effets spatiaux sont corrélées à plus de 85 % entre les trois spécifications que nous proposons. Cependant, nous pouvons distinguer les faits suivants.

La prise en compte de l'information sur les temps de transport montre que les effets de dépendance sont beaucoup plus polarisés. En effet, la figure 2 (b) montre que les externalités sont beaucoup plus prononcées que celles de la figure 2 (c). La mauvaise qualité des infrastructures renforce pour certaines régions les effets négatifs des externalités sur l'activité d'innovation. En effet, pour les régions isolées, les estimations des effets spatiaux entre les modèles (2) et (3) diminuent fortement pour les régions localisées dans l'Ouest et le Sud-Ouest de la France ou encore dans le Sud de l'Italie. Pour ces régions à faible activité innovante, les temps de transport avec leurs régions voisines sont relativement élevés. Il existe aussi de nombreuses régions où la prise en compte du temps de transport réduit la significativité des estimations. C'est principalement le cas en Irlande du Nord et en Écosse où la qualité des réseaux de communication est souvent défectueuse et la prise en compte de cette information détériore les effets locaux d'externalité.

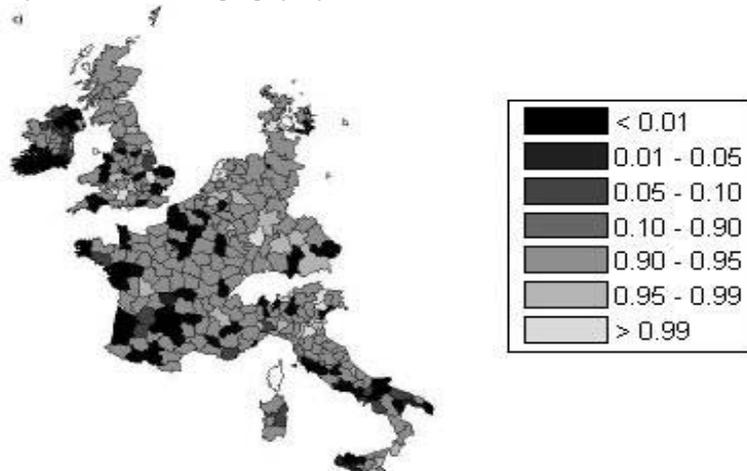
Nous avons donc mis en évidence l'importance de l'information apportée par la qualité des infrastructures de communication lorsque nous cherchons à estimer les externalités. En comparaison avec la spécification simple de dépendance basée uniquement sur une distance géographique, la prise en compte du temps de transport renforce les effets de dépendance entre régions. Cela implique que les externalités locales entre régions représentent un élément explicatif important du taux de croissance de l'activité d'innovation. Ce résultat suggère que l'amélioration des infrastructures de communication peut stimuler l'innovation en facilitant les externalités de connaissance.

2. Probabilité *a posteriori* d'observer un effet spatial positif ou négatif

(a) à partir de la proximité technologique (b) en tenant compte des temps de transport



(c) à partir de la distance géographique



Ces résultats suggèrent une forte polarisation de l'activité innovante en Europe. La figure 2 (a) représente également les estimations des effets spatiaux basés sur la proximité technologique entre les régions contiguës. Les régions doivent en effet posséder les technologies adéquates afin de profiter des connaissances venant des régions voisines. Ainsi, les effets de dépendance prédominants sont localisés dans les régions à forte activité économique telles que le sud de l'Allemagne, les Pays-Bas ou encore le sud de l'Angleterre. De plus, la proximité technologique permet de mettre en évidence des effets d'externalité qui n'apparaissent pas lorsque les effets spatiaux ne tenaient compte que du temps de transport (région parisienne). Ainsi, dans le centre de l'Angleterre ou encore le nord de l'Italie, les activités innovantes entre régions ont lieu dans des domaines d'activité similaires et présentent alors de fortes dépendances spatiales.

Ces résultats confirment ainsi l'existence de centres de polarisation de l'activité innovante en Europe. Cette concentration d'activité est en partie stimulée par la proximité géographique mais aussi technologique.

4. Conclusion

La modélisation des externalités de connaissance à partir de l'information relative à la proximité technologique et à l'écart de productivité nous a permis de mettre en évidence de nouvelles zones de polarisation de l'activité innovante en Europe. Pour bénéficier des phénomènes de diffusion, il est nécessaire que les régions possèdent un certain niveau de capacité d'absorption de connaissance.

La modélisation bayésienne hiérarchique est introduite afin d'estimer simultanément les effets spatiaux et les effets d'hétérogénéité. L'utilisation d'une loi de Student pour modéliser ces effets non observables indépendants nous a permis de tenir compte des observations extrêmes en contrôlant l'épaisseur des queues de distribution. L'utilisation des méthodes de Chib et Jeliazkov pour comparer différentes spécifications a alors montré que les résidus possèdent des queues de distribution plus épaisses que ce que donnerait l'hypothèse de normalité. Ces résultats confirment ainsi l'importance de ne pas spécifier systématiquement les termes d'erreurs par des lois normales lorsque nous estimons des modèles spatiaux.

Les estimations des effets spatiaux révèlent que les externalités de connaissance sont fortement concentrées en Europe, dans quelques régions se situant principalement dans la « banane bleue » qui s'étend de Londres à Milan (Lever, 1999).

Références bibliographiques

- ANSELIN L., A. VARGA et Z. ACS, 1997 : « Local geographic spillovers between university research and high technology innovations », *Journal of Urban Economics*, 42, 422-48.
- AUDRETSCH D. B. et M. P. FELDMAN, 1996 : « Knowledge spillovers and the geography of innovation and production », *American Economic Review*, 86, 630-640.
- AUDRETSCH D. B. et M. P. FELDMAN, 2004 : « Knowledge Spillovers and Geography of Innovation », dans V. Henderson et J.F. Thisse (Ed.), *Handbook of Urban and Regional Economics*, Elsevier, Amsterdam, Vol. 4, pp. 2713-2739.
- AUTANT-BERNARD C., 2001 : « The Geography of knowledge spillovers and technological proximity », *Economics of Innovation and New Technology*, 10, 237-254.
- BALDWIN R.E. et R. FORSLID, 2000 : « The core-periphery model and endogenous growth : stabilizing and de-stabilizing integration », *Economica*, 67, 307-324.
- BOTTAZZI L. et G. PERI, 2002 : « Innovation and spillovers : evidence from European regions », *European Economic Review*, 47, 687-710.
- BRÜLHART M. et J. TORSTENSSON, 1996 : « Regional integration, scale economies and industry location in the European Union », *CEPR Discussion Paper*, n° 1435.
- COMBES P.P. et M. LAFOURCADE, 2005 : « Transport costs: measures, determinants and regional policy implications for France », *Journal of Economic Geography*, 5, 319-349.
- CHIB S., 1995 : « Marginal likelihood from the Gibbs output », *Journal of the American Statistical Association*, 90, 1313-1321.
- CHIB S. et I. JELIAZKOV, 2001 : « Marginal Likelihood from the Metropolis-Hastings Output », *Journal of the American Statistical Association*, 96, 270-281.
- CRESSIE N., H. S. STERN et D. R. WRIGHT, 2000 : « Mapping rates associated with polygons », *Journal of Geographical Systems*, 2, 61-69.
- EL OUARDIGHI J., 2002 : « Dépendance spatiale sur données de panel, application à la relation brevets-R&D au niveau régional », *Revue d'analyse économique*, 78, 67-86.
- ENGLMANN F. C. et U. WALZ, 1995 : « Industrials centers and regional growth in the presence of local inputs », *Journal of Regional Science*, 35, 3-27.

- ERTUR C. et W. KOCH, 2007, « Growth, technological interdependence and spatial. externalities: theory and evidence », *Journal of Applied Econometrics*, forthcoming.
- FELDMAN M. P., 1994 : *The Geography of Innovation*, Kluwer Academic Publishers.
- GEWEKE J., 1993 : « Bayesian Treatment of the Independent Student-t Linear Model », *Journal of Applied Econometrics*, 8, 19-40.
- GLAESER E.L., H. G. KALLAL, J.A. SHEINKMAN et A. SHLEIFER, 1992 : « Growth in Cities », *Journal of Political Economy*, 100, 1126-1151.
- GRILLICHES Z., 1979 : « Issues in Assessing the Contribution of Research and Development to Productivity Growth », *Bell Journal of Economics*, 10, 92-116.
- HAN C. et B. CARLIN, 2001 : « MCMC Methods for Computing Bayes Factors: a Comparative Review ». *Journal of the American Statistical Association*, 96, 1122-1132.
- HALL B., J. HAUSMAN et Z. GRILLICHES, 1984 : « Patents and R&D: searching for a lag structure », *NBER Working Paper*.
- JACOBS J., 1969 : *The Economy of Cities*, (New-York: Vintage).
- JAFFE A. B., 1986 : « Technological opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value », *American Economic Review*, 76, 984-1001.
- JAFFE A. B., 1989 : « Real Effects of Academic Research », *American Economic Review*, 79, 957-970.
- JAFFE A. B., M. TRAJTENBERG et R. HENDERSON, 1993 : « Geographic localization of knowledge spillovers as evidenced by patent citations », *Quarterly Journal of Economics*, 108, 576-598.
- KOOP G., 2003 : *Bayesian Econometrics*. John Wiley & Sons.
- LANGE K. L., R. J. A. LITTLE et J.M. G. Taylor : 1989, « Robust statistical modeling using the t-distribution », *Journal of the American Statistical Association*, 84, 881-896.
- LE GALLO J., 2004 : « Space-time analysis of GDP disparities across European regions : a Markov chains approach », *International Regional Science Review*, 27, 138-163.
- LE GALLO J. et C. ERTUR, 2003 : « Exploratory Spatial Data Analysis of the distribution of regional per capita GDP in Europe, 1980-1995 », *Papers in Regional Science*, 33, 343-370.
- LESAGE J. P., 1997 : « Bayesian Estimation of Spatial Autoregressive Models », *International Regional Science Review*, 20, 113-129.

■ Olivier Parent

- LESAGE J. P., 2004 : « Spatial Regression Models », in *Numerical Issues in Statistical Computing for the Social Scientist*, John Wiley & Sons, Inc., Micah Altman, Jeff Gill and Michael McDonald (eds.), pp. 199-218.
- LEVER W. F., 1999 : « Competitive cities in Europe », *Urban Studies*, 36, 1029-1044.
- MARSHALL A., 1890 : *Principles of Economics*. London: Macmillan.
- MARTIN P. et G. I. P. OTTAVIANO, 1999 : « Growing Locations: Industry Location in a Model of Endogenous Growth », *European Economic Review*, 43, 281-302.
- MAURSETH P. B. et B. VERSPAGEN, 2002 : « Knowledge spillovers in Europe. A patent citations analysis », *Scandinavian Journal of Economics*, 104, 531-545.
- PACI R. et S. USAI, 2000 : « The role of specialisation and diversity externalities in the agglomeration of innovative activities », *Rivista Italiana degli Economisti*, 5, 237-268.
- PARENT O. et J. P. LESAGE, 2007 : « Using the Variance Structure of the Conditional Autoregressive Spatial Specification to Model Knowledge Spillovers », *Journal of Applied Econometrics*, forthcoming.
- PARENT O. et S. RIOU, 2005 : « Bayesian analysis of knowledge spillovers in European regions », *Journal of Regional Science*, 45, 747-775.
- PUGA D., 2002 : « European regional policy in light of recent location theories », *Journal of Economic Geography*, 2, 372-406.
- SAXENIAN A., 1994 : *Regional advantage. Culture and Competition in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.
- VAN POTTELSBERGHE DE LA POTTERIE B. et F. LICHTENBERG, 2001 : « Does foreign direct investment transfer technology across borders? », *the Review of Economics and Statistics*, 83, 490-497.
- VERSPAGEN B. et W. SCHOENMAKERS, 2000 : « The spatial dimension of knowledge spillovers in Europe: evidence from firms patenting data », *ECIS Working Papers 00.07*, Eindhoven Centre for Innovation Studies, Eindhoven University of Technology.

ANNEXE

Estimation MCMC

L'estimation bayésienne de ce modèle requiert la spécification de la loi jointe *a posteriori* des paramètres suivants : les effets fixes β , le paramètre ρ mesurant l'intensité de la dépendance spatiale, les effets spatiaux ϕ ainsi que leur variance σ_ϕ^2 , et les variances des erreurs σ_v^2 et V . Cressie *et al.* (2000) ont prouvé que la loi jointe *a posteriori* de ce modèle est propre, mais en restreignant le terme d'erreur à un bruit blanc gaussien. Cette propriété de propreté de la loi jointe *a posteriori* est respectée pour l'extension que nous proposons puisque nous ne spécifions que des lois propres pour le mélange gaussien de termes d'erreur. La méthode d'échantillonnage de Gibbs n'est applicable que lorsque les lois *a posteriori* correspondent à des lois connues et facilement simulables. Ce modèle étant plus complexe, l'introduction des algorithmes de Metropolis-Hastings est nécessaire pour la simulation de certaines lois *a posteriori*.

Avant d'analyser les méthodes d'estimation, nous allons définir les lois *a posteriori* à partir d'un modèle hiérarchique distinguant trois niveaux. Dans le premier niveau (8), on suppose que les observations sont distribuées selon une loi de Student (ou mélange de lois normales) dont la moyenne est caractérisée par les effets fixes (effets des variables explicatives) et les coefficients aléatoires non observables. Ces coefficients aléatoires sont spécifiés dans le deuxième niveau (9) et les lois *a priori* sont définies au troisième niveau (10) :

$$\mathbf{y} = \alpha + \mathbf{X}\beta + \phi + v, \quad (8)$$

$$\phi \sim N(0, \sigma_\phi^2 B_\rho^{-1} B_\phi^{-1}), \quad (9)$$

$$v \sim N(0, \sigma_v^2 \mathbf{V}), \quad \text{où } \mathbf{V} = \text{diag}(v_1, v_2, \dots, v_k)$$

$$\mathbf{V} \sim Ga^{-1}\left(\frac{\lambda_v}{2}, \frac{\lambda_v}{2}\right),$$

$$p(\beta) \sim N(f, T_k),$$

$$p(\alpha) \sim N(g, S),$$

$$p(\rho) \sim U(\psi_{min}^{-1}, \psi_{max}^{-1}), \quad (10)$$

$$p(\sigma_\phi^2) \sim Ga^{-1}(a, b),$$

$$p(\sigma_v^2) \sim Ga^{-1}(c, d),$$

$$p(\lambda) \sim Exp(\lambda_0),$$

Nous en déduisons le noyau de la distribution jointe :

$$p(\beta, \alpha, \rho, \phi, v, \sigma_\phi^2, \sigma_v^2, \mathbf{V}, \lambda | \mathbf{y}) \propto p(\mathbf{y} | \beta, \alpha, \phi, \sigma_v^2, \mathbf{V}) \quad (11)$$

$$p(\phi | \rho, \sigma_\phi^2) p(\mathbf{V} | \lambda) p(\lambda) p(\sigma_v^2) p(\sigma_\phi^2) p(\beta) p(\alpha) p(\rho),$$

où l'expression $p(\mathbf{y} | \alpha, \beta, \phi, \sigma_v^2, \mathbf{V})$ représente la vraisemblance qui, d'après le premier niveau (8), correspond à :

$$p(\mathbf{y} | \alpha, \beta, \phi, \sigma_v^2, \mathbf{V}) \propto \sigma_v^{-n} \prod_{i=1}^n v_i^{-1/2} \exp \left[-\sum_{i=1}^n \frac{(y_i - \alpha - x_i' \beta - \phi_i)^2}{2\sigma_v^2 v_i} \right]. \quad (12)$$

L'évaluation analytique des estimations des paramètres d'intérêt à partir de la loi jointe (11) nécessiterait le calcul d'une intégrale dont la dimension est trop élevée. Les techniques de simulation MCMC sont alors requises pour effectuer des tirages à partir de cette loi jointe. Ces techniques reposent sur l'évaluation des lois *a posteriori* suivantes (voir Parent et LeSage, 2007):

1. $p(\beta | \mathbf{y}, \alpha, \phi, \mathbf{V}, \sigma_v^2),$
2. $p(\alpha | \mathbf{y}, \phi, \mathbf{V}, \sigma_v^2, \beta),$
3. $p(\sigma_v^2 | \mathbf{y}, \phi, \mathbf{V}, \beta, \alpha),$
4. $p(v_i | \mathbf{y}, v_{-i}, \sigma_v^2, \beta, \alpha, \phi, \lambda),$
5. $p(\phi_i | \mathbf{y}, \phi_{-i}, \rho, \sigma_\phi^2),$
6. $p(\rho | \sigma_\phi^2, \phi),$
7. $p(\sigma_\phi^2 | \phi, \rho),$
8. $p(\lambda | \mathbf{V}).$

Les effets spatiaux sont construits à partir des différents indices de proximité précédemment établis et les termes d'erreur suivent le mélange de lois normales (7) robuste aux observations extrêmes. Nous introduisons dans ce modèle une loi *a priori* pour le nombre de degrés de liberté λ . Geweke (1993) propose de spécifier une loi exponentielle :

$$p(\lambda) = G(1, \lambda_0) \equiv \text{Exp}(\lambda_0). \quad (13)$$

Koop (2004) en déduit la loi *a posteriori* suivante pour le paramètre λ :

$$p(\lambda | \mathbf{V}) \propto \mathbf{p}(\mathbf{V} | \lambda) \mathbf{p}(\lambda) \\ \propto (\lambda/2)^{N\lambda/2} (\lambda/2)^{-N} \exp \left\{ \lambda_0 + \frac{1}{2} \sum_{i=1}^N [\ln(\mathbf{v}_i) + \mathbf{v}_i^{-1}] \right\}. \quad (14)$$

Une étape de Metropolis-Hasting est utilisée pour générer des échantillonnages à partir de cette loi *a posteriori* $p(\lambda | \mathbf{V})$.

La principale méthode bayésienne de comparaison de modèles est fondée sur le facteur de Bayes (BF, de l'acronyme anglo-saxon « Bayes Factor »), équivalent au rapport de la vraisemblance marginale des modèles à comparer. De nombreuses études récentes ont permis l'évaluation de ces vraisemblances marginales, qui restent cependant difficiles à estimer pour les modèles hiérarchiques. Han et Carlin (2001) ont

récemment proposé une étude exhaustive des différentes techniques d'estimation du facteur de Bayes. La plus connue, et la plus simple, a été proposée par Chib (1995) : elle ne requiert que l'utilisation de l'échantillonnage de Gibbs. Cependant nous allons voir que cette technique, pour évaluer la vraisemblance marginale, nécessite le calcul de la forme exacte de la vraisemblance des observations, des lois *a priori* et des lois *a posteriori*. Chib et Jeliazkov (2001) ont généralisé cette méthode aux cas où les lois *a posteriori* ne sont pas connues, en introduisant des algorithmes de Metropolis-Hastings.

Après avoir défini les deux premiers niveaux du modèle hiérarchique, nous allons spécifier les lois *a priori* correspondant au troisième niveau.

En l'absence d'information *a priori*, nous adoptons des lois relativement diffuses pour les effets β des variables explicatives et la constante α avec pour paramètres $f = g = 0$ et $T = S = 10^5$. Nous spécifions également des lois Gamma diffuses dont les valeurs des paramètres correspondent à $a=c=.5$ et $b=d=.0005$. L'intensité de la dépendance spatiale ρ étant *a priori* inconnue, nous choisissons une loi uniforme définie sur l'intervalle $(\psi_{min}^{-1}, \psi_{max}^{-1})$, où $\psi_{max} < 0$ et $\psi_{min} > 0$ sont respectivement la plus faible et la plus forte valeur propre de la matrice de poids W . Pour le modèle 1, nous avons vu au paragraphe précédent que le nombre de degrés de liberté suit une loi exponentielle de paramètre λ_0 . Nous choisissons $\lambda_0=0.25$ accordant un poids *a priori* sensiblement équivalent entre les queues fines et les queues épaisses.

